PHILIPS TECHNICAL REVIEW

2011-11

N. V. PHILIPS' GLOEILAMPENFABRIEKEN

Bibl. Nat. Lab.

VOLUME 15

1953-1954

PHILIPS RESEARCH LABORATORIES - EINDHOVEN

CONTENTS

VOLUME 15, 1953/1954

LIGHT

	Page
The equivalent circuit of a gas discharge lamp, by C. Zwikker	161
Photometry at low luminance levels, by W. de Groot	182
The specific light output of "Photoflux" flash-bulbs, by L. H. Verbeek	317
Photography of the eye with the aid of electronic flash-tubes, by J. E. Winkelman and N. Warmoltz	342

TELECOMMUNICATIONS

A large-screen television projector, by J. Haantjes and C. J. van Loon	27
The troposphere as a medium for the propagation of radio waves, by H. Bremmer, I	148
II	175
A signal generator for the testing of television receivers, by B. W. van Ingen Schenau	205
The focusing of television picture-tubes with Ferroxdure magnets, by J. A. Verhoef	214
The flying-spot scanner by F. H. J. van der Poel and J. J. P. Valeton	221
A cathode-ray tube for flying-spot scanning, by A. Bril, J. de Gier and H. A. Klasens	233
A sinusoidal *RC*-oscillator for measurements in the frequency range 20-250 000 c/s, by J. D. Veegens and E. Prado	240
An improved ion-trap magnet, by W. F. Niklas	258
A cold cathode gas-discharge tube as a switching element in automatic telephony, by J. Domburg and W. Six	265
Radio-controlled models, by A. H. Bruinsma	281
A line converter for the international exchange of television programmes, by J. Haantjes and Th. G. Schut	297
An instrument for measuring group delay, by H. J. de Boer and A. van Weel	307

X-RAYS AND NUCLEAR PHYSICS

The 15 million electron-volt linear electron accelerator for Harwell, by C. F. Bareford and M. G. Kelliher	1
A. Choice of parameters and general design	2
B. Detailed description of main components and circuits	10
C. Testing and performance	22
The Ferroxcube transformer core used in the Brookhaven Cosmotron, by F. G. Brockman and M. W. Louwerse	73
A time-of-flight neutron spectrometer, by P. E. Trier, J. C. Hammerton and E. Wolfendale	325

SOUND

An electronic hearing aid, by P. Blom	37
The principle of the magnetic recording and reproduction of sound, by W. K. Westmijze	84

MEASURING EQUIPMENT

Three methods of measuring magnetic fields,
I. Measurement based on the generator principle, by B. F. Jürgens 49
II. Measurement of the field on the axis of magnetic electron lenses, by A. C. van Dorsten and A. J. J. Franken . . 52
III. Measurement by the proton resonance method, by H. G. Beljers 55

A pre-amplifier for use with electronic voltmeters and oscilloscopes, by F. G. Peuscher and J. van Holthoon 169

An automatic recording potentiometer for industrial use, by H. J. Roosdorp 189

A signal generator for the testing of television receivers, by B. W. van Ingen Schenau . . 205

A sinusoidal *RC*-oscillator for measurements in the frequency range 20-250 000 c/s, by J. D. Veegens and E. Prado 240

An instrument for measuring group delay, by H. J. de Boer and A. van Weel 307

MATERIALS AND THEIR RESEARCH

The efficiency of fluorescence in cathode-ray tubes, by A. Bril and H. A. Klasens 63

The Ferroxcube transformer core used in the Brookhaven Cosmotron, by F. G. Brockman and M. W. Louwerse 73

Metal-detectors, by E. Blasberg and A. de Groot 97

Imperfections in matter, by G. W. Rathenau 105

Preparation and casting of metals and alloys under high vacuum, by J. D. Fast, A. I. Luteijn and E. Overbosch 114

A remarkable etching of copper, by J. J. de Jong 238

Lattice imperfections and plastic deformation in metals, by H. G. van Bueren,
I. Nature and characteristics of lattice imperfections, notably dislocations . . 246
II. Behaviour of lattice imperfections during deformation 286

MISCELLANEOUS

Metal-detectors, by E. Blasberg and A. de Groot 97

Imperfections in matter, by G. W. Rathenau 105

The determination of droplet size in arc welding by high-speed cinematography, by P. C. van der Willigen and L. F. Defize 122

Modern casting techniques, by H. J. Meerkamp van Embden 133

The equivalent circuit of a gas discharge lamp, by C. Zwikker 161

An apparatus for protection against dangerous voltages in welding equipment, by F. H. de Jong and D. W. van Rheenen 199

Radio-controlled models, by A. H. Bruinsma 281

BOOK REVIEWS

C. Zwikker (Editor), Fluorescent lighting . . 129

P. J. Oranje, Gas discharge lamps 129

F. Kerkhof and W. Werner, Television . . . 146

A. G. W. Uitjens, Television receiver design, I: I.F. stages 147

G. D. Rieck and L. H. Verbeek, Artificial light and photography 188

N. S. Markus and J. Otte, Data and circuits of radio receiver and amplifier valves . . . 213

H. Zijl, Manual for the illuminating engineer on large size perfect diffusors 262

VOL. 15 No. 1, pp. 1-36 JULY 1953

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
THE PHILIPS INDUSTRIES

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, NETHERLANDS

THE 15 MILLION ELECTRON-VOLT LINEAR ELECTRON ACCELERATOR FOR HARWELL

by C. F. BAREFORD *) and M. G. KELLIHER **).

621.384.622.2

A. CHOICE OF PARAMETERS AND GENERAL DESIGN.

B. DETAILED DESCRIPTION OF MAIN COMPONENTS AND CIRCUITS.

C. TESTING AND PERFORMANCE.

This is the second article in the series on linear electron accelerator development published in this Review. Since the appearance of the first article, in which the basic principles of the travelling wave linear accelerator were described by D. W. Fry of A.E.R.E., Harwell, the present 15 MeV machine designed by the Mullard Research Laboratories has been installed at Harwell and is now in full time operation. Used in conjunction with a suitable neutron target, very intense bursts of neutrons are produced and the machine then provides the most intense pulsed neutron source available (so far as is known) anywhere in the world. The large variety of problems that had to be solved in this comparatively novel type of apparatus warranted a detailed treatment, to which more space has been devoted than is usual with the subjects dealt with in this Review.

Introduction

The linear electron accelerator is an instrument of considerable interest to the nuclear physicist. Apart from constituting a comparatively simple and efficient means of producing electrons of energy in the 10 million electron-volt range or higher, the linear accelerator, as was pointed out by Sir John Cockcroft [1]) in 1949, owing to the short duration and the large current in the electron pulse, is valuable as a means of generating short bursts of neutrons for use in "time-of-flight" neutron spectroscopy. This method has gained considerable importance for the investigation of neutron capture or scattering cross-sections of materials used in nuclear reactors.

The first two travelling wave type linear accelerators designed and built by Mullard Research Laboratories were destined mainly for nuclear physics work of this nature at the Atomic Energy Research Establishment at Harwell. The first of these was a 3.5 MeV accelerator similar to the one developed by D. W. Fry and co-workers and described in a previous article in this Review [2]). The second (*fig. 1*), whose installation at Harwell was completed during September 1952 and which is the subject of this article, was designed for an energy of 15 MeV.

Although this machine (as well as the smaller one) was destined for nuclear physics work, the design was influenced by the consideration that linear electron accelerators are likely to assume a very important role as X-ray sources in radio-

*) Formerly with Mullard Research Laboratories, Salfords, England; now Chief Superintendent, Long Range Weapons Establishment, Salisbury, Australia.
**) Mullard Research Laboratories, Salfords, Surrey, England.

[1]) J. D. Cockcroft, Nature **163**, 869, 1949.

[2]) D. W. Fry, Philips tech. Rev. **14**, 1-12, 1952 (No. 1). This article will further be quoted as I.

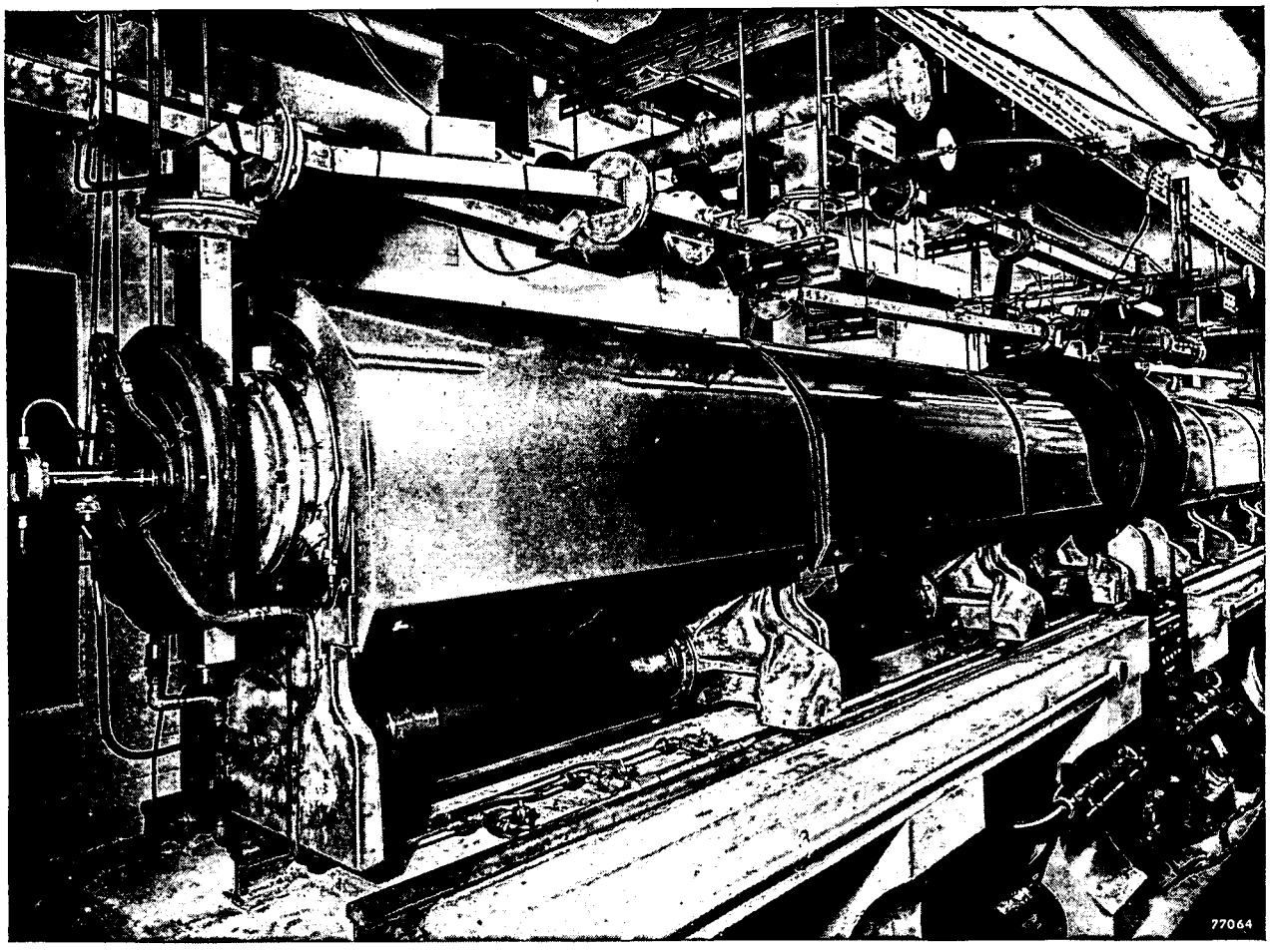


Fig. 1. View of the 15 MeV-accelerator, mounted in concrete shelter at the Atomic Energy Research Establishment, Harwell, as seen from the target end. (This photograph and that of fig. 13*b* are reproduced by permission of the Comptroller of H. M. Stationary Office.)

therapy. In fact, it seems that in the future the most widespread use of the linear accelerator may well be in clinical applications. It will be seen in this description that the 15 MeV accelerator was designed in such a way that it can be used without major modification either as a tool for the nuclear physicist or as a clinical instrument for the radiotherapist.

A. CHOICE OF PARAMETERS AND GENERAL DESIGN

Choice of operating parameters

General

The problem of the design of a linear accelerator for a given purpose depends on a large number of factors, and as in other problems of this nature the final design must inevitably be a compromise. While most of the factors which affect the design have been discussed in I, it will be desirable to recapitulate while discussing the choice of design parameters for the 15 MeV accelerator.

The preliminary data for the design were the following: the performance aimed at was an energy of 15 MeV with 25 mA beam current in the electron pulse; the length of the corrugated waveguide should not greatly exceed 6 m, as this was the space available in the concrete shelter at Harwell in which the machine was to be housed in order to protect personnel from the intense radiation produced; finally the most suitable power source available was a magnetron operating at 3000 Mc/s with 1.8 MW output in the pulse.

In principle a higher operating frequency would be an advantage. This is readily seen from the relationship (valid when the phase velocity is in the region of the velocity of light) between the maximum accelerating field E and the power flux W in the corrugated waveguide through which the

wave and the electrons travel (see I, eq. (2)):

$$E = (480\ W)^{\frac{1}{2}}\ \lambda/\pi a^2, \quad \ldots\ldots \quad (1)$$

where a is the inner radius of the irises in the corrugated guide. For a given value of a/λ, which is the main parameter of the guide, a higher frequency (shorter wavelength λ) would mean a higher accelerating field per unit power flux (i.e. a higher series impedance of the guide, see I). In the future suitable magnetrons for higher frequencies may become available but at the time when this equipment was designed the pulse powers available at higher frequencies, e.g. 100 kW at 10 000 Mc/s, were much too small.

It was pointed out in I that a choice has to be made between a long guide with low series impedance which is insensitive to frequency changes and a short high impedance guide which is relatively frequency sensitive. This situation is improved by the use of the feedback principle, since the power entering the guide is then higher than the magnetron power and therefore gives a higher energy, with the same impedance, in a given length. It may also be noted that the same electron energy may be achieved with the same length of guide but with a lower impedance (i.e. a less frequency sensitive guide). Even so it may well happen that the frequency sensitivity of the guide will be greater than is compatible with the given stability of the high frequency source. In this case it will be necessary to divide the given total length of guide into two or more independent systems fed from a common source, thus permitting correction of the accumulated phase error at the end of each system.

The decision whether the available length of guide should be subdivided in this way, is complicated by the problem of how much of the maximum accelerating field in the corrugated waveguide can be usefully employed. To emphasize the importance of this point it is necessary to enlarge upon the motion of wave and electron in the waveguide. If it were possible to obtain a high frequency source of very great stability (1 in 10^5) and if it were possible to machine and assemble the corrugated waveguide such that electron and wave velocities are substantially equal at all points of the electron path (machining tolerances $\pm 2\ \mu$ on some dimensions) then all electrons injected at the correct velocity and appropriate phase could travel in the wave crest, i.e. they could be accelerated in the *maximum* axial field. In practice magnetron frequency stability is at best of the order of 1 in 10^4, and it is not economic to machine corrugated waveguides to tolerances closer than $\pm 15\ \mu$. It is therefore clear that the electron must be made to travel in a region of increasing axial field such that changes in guide dimensions or frequency which cause the wave to accelerate more rapidly than the electron, do not cause the electron to move into a region of instability (i.e. of decreasing field *behind* the wave crest). With the guide impedance used up to the present time, it has been found possible to machine and assemble one metre lengths of waveguide such that the phase error of the position of the electrons after traversing one length will not exceed 15°. The accumulation of phase errors due to the mechanical inaccuracy can be avoided by inserting correcting sections of guide at the end of each metre length. Magnetron frequency changes of 1 in 10^4, i.e. 0.3 Mc/s, produce phase errors of the same order of magnitude, so that the total phase error in, for example, a 3 m length would be in the region of 30°. Current practice therefore is to make the electrons lead the wave crest by about 35°, i.e. to make the stable phase for the electrons $\Theta = 55°$. This means that the actual accelerating field on the electron

$$E_z = (480\ W)^{\frac{1}{2}}\ (\lambda/\pi a^2) \sin \Theta = (pW)^{\frac{1}{2}} \quad . \quad (2)$$

(using the abbreviation $p = 480(\lambda/\pi a^2)^2 \sin^2 \Theta$), will amount to about 80% of the maximum value[3]).

Calculation of performance

For a chosen value of the waveguide parameter a/λ, integration of E_z over the length L of the guide, taking into account the attenuation of power W along the guide, gives the total energy V gained by the electrons. If the beam loading of the accelerator is light, i.e. if the ratio of power transferred to the electron beam to power lost as heat in the walls of the corrugated waveguide is small, the power flux W at any point may be expressed in terms of W_e, the power entering the guide, by

$$W = W_e\ e^{-2\alpha z} \quad \ldots\ldots \quad (3)$$

where α is the voltage attenuation coefficient expressed in nepers/metre and z is the distance along the guide. Then,

$$V = \int_0^L E_z\, dz = \int_0^L (pW_e)^{\frac{1}{2}}\ e^{-\alpha z}\, dz,$$

[3]) p might be termed the "reduced impedance" of the waveguide: when $\Theta = 90°$, p is equal to the series impedance of the guide.

so that the energy gain per unit length is:

$$\frac{V}{L} = (pW_e)^{\frac{1}{2}} \frac{1-e^{-aL}}{aL}. \quad . \quad . \quad . \quad . \quad (4)$$

In most cases of practical interest, however, the beam loading cannot be neglected. In fact, the ratio of the power absorbed in the beam to that lost in the walls may be of the order of $^1/_4$. A recalculation of (3) and (4) taking into account the power absorbed by the electron beam (beam current I) gives the equations:

$$W^{\frac{1}{2}} = W_e^{\frac{1}{2}} e^{-az} - \frac{Ip^{\frac{1}{2}}}{2a}(1-e^{-az}) \quad . \quad . \quad (5)$$

and

$$\frac{V}{L} = (pW_e)^{\frac{1}{2}} \frac{1-e^{-aL}}{aL} - \frac{Ip}{2a}\left(1 - \frac{1-e^{-aL}}{aL}\right) \quad (6)$$

for the variation of the power flux along the guide and the energy gain per unit length, respectively.

Equations (5) and (6) can be derived in the following way: The power loss in a length dz of the guide is the sum of the power transferred to the electron beam and the copper loss:

$$dW = -IE_z dz - W(1-e^{-2adz}) = -IE_z\, dz - 2aW dz.$$

Substituting for E_z in terms of W by means of equation (2) and integrating between 0 (where $W = W_e$) and z gives equation (5).

Substituting for W in terms of E_z, again by means of eq. (2), gives:

$$E_z = (pW_e)^{\frac{1}{2}} e^{-az} - \frac{Ip}{2a}(1-e^{-az}),$$

which when integrated over the length L and rearranged, gives equation (6).

From eq. (6) the electron energy to be expected with given values of p, a and L can be computed [4]) provided that the power flux W_e at the beginning of the corrugated guide is known. In an accelerator without feedback this is of course equal to the magnetron power, W_m. In the case where feedback is used (*fig. 2*) W_e is calculated as follows. If it is assumed that there is no loss of power in the feedback bridge itself, but a reduction by a factor k at the input and output feeds to the corrugated guide (see fig. 2), then:

$$W_e = kW_m + k^2W_0,$$

where W_0 is the power flux at the output end of the accelerator. W_0 can be expressed in terms of W_e using equation (5), putting $z = L$. Eliminating W_0 from these two equations gives W_e, and the electron energy can be computed.

Conclusions from the calculations

In this way the accelerator performance to be expected for various designs of corrugated guide was calculated. A few results are illustrated in *figs. 3* and *4*. Fig. 3 is valid for $a/\lambda = 0.2$ and a *single section* waveguide with feedback. Fig. 3*a* shows the variation of electron energy with beam current for various lengths of guide [5]); fig. 3*b* shows the variation of the electron energy with length, for a beam current of 25 mA. In the same diagram is plotted the frequency shift Δf which produces a phase error of 35°.

Figs. 4*a* and *b* provide similar information for a corrugated guide with $a/\lambda = 0.168$ and which consists of *two sections* in series, each with its own feedback loop.

These graphs make it clear that, if the effect of beam loading had been neglected, the error in the estimation of the final electron energy would be 1.5 to 2 MeV.

Considering figs. 3 and 4 in more detail it will be seen that for $a/\lambda = 0.2$ the maximum beam current which can be achieved at 15 MeV in a guide 6 metres long is 12 mA. Alternatively at 25 mA the energy is 14.15 MeV. Under these conditions the permissible frequency change of the magnetron (for a phase shift of 35°) is $\pm$ 0.527 Mc/s. Again by increasing the length, 25 mA at 15 MeV

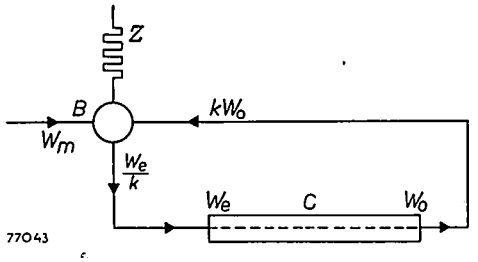


Fig. 2. Power flux in a linear accelerator with feedback. C = corrugated waveguide, B = feedback bridge, Z = dummy load, W_m = power arriving from the magnetron, W_e = "circulating power", W_0 = attenuated power at output end of waveguide.

[4]) The form of eq. (6) is convenient for computation since the function $(1-e^{-x})/x$ has been tabulated by Miller and Rosebrugh, Trans. Roy. Soc. Canada (2) **9**, sect. III, 73-107, 1903, and by G. F. Hardy, Trans. Fac. Actuaries **8**, 57-86, 1918.

[5]) These curves do not represent the variation of energy with beam loading of a practical accelerator. Each point on the curves represents the performance of an individual accelerator in which the corrugated guide has been designed to keep the wave in phase with the electron. Further, it is assumed that the feedback ratio n of the feedback bridge is always maintained at the correct value kW_0/W_m. Since W_0 will vary with the beam loading and n cannot be adjusted during operation, the latter condition will not be fulfilled in a practical case. Therefore, the curve of variation of energy with beam loading in a practical accelerator will be tangential to the appropriate curve in fig. 3*a* at the beam current for which it is designed but at lower and higher currents the electron energy will be less than that indicated in the curve.

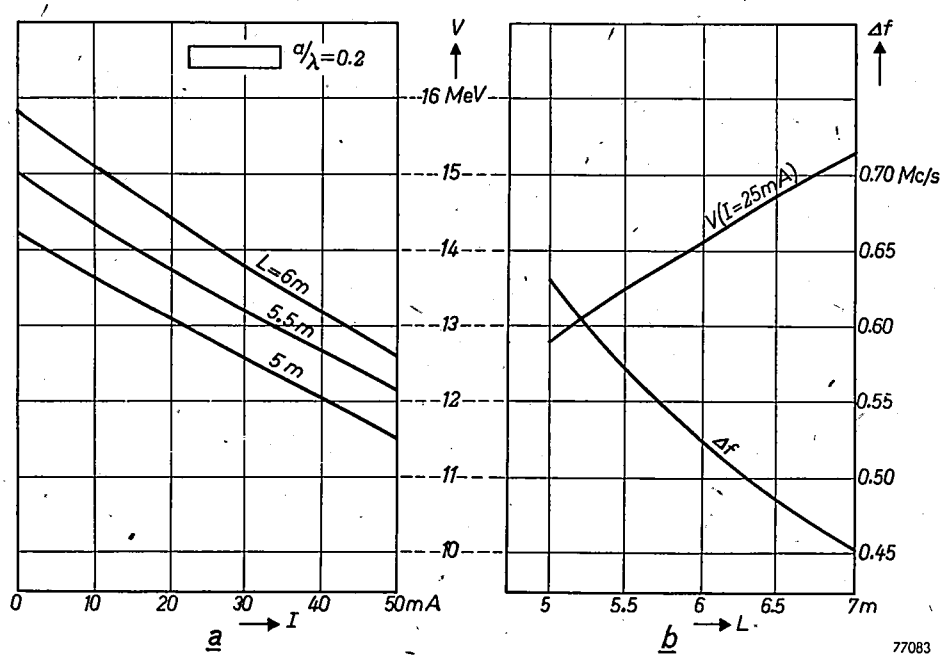


Fig. 3. Predicted performance of linear accelerators with feedback, in *one section* and with waveguide parameter $a/\lambda = 0.2$. Magnetron frequency 3000 Mc/s, magnetron pulse power 1.8 MW; the power loss at each doorknob feed is 5%, or $k = 0.95$.
a) Variation of predicted electron energy V with rated beam current I, for different values of accelerator length L.
b) Predicted energy V for 25 mA rated beam current and magnetron frequency deviation Δf for a phase error of 35°, as functions of accelerator length L.

is obtained with 6.73 metres, but the permissible frequency change is now only ± 0.47 Mc/s. While the frequency sensitivity at 6 metres, with $a/\lambda = 0.2$, is on the borderline, an accelerator could be constructed provided the target performance of either electron energy or beam current is relaxed; on the other hand the target performance can be realised at 6.73 metres but the frequency sensitivity would make the machine difficult to operate over a long period without constant attention.

For $a/\lambda = 0.168$ and the *two section* accelerator it will be seen that a 6 metre length gives 25 mA at 15 MeV with a permissible frequency change of ± 0.62 Mc/s. These were the operating conditions

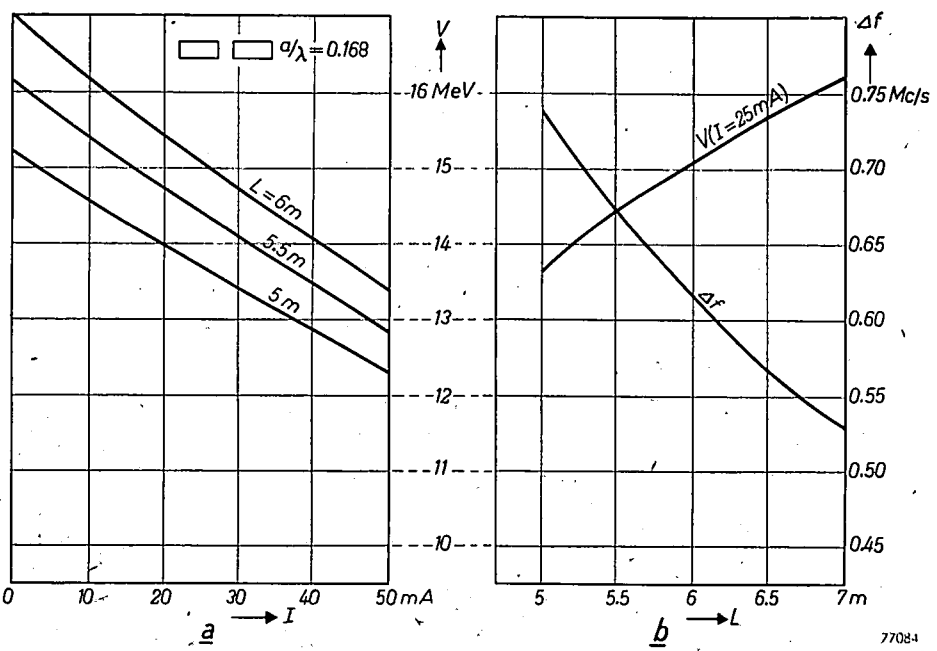


Fig. 4. Same as fig. 3, for linear accelerators with feedback, in *two sections* and with waveguide parameter $a/\lambda = 0.168$.

which were finally adopted. With this choice of a/λ the iris radius of the corrugated guide will be 1.68 cm.

Somewhat more complex conditions apply in the first 30 cm of the guide, where the bunching of the electrons (cf. I) must be accomplished. This has not been taken into account in the above calculations.

Main components of the accelerator

Fig. 5 is a schematic diagram of the operation of the accelerator and shows the relationship of the various waveguide components. *Fig. 6* shows an artist's impression of the layout of the accelerator in the shelter at Harwell; the control desk and general power supply rack are not shown in this drawing.

Referring to figs 5 and 6, a short survey of the main components of the apparatus will be given as an introduction to the detailed description that is to follow.

The high frequency power produced by the magnetron is fed through a rectangular waveguide to a Tee junction (T) where it is divided into two equal parts; the two outputs are fed to the *feedback bridges* B_1 and B_2 associated with the first and second accelerator sections, C_1 and C_2 respectively. The phase of the wave fed to B_2 can be adjusted by means of a *phase shifter* P_0. Power is fed from each feedback bridge to the input of the appropriate corrugated waveguide section (C_1 and C_2) through the "doorknob" feeds D_i. By adjusting P_0 the wave fed to section C_2 is phased precisely to pick up the electron bunches arriving from section C_1. The residual power of the attenuated wave reaching the end of each section is brought via the output doorknob feed D_0 back to the appropriate feedback bridge where it is combined with the incoming power from the magnetron. The phase of the power fed back is controlled by the *phase shifters* P_1 and P_2; a check on the correct phasing is made by monitoring the power levels by means of *waveguide thermocouples* near the input and output feeds.

The *magnetron* (type VX4061) is operated at a power output of 1.8 MW in the pulse, as mentioned above. Its frequency is adjusted by means of a mismatch probe and a line lengthener (ceramic

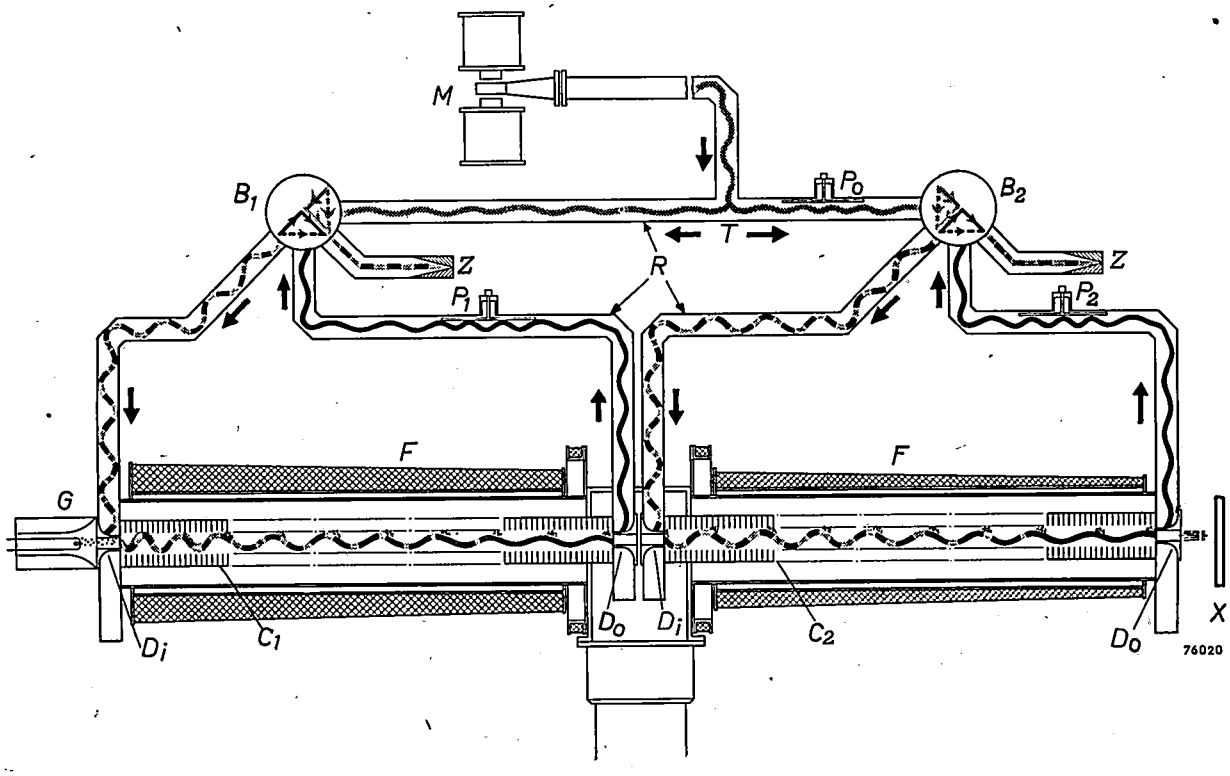


Fig. 5. Schematic diagram of operation of the 15 MeV accelerator. C_1, C_2 = the two sections of the corrugated waveguide, with input and output doorknob feeds D_i and D_0, electron gun G, and focusing coils F. M = magnetron. R = rectangular waveguide system, with dividing Tee T, feedback bridges B_1, B_2, dummy loads Z and phase shifters P_0, P_1, P_2. The power arriving from the magnetron (W_m) is indicated by the grey wave, the attenuated power leaving each accelerator section (W_0) by the black wave; these powers are combined in the feedback bridges and fed to the input doorknob of each accelerator section (black and grey wave). The electrons after being bunched in the first waveguide section travel in a stable phase position somewhat ahead of the wave crest, towards the target X.

wedge phase shifter [6]). The magnetron requires a magnetic field of the order of 0.15 Wb/m² (1500 gauss). This field should be adjustable, as for any given valve there is an individual optimum value. Hence an electromagnet, fed with stabilized current, is preferred to a permanent magnet. The required high frequency power pulses are obtained from the magnetron by feeding to it voltage pulses in the region of 40-50 kV and of 2 μsec duration from the *modulator* (*Mod* in fig. 6). The power in the pulse which the modulator has to deliver depends on the individual magnetron sample, i.e. on the actual efficiency, and may be as large as 5 MW (for an inefficient magnetron).

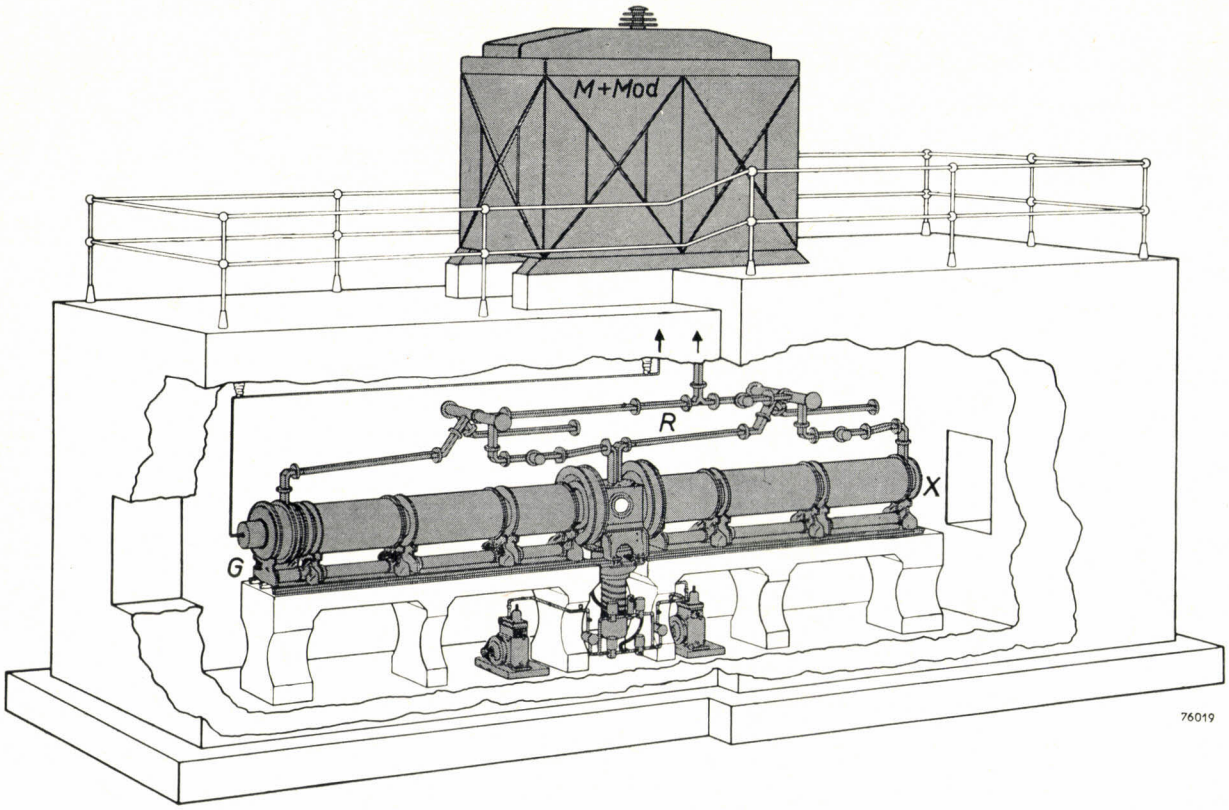


Fig. 6. Artist's view of the layout of the 15 MeV accelerator in the shelter at Harwell. The components are indicated by the same letters as in fig. 5. On top of the shelter are the magnetron *M* and modulator *Mod*. In practice the magnetron has been housed in a separate unit. Note the high-voltage line to the electron gun.

The modulator design is essentially similar to that of a high power radar modulator and is based on charging a pulse forming network and discharging it through a pulse transformer via an ignitron. It differs, however, in that the condition of freedom from time jitter is of considerably greater importance. In fact, an uncertainty in neutron energy in time-of-flight experiments [7]) is occasioned by uncertainty in the actual time of the modulator pulse, and this error should be small in comparison with other sources of error, for example the duration of the electron pulse. At present the *electron gun* which supplies electrons to the corrugated waveguide is pulsed simultaneously with the magnetron by the modulator. It is possible, however, that shorter electron pulses will be used in the future. It was therefore decided to tolerate a jitter of the main modulator of not more than about 0.1 μsec.

A special circuit involving the use of a "magnetic diode" was developed for stabilization of the gun filament current, as experience with the 3.5 MeV accelerator had indicated that both beam current and X-ray output are very sensitive to small filament current changes.

The considerable defocusing forces acting on the electron beam, especially in the first part of the corrugated guide, are neutralized by the use of

[6]) We cannot in this article enlarge on the details of standard magnetron and waveguide techniques. For such details the reader may refer, for instance, to C. G. Montgomery, R. H. Dicke and E. M. Purcell, Principles of microwave circuits, Mass. Inst. Technol. Radiation Series No. 8, McGraw Hill, New York 1948. See also W. Opechewski, Philips tech. Rev. **10**, 13-25 and 46-54, 1948, and A. E. Pannenborg, Philips Res. Rep. **7**, 131-157, 169-188 and 270-302, 1952.

[7]) T. I. Taylor and W. W. Havens, Nucleonics **5**, 4, Dec. 1949.

focusing coils covering the entire length of the guide and producing an axial magnetic field whose value is just sufficient to constrain the electrons to move in a helix about the corrugated guide axis.

The guide must be evacuated to a pressure lower than 10^{-5} mm Hg, to give a mean free path greater than the length of the electron path (6 m) to avoid loss of beam current by collision with gas molecules. This low pressure will at the same time eliminate the danger of sparking in the corrugated guide, where peak electric fields of the order of 30 kV/cm will be present.

Sparking must also be taken into account in the *rectangular waveguide*. In this case it could be eliminated either by evacuating (to about 10^{-4} mm Hg) or by pressurizing to 2 or 3 atmospheres [8]). While there is little doubt that a pressurized waveguide system would have been very convenient (small leaks could then have been disregarded), it would then have been necessary to develop a waveguide window capable of withstanding a differential pressure of 2 to 3 atmospheres and at the same time of passing up to 1 kW mean power ($^1/_2 \times 1.8$ MW with a duty cycle of up to 1 : 1000) without undue losses. As a result it was decided to evacuate the whole system right up to the magnetron window.

The *vacuum system* is pumped by a single nine inch diffusion pump whose unbaffled pumping speed is 1500 litres/sec. While it would have been possible to evacuate the corrugated waveguide from one or both ends, the high pumping impedance of an iris loaded structure would give a considerable pressure gradient along the guide. Furthermore it is not improbable that the pressure near the end away from the pump (or the centre) would be too high for satisfactory operation. To avoid these difficulties it was decided to pump the corrugated guide *radially;* it is therefore enclosed in a vacuum envelope, each cell of the guide (see below) being perforated by four holes round its circumference.

To conclude this short survey of the main components and returning once more to fig. 6, it should be noted that the modulator and magnetron have been positioned on top of the concrete shelter. Thus the magnetron and its associated waveguide components have been mounted as close as possible to the accelerator in order to minimize attenuation of power from the magnetron in the rectangular waveguide. At 3000 Mc/s, the power loss in a run of rectangular waveguide, say, 30 feet long, is of the order of 10%. Whereas this is negligible in a radar application it would be quite serious for the satisfactory operation of the accelerator: 10% change in power level would in fact represent a 5% change in peak field and therefore in electron energy.

As the intensity of the radiation from the accelerator is very high (in the region of 2000 röntgen per min at 1 m distance from the target), the walls of the shelter have to be sufficiently thick to reduce the radiation to a reasonable level at all points where personnel may be working continuously. The walls and roof of the shelter have been constructed from barytes concrete, which owing to its high density strongly absorbs X-rays, and the walls are $4\frac{1}{2}$ feet thick in the vicinity of the target.

General mechanical construction

The chief point borne in mind in the mechanical design of the accelerator was that the machine should be easy to assemble. In addition it should be possible to dismantle the equipment for servicing, as required, and subsequently reassemble without the necessity of renewing the initial setting up adjustments. Finally, while it is convenient to support the corrugated waveguide (with its vacuum envelope and focusing coils) from below, it will be essential for medical application of the accelerator to have a clear space below the target end to accommodate the patient.

To satisfy these requirements it was decided to support each metre length of the guide on a separate three-wheeled trolley.The six trolleys are run into position on rails laid on two reinforced concrete tables. Mounted between these two tables is a fixed centre member which carries the diffusion pump (cf. fig. 6). This set-up also allows the trolleys (or only those nearest to the target end) to be supported from above if desired.

Fig. 7 shows the centre member A and the two trolleys T_1 and T_2 adjacent to it. The centre member is machined from an aluminium forging; aluminium was chosen rather than brass for lightness and economy, and as this member forms part of the vacuum envelope a forging is preferred to a casting where the risk of blow holes or porosity exists. The centre member serves also to support the doorknob feeds D associated with the two sections C_1, C_2 of the corrugated guide. The trolleys consist basically of two aluminium end castings which are held apart at the top by the focusing coil former F and at the bottom by a tubular member K. One end of the trolley carries two single flanged wheels and the other end one double flanged centre wheel. Owing to this three-point-support the system will be

[8]) The two possibilities correspond with choosing a point sufficiently high up either to the left or to the right of the minimum of the well known Paschen curve.

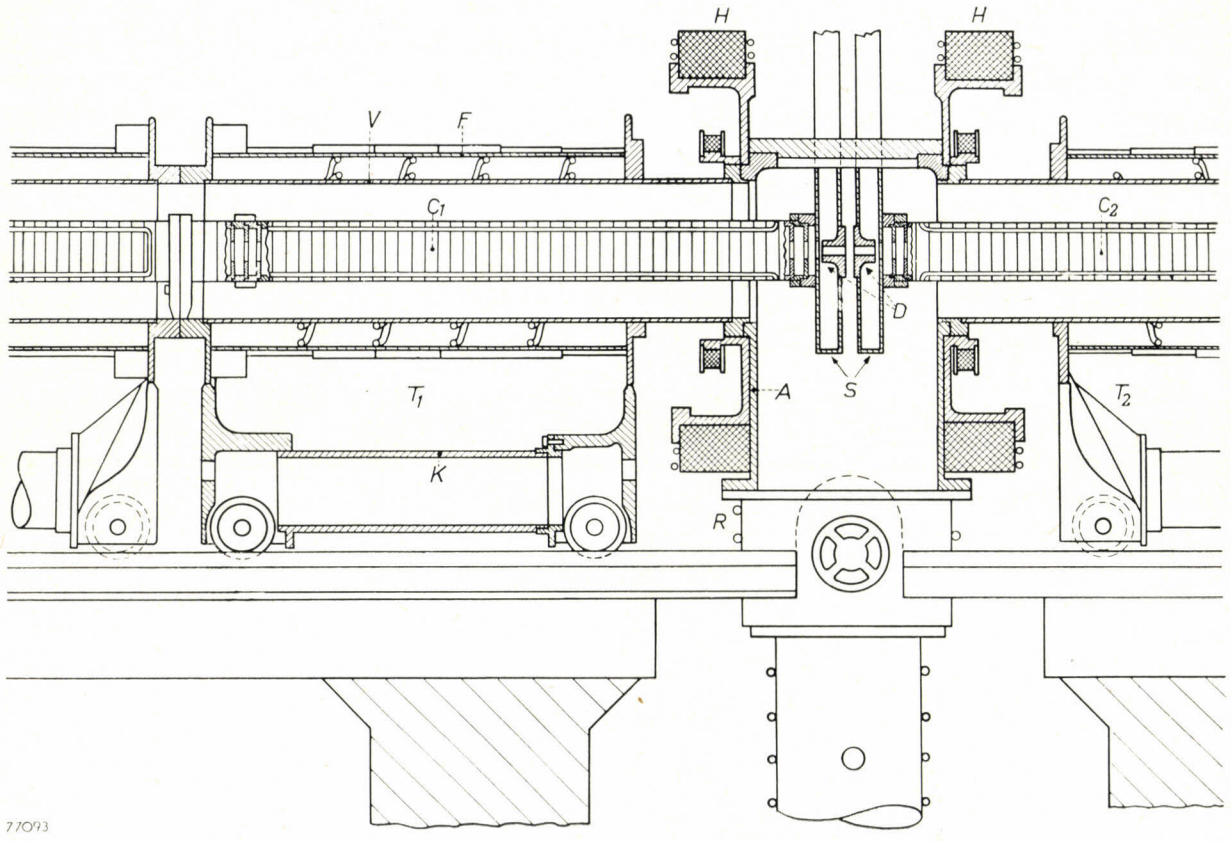


Fig. 7. Centre member A and adjacent trolleys T_1, T_2 of accelerator. F = focusing coil former. K = tubular member. C_1, C_2 = corrugated waveguide sections. V = vacuum envelope. D = doorknob feeds. S = slots in rectangular waveguide. H = Helmholtz coils. O = oil diffusion pump. R = section of envelope with special cooling.

strain-free whatever be the position of the supporting rails. The accuracy required in laying down the three rails, which are rigidly attached to the concrete tables, is determined by the range of the adjustment provided for aligning the trolleys.

Fig. 8 shows a complete trolley prior to connection to the main system. The tubular sections of vacuum envelope consist of $\frac{1}{4}$ inch wall copper tube to which are hard-soldered flanges of strong, nonmagnetic metal ("Delta" bronze, which is much stronger than brass). Each section of vacuum envelope is supported at its flanges inside the focusing coil former by three adjusting screws at each

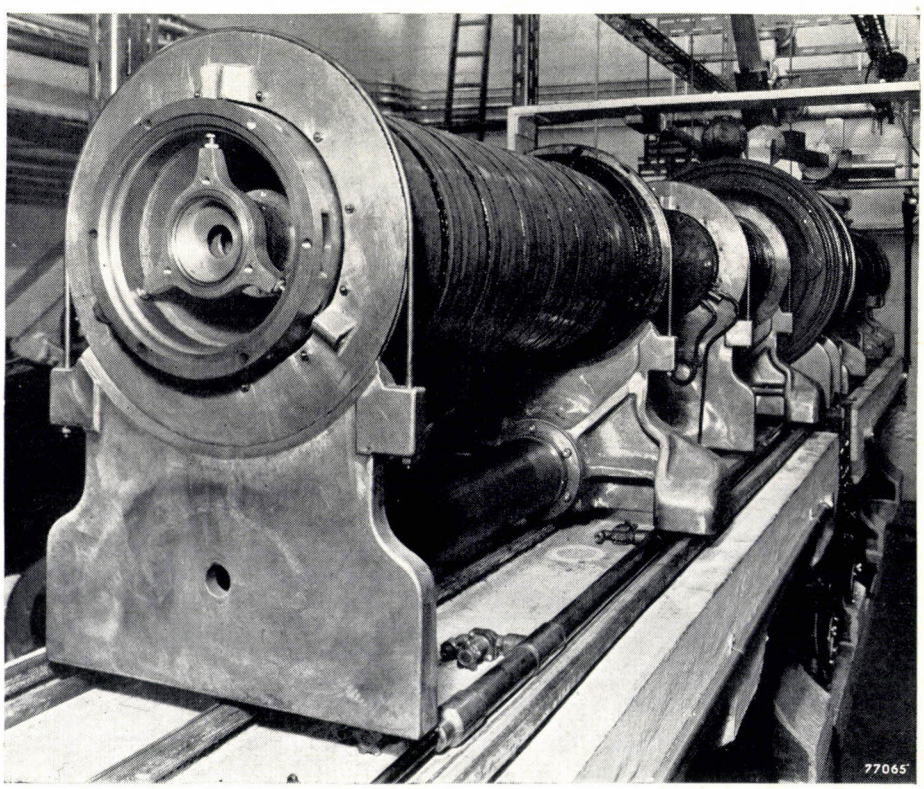


Fig. 8. Complete trolley prior to connection to the main system. The corrugated waveguide supported by a three-legged spider in its vacuum envelope is clearly seen.

end which permit accurate alignment. Similarly each metre section of the corrugated waveguide is in its turn supported in the vacuum envelope at each end by a three-legged spider with adjusting screws.

In this description some of the detailed features of the design have already been mentioned. The second part (B) of this article will continue along these lines, describing the main components listed above and discussing a large variety of detail problems that had to be solved in their development. The third part (C) will deal with the methods of testing and the actual performance of the accelerator.

B. DETAILED DESCRIPTION OF MAIN COMPONENTS AND CIRCUITS

Corrugated waveguide

Mechanical construction

As the successful operation of the accelerator is dependent to a very important extent on the correct performance of the corrugated waveguide, its method of construction required careful consideration in relation to dimensional accuracy and, of course, manufacturing costs.

In particular the surface finish and radio frequency joints between adjacent parts must be such as to give a low value of attenuation, the diameter of the guide must vary with distance along the guide to give the predicted variation of phase velocity [9]) and lastly the inner hole diameter and corrugation pitch must not exhibit irregularities which would give rise to standing waves. The importance of the latter point does not stem from the power loss involved (this amounts to only 0.06 dB for a voltage standing wave ratio of 0.8); the main point is that it severely restricts the tuning range of the magnetron.

The corrugated guide is composed of cells, whose general shape is illustrated in *fig. 9a*. For the major part of the length of the guide a corrugation pitch (length of cells) of 2 cm was chosen as a suitable compromise between the required uniformity of field between corrugations and the copper loss in the dividing walls (cf. I). For the first section of the guide, however, where the electrons must be bunched and where the phase velocity of the wave must vary from 40% to 90% of the velocity of light within a distance of only 30 cm, a smaller pitch is necessary. Thus the first 30 cm of the guide are made up of cells with 1 cm pitch. Next come two special 1 cm cells that serve as a matching section between 1 cm and 2 cm cells. The rest of the guide consists of 2 cm cells. The variation of phase velocity after the first 30 cm guide length is from 90% to 99.8% of the velocity of light at the end of the first three metre length, and then to 99.95% at the target end. The variation of cell diameter (outer diameter of the corrugations) required to produce the variation of phase velocity from 90% to 99.95% of the velocity of light is only 0.2 mm. The cells are spigotted together as shown in fig. 9*b*; a groove is provided at the corner of the male spigot for the insertion of a solder ring.

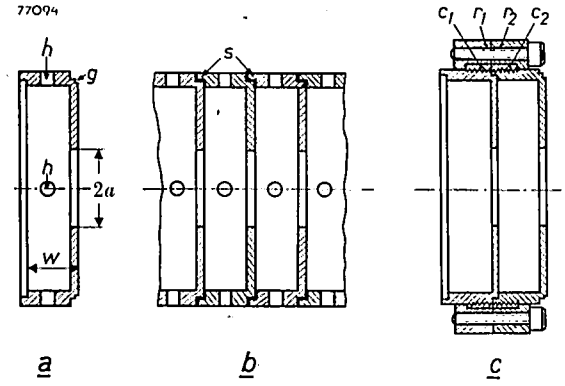


Fig. 9. *a*) One of the 316 cells from which the corrugated guide is composed. Two types of cells are used, viz, of pitch $w = 2$ cm and 1 cm respectively. g = groove for solder ring. h = pumping holes. a = inner iris radius.
b) Two cells showing spigotting action. s = solder fillet.
c) Clamping of two guide sections by rings r_1, r_2 registering on collars c_1, c_2.
The joints shown in (*b*) and (*c*) are important since good radio frequency contact is required.

The method of machining and assembly involved considerable experimental investigation. High conductivity copper strip 5 inch wide and 0.25 inch thick is used as the basic raw material. From this, cold pressings in the shape of a cup (*fig. 10*) are made using a 75 ton press. The pressings are machined to within 0.25 mm of final dimensions. The considerable stresses produced in the material by the cold working must then be relieved in order to avoid the risk of dimensional changes during the subsequent fine machining. The stress relieving is effected by heating to 800 °C in an atmosphere of "mixed gas" (90% N_2, 10% H_2). This process

[9]) The detailed design of the waveguide was carried out by W. Walkinshaw of the Atomic Energy Research Establishment, Harwell.

performs at the same time the function of rendering the copper in the neighbourhood of the surface effectively oxygen free [10]), the importance of which will be evident presently. The cells are then fine machined to tolerances of $\pm 10\ \mu$ at 20 °C on all internal dimensions and on the spigot. It will be appreciated that control of temperature during measurement of the cells is of great importance since an error of 5 °C during the measurement of the guide diameter would suffice to throw the dimension outside tolerance.

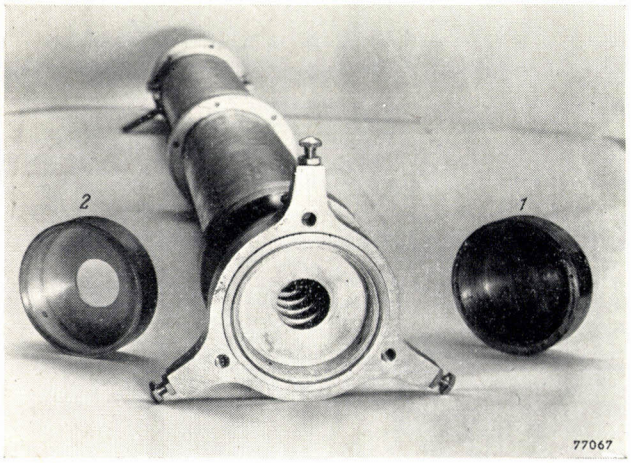


Fig. 10. A length of corrugated guide with centering spider. *1* = cold pressed cup from which a cell (*2*) is machined. Note the pumping holes in *2*.

After machining the cells are degreased and assembled in a stack with copper-silver eutectic solder rings inserted in the grooves provided. The stack is then heated to 780 °C in an atmosphere of pure nitrogen; mixed gas is admitted for a few minutes at the soldering temperature, the hydrogen acting as a flux for cleaning the surface in the soldering process. Finally the guide is cooled slowly in a nitrogen atmosphere. In this way a very satisfactory radio frequency joint is obtained. It should be noted that the surface of the copper in the soldering process must be oxygen free since otherwise the hydrogen flux will tend to cause blisters on the surface of the copper with a consequent serious increase in RF attenuation.

The phase velocity of each section of assembled waveguide is checked by treating the system as a standing wave circuit. A short circuiting plate is placed at one end and RF power at 3000 Mc/s fed to the other; the wavelength and hence the phase velocity is measured by introducing a small loop into one of the pumping holes of each cell in turn and plotting the field pattern. As a result of these measurements of phase velocity, correcting sections of 10 cm length were inserted at four places in the corrugated guide, in the 15 MeV accelerator.

Using the same set-up the attenuation is measured by feeding power into and out of the guide through doorknob transformers and observing the mean power level at points along the guide.

Adjacent lengths of corrugated guide are spigotted and are clamped together by means of a pair of rings, each of which registers on a collar on the end cell as shown in fig. 9*c*. At the end of metres no. 1, 2, 4 and 5 these rings are replaced by the three-legged spiders mentioned earlier and which serve to support and to centre the length of the guide in the vacuum envelope. A length of guide with a spider is shown in fig. 10. The ends of each three metre length are supported by the doorknob transformers.

Water-cooling

At the maximum pulse repetition frequency the power dissipation in the corrugated guide can amount to as much as 300 W per metre length. As the area of contact of the spider adjusting screws with the vacuum envelope is necessarily very small, the only effective points of heat transfer by conduction are at the doorknob feeds. In a few hours running quite high temperatures could therefore be developed in the guide and since a change in mean temperature of only 10 °C produces a 30° phase error between wave and electron, it is necessary to water-cool the guide. This is effected by cooling pipes which have a large area of contact with the guide and which keep the temperature rise within 2 °C.

Since the guide and the vacuum envelope are in effect rigidly clamped together at their ends, any temperature variation of the envelope would cause the guide — which remains at constant temperature — to be in tension or compression, which results in undesirable mechanical strains. (A phase error would also result, but this would be very small since only the length would be affected, where dimensional errors are only 1/20 as critical as on the diameter.) The undesirable effect could be avoided by providing an expansion joint in the vacuum envelope but it is mechanically simpler to pass the cooling water from the guide to a helix soldered to the vacuum envelope so that both are maintained at effectively the same temperature. Before going to waste the water passes to the focusing coil system (to be described presently).

Each metre length of the accelerator has its own water circuit performing this triple cooling function.

[10]) Brit. Pat. Application 23322 of 1952.

Electron gun

The electron gun is a simplified Pierce gun which injects into the corrugated waveguide a substantially parallel stream of electrons. The electrode geometry is shown in *fig. 11a*. A directly heated flat spiral tungsten cathode is situated just inside the apex of a truncated conical focusing electrode. The anode is at earth potential whilst the cathode is pulsed to —45 kV by the main modulator. The cathode structure is supported and insulated by a plate of "toughened" glass.

Experience with the 3.5 MeV accelerator showed that while the average cathode life was in the region of 150 hours, the performance of the accelerator did not give any indication of cathode deterioration until fracture actually occurred. Thus it may well happen that the cathode must be changed in the middle of a long experiment. If cathode replacements required opening-up the whole vacuum system, about two hours would be necessary to re-establish the vacuum after the operation. To avoid this waste of time it was decided to incorporate a vacuum valve between the main accelerator envelope and the electron gun [11]). The valve consists of a slide plate in contact with the flat anode of the gun (fig. 11*b*). Corresponding holes for the passage of electrons and for pumping are drilled in the anode block and slide plate. These holes are in register when the valve is open. When the slide plate is translated to the closed position and pressed on to the anode, the holes in the latter are shut off by "Neoprene" O-rings held in self-retaining grooves in the slide plate. The mechanism for moving and compressing the slide plate is shown and explained in *fig. 12*.

After the cathode has been replaced the gun space is pumped down to about 0.04 mm Hg by one of the rotary pumps of the vacuum system and the valve is then opened slowly. The resultant

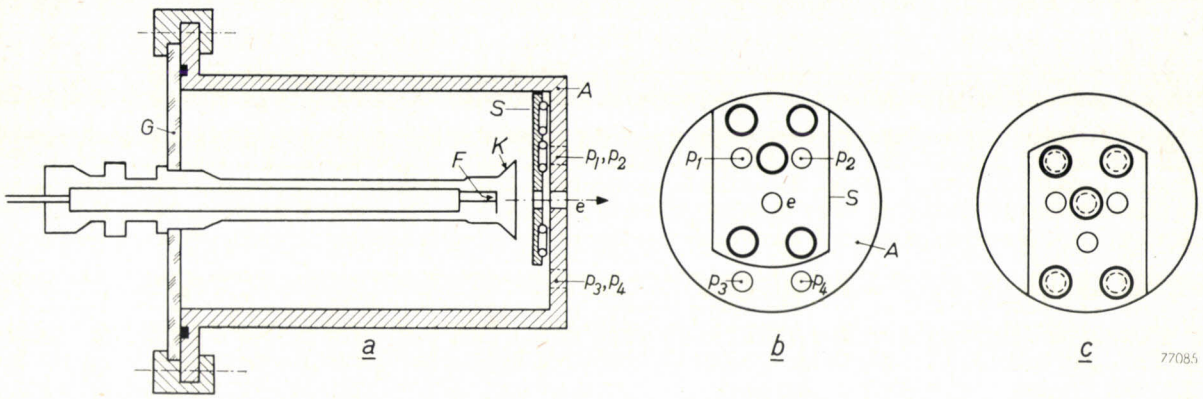


Fig. 11. *a*) The electron gun is of the Pierce type. F = replaceable, directly heated flat spiral tungsten filament. K = focusing electrode. A = flat anode with hole e for the passage of electrons. G = insulating glass plate. S = sliding plate of vacuum valve. *b*) and *c*) Mechanism of vacuum valve. In the open position (*b*) three holes — e for the electrons and p_1, p_2 for pumping — of both anode plate and slide plate are in register; pumping holes p_3, p_4 of the anode are open. In the closed position (*c*) the five holes of the anode plate are shut off by "Neoprene" rings held in self-retaining grooves in the slide plate.

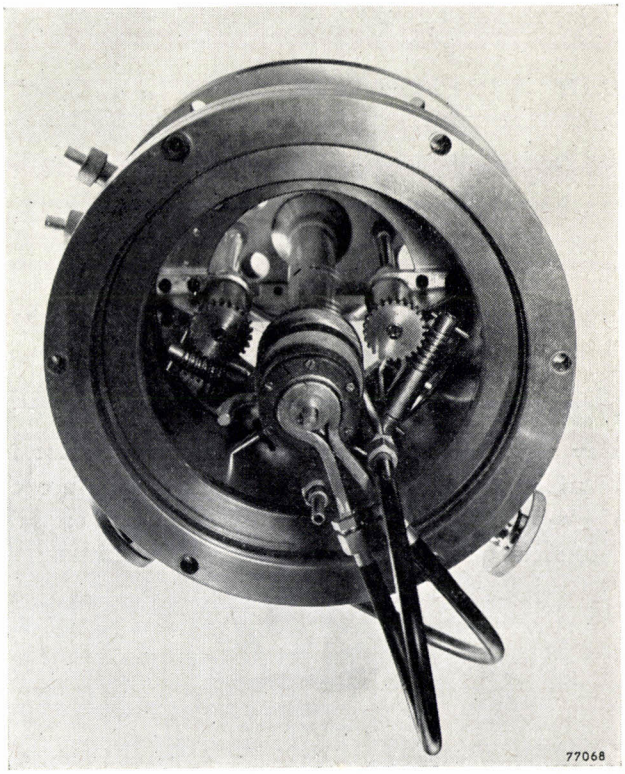


Fig. 12. Electron gun. Through the glass plate supporting and insulating the cathode, the mechanism for actuating the vacuum valve may be seen. The slide plate (with holes) is seen at the far end. The translational movement of the plate is effected by a rack and pinion and the compression by a screw thread. Both movements are controlled by shafts passing through vacuum-tight bearings (Wilson seals) in the gun envelope; the driving knobs are seen at the bottom.

[11]) Brit. Pat. Application 1060 of 1953.

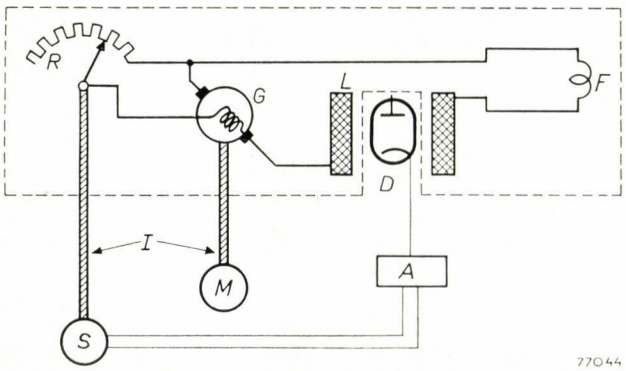
rise in pressure in the main system is very small and the pressure rapidly falls back to its normal value.

It has been stated above that it is desirable to stabilize the gun filament current. Since the cathode is pulsed to 45 kV the problem of stabilization requires special measures. The system adopted is shown in *fig. 13 a* and *b*. A D.C. generator, driven through an insulating shaft by an A.C. motor, feeds the gun filament through an 8 inch diameter cylindrical coil. A magnetic diode (which has a D.C. magnetron characteristic [12])) placed inside and on the axis of the coil examines the field and indicates (by its anode current) whether the field differs from that corresponding to the filament working current. This information is fed via a push-pull amplifier to a servo motor, which drives through suitable gearing and an insulating shaft a variable resistor in the field circuit of the D.C. generator. In this way the filament current can be maintained at its working value. To allow adjustment of the current to lower values, a subsidiary coil surrounding the diode is provided. By passing a current through this coil the field at the magnetic diode is increased and the main current will decrease since the diode will tend to maintain the *total* field at the same value as before.

The only components of the power supply which are at high potential, are the generator, its field resistor and the field coil. A slipping clutch and limit switches are provided to avoid damage to the field resistor if the gun filament should go open or short circuit.

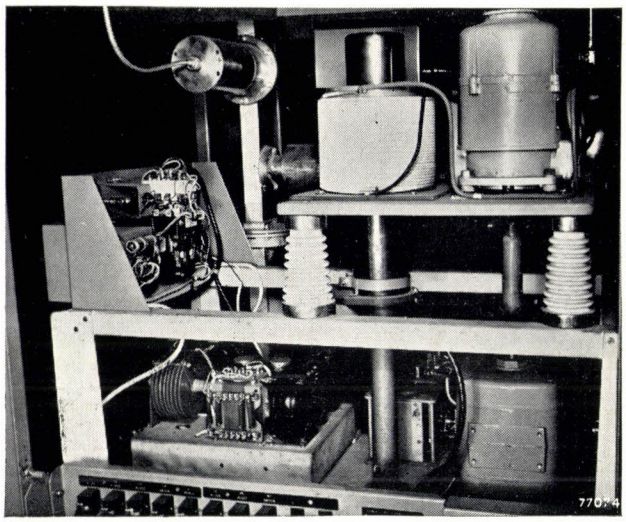
This servo control system has proved very satisfactory in practice and gives a filament current stable to better than $\pm 1\%$.

An alternative and relatively simple method involves feeding the filament through a bifilar choke. Conventional means for stabilization using components at or near earth potential can then be adopted. The choke, however, would have a large capacitance to earth. In the particular case of the accelerator at Harwell it is hoped in the future to pulse the electron gun by a separate modulator having a pulse duration considerably shorter than the 2 μsec employed at present (see page 7). In the interests of economy in this subsidiary modulator it was desirable to keep the capacitance to earth of the gun and its power supply as low as possible.



a

Fig. 13. *a*) Circuit for stabilization of the gun filament current. Only the components enclosed in the dotted line are at high potential (—45 kV during the gun pulse). G = D.C. generator driven by motor M. F = gun filament. L = field coil. R = variable resistor. D = magnetic diode measuring the field in the centre of coil L. A = push-pull amplifier. S = servo motor. I = insulating shafts.



b

b) Photograph of the stabilizer.

[12]) Cf. H. B. G. Casimir, A magnetron for D.C. voltage amplification, Philips tech. Rev. **8**, 361-367, 1946.

Focusing system

The defocusing forces acting on the electron beam in the corrugated waveguide decrease as the velocity of the electrons approaches the velocity of light (see article I). As a result, the axial magnetic field required for focusing must be highest near the electron gun end of the guide. The calculated value for the 15 MeV accelerator varies between 0.0750 Wb/m² at the gun end to 0.0125 Wb/m² at the target end.

The coil system for producing this focusing field was designed round the use of windings 4 inches long and 10 inches inside diameter, fed by a current in the region of 3 A; the number of turns (and hence the outer diameter of the windings) must vary along the length of the accelerator. The windings are assembled on cylindrical bobbins just under 1 metre long. This subdivision of the

coil system makes for easy access to the water-cooling connections of the vacuum envelope at the end of each metre length. To bridge the gap created by the centre member of the accelerator and the diffusion pump (cf. fig. 7) a pair of large Helmholtz coils have been used. To avoid carrying the focusing coils beyond the ends of the accelerator (or alternatively of providing very large coils) in order to maintain the calculated field in the region of the ends, mild steel end plates have been used to provide a local concentration of field.

As the total power dissipation in the focusing coils is in the region of 6 kW, water-cooling has been provided. The cooling pipes are attached to the cylindrical coil formers which constitute the body of the trolley assemblies as shown in figs. 7 and 8. In the region of the gun end, where the power dissipated is considerable, cooling of the cheeks of the coils is also provided. The Helmholtz coils are cooled on one end cheek and on the outer cylindrical surface.

In spite of the water-cooling the mean temperature and hence the resistance of the coils rises during operation and this coupled with mains voltage variations could easily produce a 25% change in focusing field. Stabilization of the coil current was therefore necessary. In this instance it was decided to supply the coils from six independently variable stabilized power units, as experience with the earlier accelerator had indicated that a slightly higher beam current could be obtained by local changes in the focusing field from its predicted value. The six sections of the coil system were chosen such that each power unit supplies approximately the same wattage. This means of course that the coil divisions become progressively longer in the direction of the target; this feature is convenient since it is to be expected that the actual local value of field will be more critical in the low velocity region.

The field stability required was estimated to be in the region of $\pm 5\%$. A magnetic diode circuit was used for the stabilizer units as this was considered to be convenient and economical even though no high voltage problems are involved and the precision achieved is better than required.

Vacuum system

The diffusion pump gives a pressure of 2×10^{-6} mm Hg measured at the top of the centre member of the corrugated guide, under normal running conditions. Although the pumping conditions for the corrugated waveguide are not ideal, owing to the small diameter of the pumping holes in the cells, the actual pressure in the path of the electron beam is probably not greatly in excess of the measured figure and satisfactory operation of the accelerator is achieved provided the pressure, as measured, is not in excess of 4×10^{-6} mm Hg. Pumping conditions for the long rectangular guide with its small cross section are even worse but fortunately the vacuum requirements in this guide are not so stringent. Some pumping of the rectangular guide will occur through the doorknob cones and the corrugated guide. This in itself is undesirable since it tends to increase the pressure in the corrugated guide. Slots have therefore been made in the rectangular guide sections of the doorknob feeds (S in fig. 7). Pumping through these slots provides a satisfactory operational pressure throughout the whole rectangular guide system.

In order to avoid so far as possible the condensation of pump fluid on the water-cooled walls of the vacuum envelope and the corrugated waveguide the cylindrical section of envelope R in the neighbourhood of the pump baffle is cooled by water at least 5 °C colder than the main water supply. A standard milk cooler comprising a refrigerator unit, a well lagged 40 gallon (180 litre)

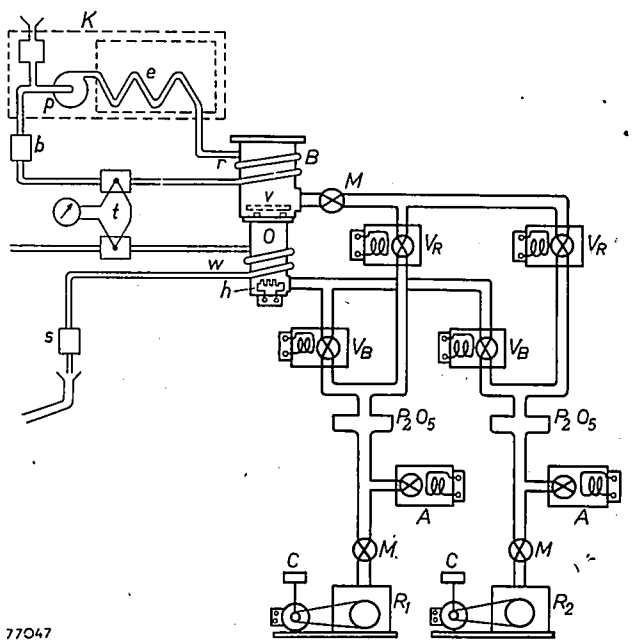


Fig. 14. Layout of the pumping system. O = oil diffusion pump with heater h, water cooling w and water flow switch s. R_1, R_2 = rotary pumps for backing and roughing. V_B = backing valves, V_R = roughing valves, A = air inlet valves. All these valves are magnetically operated. M = manual valves. C = centrifugal safety switches. P_2O_5 = moisture traps. B = baffle, with main valve v and refrigerated coils r; the water in this circuit is kept at least 5° C below that in the main cooling circuit by means of the refrigerator K with circulating pump p, heat exchanger e, water flow switch b (interlocked to the diffusion pump heater) and thermocouple t for monitoring temperature difference.

storage tank and a circulating pump, was adapted for this purpose. An electronic temperature controller was fitted which maintains the temperature of the water fed to the pump baffle at 5° C; this temperature is adequate in the Harwell installation, in which the main cooling water is recirculated and whose temperature never falls below 10 °C.

The layout of the pumping system is shown diagrammatically in *fig. 14* and the control panel and some of the component parts are seen in *fig. 15*.

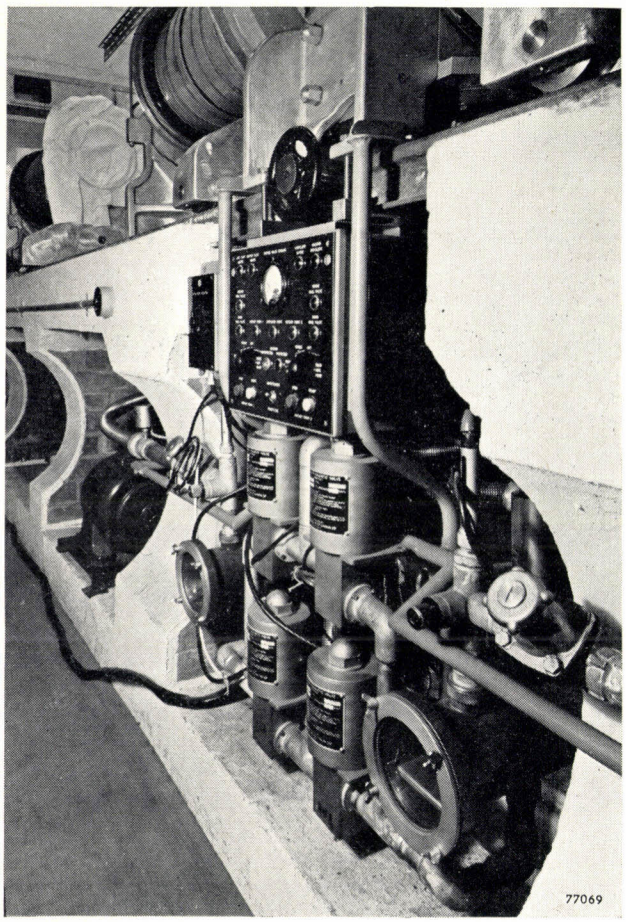


Fig. 15. Control panel and some of the component parts of the vacuum system. The four cylindrical containers under the panel are the magnetically operated valves. One of the rotary backing pumps is seen in the adjacent arch of the concrete table.

Two rotary vacuum pumps each with a displacement of 450 l/min are provided. While one of these is sufficient to maintain an adequate backing pressure for the diffusion pump, the second pump is useful when carrying out routine maintenance and also to save time when evacuating the whole system from atmospheric pressure (roughing).

The two pumps can be used either for backing or for roughing simultaneously, or each can carry out either function. In order to avoid the danger of mistakes in the control of these functions manual vacuum taps are not used; each pump is operated by means of a selector switch which controls a pair of magnetically operated valves.

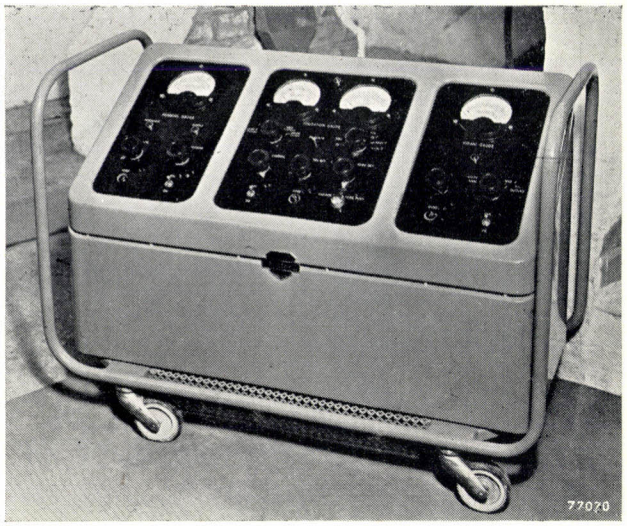


Fig. 16. Trolley housing the control units for all vacuum gauges (Pirani, Philips and triode ionization gauge).

Pirani, Philips and triode ionization gauges [13]) are provided for the measurement of rough, medium and fine vacuum respectively. The control units for all gauges are housed in the vacuum gauge trolley shown in *fig. 16*. The trolley is castor mounted and is connected to the gauge heads through a long flexible cable so that it may be moved around to facilitate leak hunting.

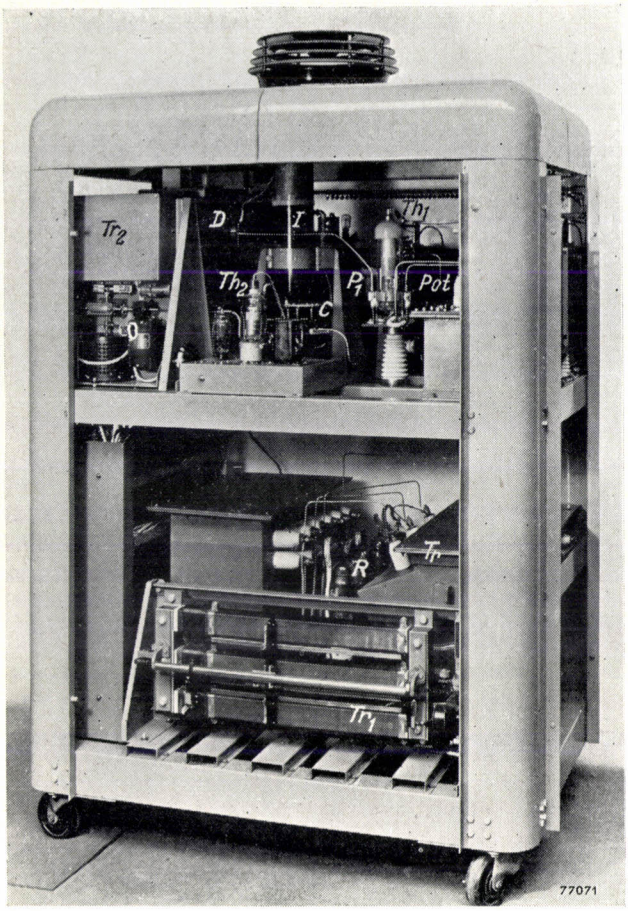
A number of protective devices connected with the vacuum system will be dealt with in a later section.

Modulator

The modulator which feeds the voltage pulses of 40-50 kV and 2 μsec duration to the magnetron and to the electron gun is shown in *fig. 17*. *Fig. 18* is a simplified circuit diagram.

The output of the main transformer *Tr* which is fed from a variable transformer Tr_1 is rectified and provides D.C. at 10-12 kV level. After smoothing, the output is fed through a ringing choke *L* and a gas triode Th_1 to the pulse forming network *D* ("delay line"). Owing to the inductance of the choke the pulse forming network will be charged to double the D.C. voltage supplied. When the ignitron *I* is triggered, the pulse forming network discharges through the ignitron and the primary of the 1:5 step-up pulse transformer P_1.

[13]) The ionization gauge control circuit is interesting in that a simple saturable reactor is used to stabilize the emission current at a preset value, in which case the plate current is a direct measure of the pressure in the system (Brit. Pat. Application 5377 of 1952).



The rectangular pulse across the secondary of this transformer is fed to the magnetron cathode. For maximum energy transfer the impedance of the primary of P_1 is matched to that of the pulse forming network so that the voltage is divided equally between these two and a pulse voltage of about 10 kV is obtained on the primary and 40-50 kV on the secondary. The pulse transformer is bifilar wound so that the heater transformer for the magnetron can be at earth potential.

If the modulator were to operate at a fixed pulse repetition frequency the introduction of the series triode Th_1 would not be necessary since the time constant of the ringing choke L could then be adjusted so that the pulse forming network is charged to the maximum voltage just at the time of the occurrence of the next pulse. However, in practice the accelerator is used at a number of pulse repetition frequencies between 100 and 500 pulses per sec and the introduction of a cut-off valve is therefore essential. Furthermore the use of a suitable time constant in the grid circuit of Th_1 ensures that the pulse forming network does not start to charge

Fig. 17. Main modulator producing the voltage pulses of 40-50 kV and 2 μsec duration fed to the magnetron and to the electron gun, with covers removed. The letters correspond to those used in fig. 18.

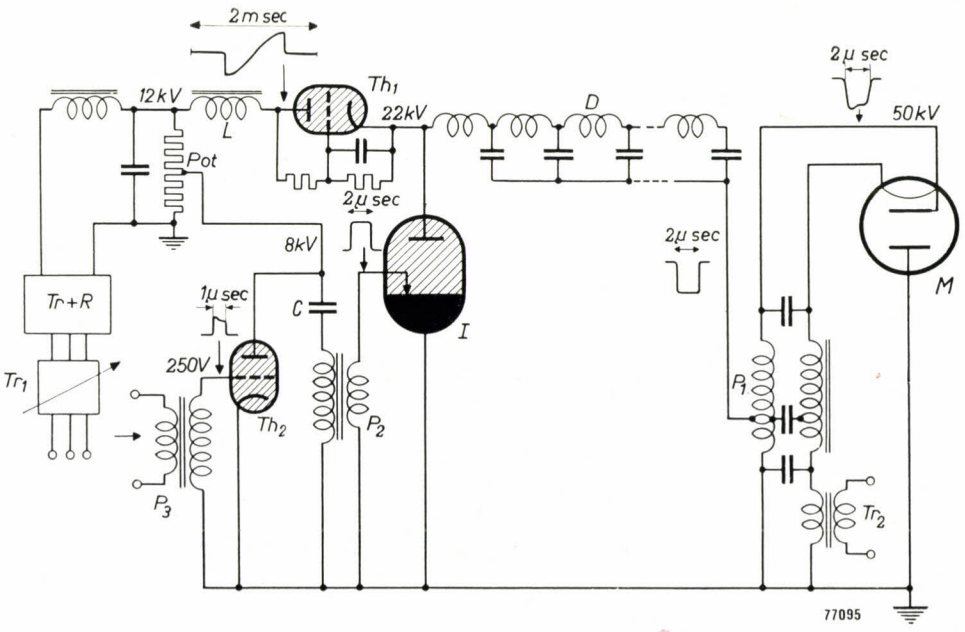


Fig. 18. Simplified circuit diagram of modulator, with waveforms obtained at several points. The pulse forming network D is periodically charged up from a 12 kV D.C. source via a ringing choke L and a cut-off gas valve Th_1. By discharging the network D through the ignitron I and the primary of step-up pulse transformer P_1, a rectangular voltage pulse is obtained on the magnetron M. The maximum pulse repetition frequency is 500 per second. The repetition frequency is controlled by triggering pulses fed through pulse transformer P_3 to the grid of the hydrogen thyratron Th_2 which on firing will discharge the capacitor C through the step-down pulse transformer P_2 thus delivering a high energy triggering pulse (2 kV, 300 A) to the igniter of the ignitron. $Tr + R$ = main transformer and rectifier. Tr_1 = variable transformer. Tr_2 = magnetron filament supply transformer.

until the ignitron discharge valve is completely de-ionized.

The choice of an ignitron as the discharge valve is rather unfortunate since it is inherently in its design prone to jitter and since its de-ionization time is relatively long. However, at the time the design was carried out, no hydrogen thyratron capable of handling 5 MW in the pulse was available. (A hydrogen thyratron to this specification is now made in the United States [14].)

In order to reduce jitter two measures have been adopted. The first aims at assisting ionization by maintaining the ignitron bulb at as high a temperature as possible subject to not excessively increasing the de-ionization time of the ignitron. Experiments with a number of samples showed that a reasonable compromise temperature is 18 °C. The cooling water which is fed to the ignitron is therefore warmed to 18 °C by a 500 W immersion heater stabilized by an electronic temperature controller (Mullard type E 7594).

The second measure involves the use of a large triggering signal on the igniter of the ignitron. The signal is a pulse of 2 kV, 300 A obtained by the discharge of the 0.01 μF capacitor C through the step-down pulse transformer P_2 and the hydrogen thyratron Th_2. The capacitor C is charged to 8 kV from a potential divider (*Pot*) connected to the main high voltage line. The hydrogen thyratron Th_2 receives its triggering pulses from a blocking oscillator which thus controls the pulse repetition frequency of the modulator. The large triggering pulse of 600 kW peak power on the igniter of the ignitron reduces the maximum jitter to a value in the region of 0.1 μsec and the mean value is considerably less. On the other hand there is some evidence that the use of such a large triggering signal reduces somewhat the ignitron life and it becomes progressively more difficult to fire.

As a time somewhat longer than 1/10 μsec is necessary for oscillation to build up in the magnetron and the effective impedance during this period is very high (the current drawn is very small) the steep voltage rise on discharging the pulse forming network tends to produce a voltage spike on the magnetron in the early part of the pulse. The corresponding current spike would tend to produce unwanted modes of oscillation in the magnetron which generally cause sparking. The effect has been removed by the use of a choke in series and a series connected resistor and capacitor in parallel with the magnetron.

[14]) Electronics **26**, Febr. 1953, p. 158.

Rectangular waveguide system

The whole waveguide run from the dividing Tee is seen in fig. 6. As the system is evacuated right up to the magnetron, all the sections, bends and waveguide components (feedback bridges, phase shifters etc.) are connected together using standard vacuum type choke couplings. The feedback bridges, phase shifters and monitoring thermocouples will be described in some detail below. Each part of the waveguide which contains a discontinuity has been individually matched with irises to avoid reflections as far as possible: the voltage standing wave ratio (cf. note [6])) is not less than 0.98 for any individual part except for the phase shifters which cannot be matched to give this V.S.W.R. over their whole range of phase adjustment. The overall V.S.W.R. of the system (as measured above the dividing Tee, at low power and with the system at atmospheric pressure) varies between 0.8 and 0.97, depending on the position of the phase shifters.

Feedback bridges

The feedback bridges are of the type known as circular magic-Tee [15]) which is illustrated in *fig. 19a* (see also figs. 1, 5 and 6). It consists of two circular

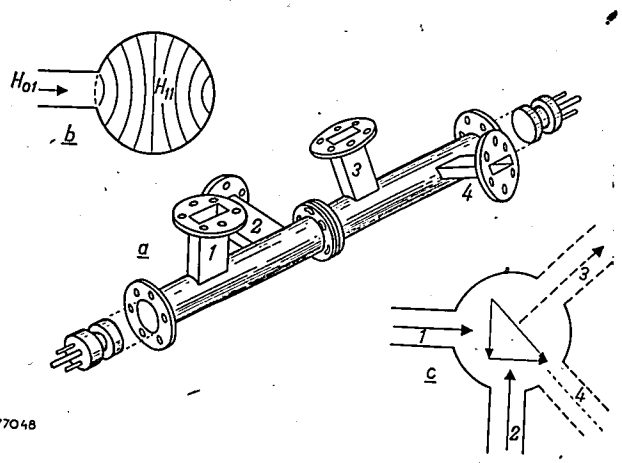


Fig. 19. *a*) Circular magic-Tee. It consists of two sections of circular waveguide, each section carrying two rectangular waveguide arms mutually at right angles and spaced axially by one wavelength. The two sections are joined by a rotatable vacuum joint and closed by an adjustable plunger for matching. *b*) Electric field lines of H_{11}-wave in the circular guide excited by H_{01}-wave in a rectangular waveguide arm. *c*) The two H_{11}-waves, excited in the circular waveguide by the rectangular waveguide arms *1* and *2*, have a resultant polarization at 45° (provided the phases are suitably adjusted). The sum of the two incident powers can then be drawn from rectangular waveguide arm *3*, and in the steady state no power will enter arm *4*.

[15]) The use of a circular magic-Tee in an accelerator feedback system was suggested by L. B. Mullett, see: L. B. Mullett and B. G. Loach, A. E. R. E. Report EL-R 93. See also: B. E. Kingdon, A circular waveguide magic-Tee and its application to high power microwave transmission, J. Brit. Inst. Rad. Engrs **13**, 275-287, 1953 (No. 5).

waveguides joined together at one end by a butt waveguide joint and a rotatable vacuum joint and closed at the other end by an adjustable non-contacting plunger and a vacuum end plate. To each section of circular guide are attached two rectangular waveguide arms, mutually at right angles and spaced axially by one wavelength in the circular guide. The principle of operation is as follows. Assume initially that equal powers are fed into two arms at right angles. Each will then give rise to an H_{11}-wave in the circular guide as shown in fig. 19*b*. (The radius of the circular guide is chosen so that a reasonably good impedance match is obtained as viewed from the rectangular guide.) The electric fields of both H_{11}-waves will be at right angles and of equal amplitude. Provided the phases are suitably adjusted the resultant polarization will be linear and at 45° to the individual polarizations. If one of the rectangular arms of the other circular section is now adjusted to be at 45° to the first two arms the sum of the two incident powers can be drawn from it; see fig. 19*c*.

When the ratio n of the two incident powers is not unity, the situation is similar but the angle between the resultant polarization and one input arm will now be $\varphi \neq 45°$. This angle is given by $\tan^2 \varphi = n$. By means of the rotatable vacuum joint the output arm from which the power is to be drawn can again be adjusted so that its axis is perpendicular to the resultant polarization. In the 15 MeV accelerator the practical value of n, i.e. the ratio of the power returned to the bridge from the guide to that supplied to the bridge by the magnetron, is $n = 1.28$. The angle φ in that case is about 46.5°. (This value of n will of course be correct only for one value of beam loading; cf. footnote [5]).

It might appear that under these ideal conditions the fourth arm, of the second circular guide section, serves no useful purpose. During the steady state of the pulse, when the power flow through the corrugated guide has reached its final value, no power will in fact enter this arm. However, during the first few tenths of a microsecond the power fed back will be less than the steady state value, and some power will enter the fourth arm. This power is dissipated in a matched load. (A similar situation will obtain even in the steady state when the beam current in the accelerator differs from its design value, but the power loss is small.)

It is convenient to use a matched water load [16]) in the fourth arm since it allows the power dissipated there to be measured during the setting up tests. This load is illustrated in *fig. 20* and some details are explained in the legend.

Comparing the circular magic-Tee with another type of feedback bridge, the rat-race, the former has the important advantage that it can be simply adapted to power addition ratios other than unity without upsetting the impedance match to the rectangular waveguide, viz, by changing the angle between the two sets of arms. On the other hand the rat-race is smaller and all arms are in the same plane, thereby eliminating the "staggered" sections of waveguide required for the circular magic-Tee. Incidentally it may be mentioned that in theory the two arms of each circular guide section could be in the same plane, but experimentally it is found that an appreciable amount of power fed into one arm then enters the arm perpendicular to it.

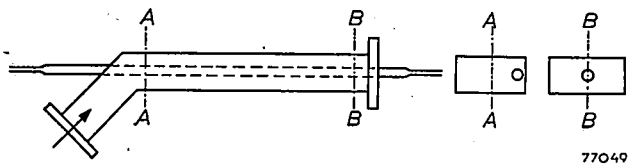


Fig. 20. Matched water load connected to fourth arm of each circular magic-Tee. A glass tube, through which the water flows, enters through a Wilson vacuum seal at one corner (see section *AA*) of a rectangular waveguide with a 45° bend and emerges centrally through a similar seal at the other end which is closed (section *BB*). Since the electric field near *AA* will be small, the discontinuity caused by the glass tube will not be important and little reflection of power will occur.

Phase shifters

The phase shifters used in this equipment consist essentially of a ceramic wedge placed in the rectangular waveguide parallel to its narrow face. The wedge is carried by two shafts and can be displaced perpendicular to the narrow face of the guide. The wedge will cause a delay in the propagation of the wave resulting in a phase change which depends on the position of the wedge and which attains its maximum value when the wedge occupies the centre of the broad face.

The displacement of the ceramic wedge involves the movement of a component inside the evacuated rectangular guide. Further, this movement must be effected by remote control as the phase shifters in the feedback loops may need to be adjusted during full power operation of the accelerator.

A magnetic coupling was developed for moving the phase shifters, which eliminates a vacuum seal on any moving member and therefore requires only a very small torque and the very minimum of maintenance. The coupling makes use of a horseshoe magnet which is supported on a bearing outside the vacuum system and which may be rotated about its axis by remote control. Magnetically coupled to

[16]) Brit. Pat. Application 24823 of 1952.

the magnet but inside a bell-housing attached to the guide and which forms part of the vacuum system is a soft iron pole piece mounted on a threaded shaft which actuates the ceramic wedge carriage. The shafts which carry the ceramic wedge are of silver steel and are separated by an odd multiple of the mean quarter wavelength in the guide in order to minimize mismatch.

Moving parts inside the vacuum are lubricated occasionally when convenient with "Apiezon" pump fluid. A unit of this type incorporated in the 3.5 MeV accelerator has been in operation for three years; it required servicing after about 18 months.

The position of the wedge in the waveguide is indicated remotely on a voltmeter by means of a potentiometer coupled through a gear box to the external rotating shaft. Limit switches are provided in the external driving mechanism to prevent the wedge jamming against the guide wall.

Monitoring thermocouples

The thermocouples used for power monitoring are of the usual waveguide type. Each thermocouple consists of a small pick-up loop in an evacuated bulb. The current induced in the loop by the wave in the guide heats the thermocouple junction, one end of which is connected to a copper foil flange, the other end being brought to an insulated top cap. The thermocouples are used in pairs, spaced by $^3/_4$ of a guide wavelength to reduce errors due to standing waves. The two thermocouples of one pair are connected in series and the polarities are arranged so that the output is taken from the two top caps.

The thermocouple mount is designed to provide a good vacuum seal on the rectangular waveguide and at the same time to allow the replacement of burnt-out thermocouples without an elaborate re-calibration procedure. The mount is shown in *fig. 21*. It will be seen that each thermocouple with its copper foil flange is clamped between a pair of annular discs; this assembly rests in a circular hole in a plate hard soldered to the waveguide and is sealed in position by a collar bolted on to the plate. For calibration a thermocouple pair is mounted in a separate waveguide section where the power flux can be measured directly using a water load. By rotating the disc assemblies the sensitivity of each thermocouple is adjusted so that the correct power values are indicated on the meter. The line of the axis of the guide is marked on the upper disc of each assembly and the thermocouples are transferred in this position to their proper place in the rectangular waveguide system. Using this method the calibration will be accurate to $\pm 5\%$ which is sufficient for the purpose.

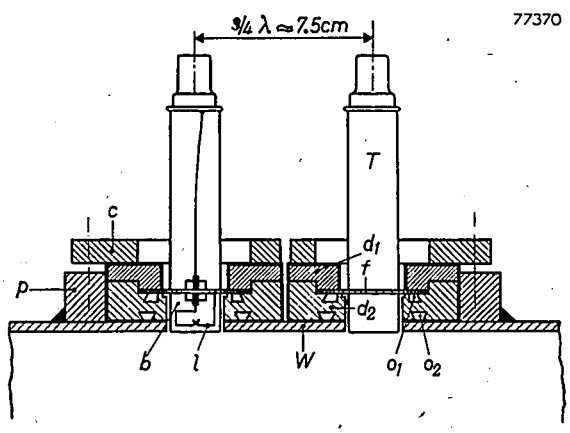


Fig. 21. Waveguide thermocouple mount. The thermocouples T are used in pairs, each being mounted in a rotatable disc assembly so that the sensitivity may be easily adjusted. l = pick-up loop. b = bulb. f = copper foil flange. d_1 and d_2 = discs. o_1 and o_2 = O-ring vacuum seals. p = plate soldered to waveguide wall W. c = collar for clamping the disc assembly in position.

Protective system

With a view to avoiding damage or loss of time in case of failure of any component or supply, a number of safety circuits have been provided. The description of the safety system will be restricted to indicating which components require protection in the event of a given failure, and the general type of indicators and circuits used.

A series of safety measures have been provided for the vacuum system. In the event of mains failure to the diffusion pump the two magnetically operated backing valves are closed so that the vacuum in the envelope is maintained, avoiding waste of time and preventing damage to the hot diffusion pump fluid. If the rotary backing pump ceases to function, due to failure of the mains supply to the pump, of the driving belt or of the motor, a centrifugal switch mechanically connected to the pump shaft switches off the diffusion pump and closes the backing valve; moreover this switch opens magnetically operated air inlet valves situated in the line to each of the two backing pumps, as a protection against oil from these pumps being sucked into the vacuum lines. Of course, the air inlet valve must not open until the backing valves have closed; this is ensured by increasing the time constant of the solenoid of the inlet valve by a large capacitor.

The same series of events will take place in the case of a general mains failure. No other automatic action is necessary in such an event since self-holding contactors are used in the switches so that

power is not re-applied automatically when the supply returns.

A rise of pressure in the vacuum system might also occur due to a vacuum leak while all components in the vacuum system are functioning normally. Such a rise of pressure, apart from cracking the pump fluid, would cause irreparable damage to the ionization gauge filament and the gun filament. Moreover, it would cause sparking in the rectangular waveguide or the corrugated guide and the resultant considerable mismatch would in turn produce severe sparking in the magnetron. To avoid these consequences the Philips gauge mentioned earlier incorporates an electronic relay system which operates if the pressure in the vacuum envelope rises above 10^{-3} mm Hg: this initiates the closing of all vacuum valves and at the same time switches off the modulator, gun filament supply, diffusion pump heater and all vacuum gauges.

The failure of cooling water supplies could cause even more costly damage to the magnetron, to the corrugated guide (due to differential expansion between it and the vacuum envelope), to the electron gun and to the diffusion pump. The water circuits therefore contain bucket-type switches, which in the event of water supply failure interrupt the mains supply.

Lastly, overloads in the three phase supply to the modulator, in the drive circuit, the main high tension circuit and the single phase supply to the focusing coils are monitored by relays. These relays in the D.C. circuits are actuated by the voltage drop across resistors and in the A.C. circuits by current transformers.

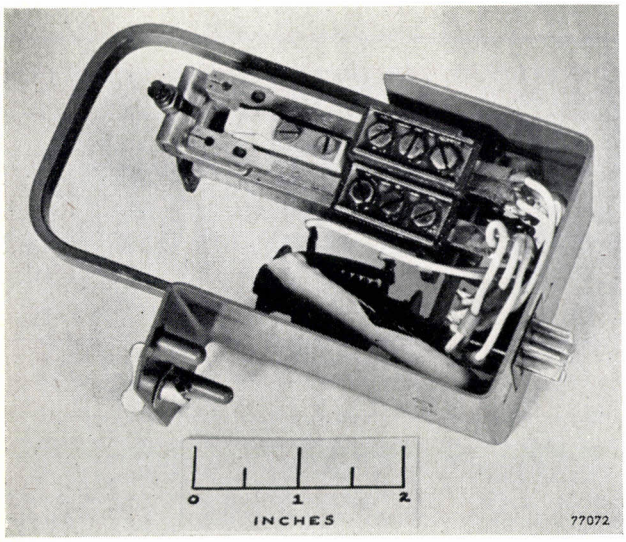


Fig. 22. Relay pack used in the protective circuits. If a relay failure occurs, the pack is unplugged and may be replaced by another one of the same type.

It has been stated that the protective system besides safeguarding the equipment from damage should minimize the loss of time due to failure of components. In any protective system of this extent

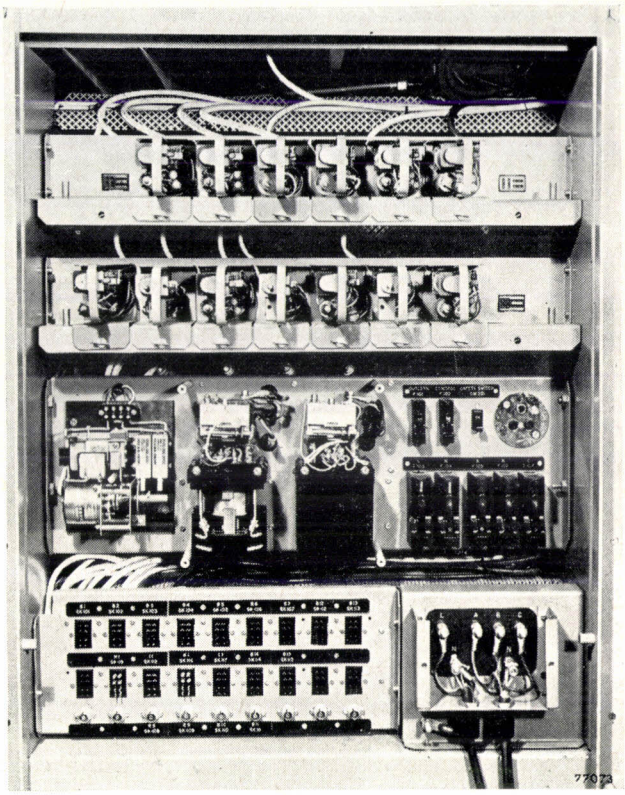


Fig. 23. View of the safety circuit panel of the modulator. Note the set of relay packs in the upper portion.

and intricacy, however, the factor of time loss due to failures in the protective system itself becomes important! The elements most liable to cause trouble are the relays with which most of the protective circuits are associated. Therefore a plug-in relay pack (*fig. 22*) has been developed which incorporates a standard Post Office type relay and which with very minor circuit modifications is used in all these circuits (*fig. 23*). Failure of any relay during service instead of necessitating a lengthy removal and replacement delay, merely requires the unplugging of the faulty pack and replacement by another of the same type. A coding system is used so that it is impossible to plug in a pack of the wrong type.

Control

As the accelerator shelter is not accessible while the apparatus is running, all adjustments required during operation must be made by remote control. In view of its position it is convenient to be able

to control the modulator remotely also[17]). The control desk (*fig. 24*) has been designed to carry out this dual purpose. It consists of seven easily removable panel units:

1) Main control panel. This includes push buttons for switching on valve heaters and anode supplies and for switching on and varying the high voltage output from the modulator. Meters are provided to monitor the voltage and current from the main rectifier circuit in the modulator. Also attached to this panel is the oscillator which generates the voltage pulses for controlling the modulator. Pulse repetition frequencies of 100, 200, 350 and 500 pulses per second may be selected by means of a switch. Alternatively an external source of triggering pulses may be connected by means of the same switch.

2) Magnetron control panel. This provides facilities for control and measurement of filament current, magnetic field, radio frequency power and frequency.

3) Controls more directly concerned with the electron beam. Here facilities for control of the gun filament current, for the measurement of the total power in the electron beam and of the X-ray output and finally for the control of phase and measurement of radio frequency power in each of the feedback loops are found.

4) Oscilloscope unit. This has a time base of either 12 μsec, 120 μsec or 12 msec. It may be used to monitor pulse voltage and current waveforms in the modulator or, for example, the current pulse produced by the scintillation from a crystal under the action of the X-rays generated. It also includes a 60 dB pulse amplifier connected to the output of a high-Q cavity wavemeter for frequency monitoring; this method of frequency measurement is useful since it demonstrates frequency changes during the magnetron pulse or the presence of even a very small amount of unwanted modes in the magnetron oscillation.

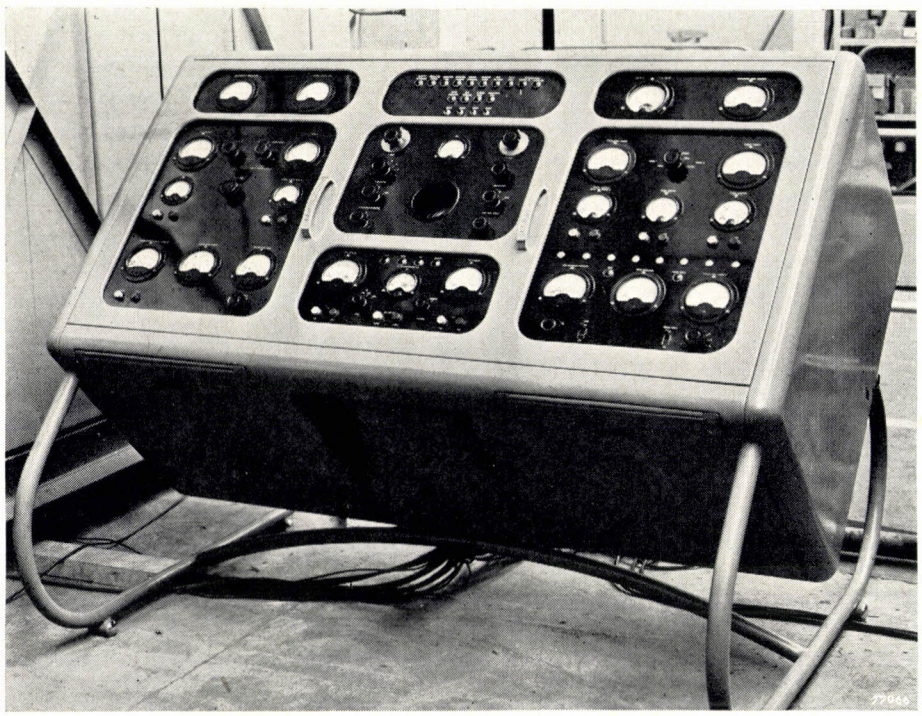


Fig. 24. Control desk of 15 MeV accelerator. It includes a main control panel (lowest centre), magnetron control panel (lower left hand), controls concerned with the electron beam (lower right hand), oscilloscope unit (middle centre), safety indicator panel (uppermost centre) and two subsidiary panels for checking the condition of the vacuum system and for various other meters.

5) Safety circuits panel. This indicates the state of the various safety circuits by a series of green lamps. A red lamp is associated with each of the four high power circuits which are protected against current overload as explained in the previous chapter. When the overload protection operates, the red lamp lights and after the fault has been noted and cancelled by means of a push button on the

[17]) In certain medical applications the modulator will be placed beside the accelerator, so that remote control is unavoidable.

same panel, the circuit can be switched on again.

6) and 7) Finally two small panels provide a remote indication of pressure from the ionization gauge circuit, of the temperature difference between the main and the refrigerated water supply and any other indications which may be required in a particular installation. At Harwell three hour-meters (not shown in fig. 24) indicating low tension circuits, high tension circuits and refrigerator compressor hours were installed in one of these small panels. The first two assist in compiling information on the valve life to be expected in various circuits.

C. TESTING AND PERFORMANCE

Measuring methods

The testing of the machine as a particle accelerator primarily involved the measurement of the electron energy and the beam power. No attempt was made to measure the beam current directly since this is more difficult and without very carefully designed measuring gear would yield no more than a very rough cross check on the value of current derived from the other two measurements.

Measurements of the X-ray intensity obtained and of the beam concentration (i.e. the radial distribution of the electrons in the beam) were also made. These characteristics will of course be of special importance when a similar machine is used for X-ray therapy.

The *total beam power* was measured by means of a constant flow water calorimeter. This consisted of a cylindrical brass container long enough to absorb all electrons and closed by a brass plug at one end and a thin copper foil at the other. The container was supported coaxially with the accelerator and with the copper foil facing the electron source so that the electrons tended to give up their energy in the water rather than in the walls of the container. The temperature rise of the water was monitored by a differential thermocouple. In these experiments the electrons emerged from the accelerator through a hole covered by a 0.002 inch (0.05 mm) beryllium copper foil. The loss of energy in this foil is very small but it causes considerable scattering. The calorimeter was therefore placed as close as possible to the foil to minimize the air gap and to prevent the loss of scattered electrons.

By adjusting the rate of flow of water so that the inlet and outlet temperatures differ by not more than 10 or 15 °C, and by arranging that these temperatures in so far as possible lie below and above ambient temperature, the heat loss to the air is made negligibly small.

The *electron energy* is determined, by the conventional method, from the radius of curvature of the electron path in a known magnetic field which is at right angles to the direction of motion. A schematic layout is shown in *fig. 25*. Two steel plates F_1 and F_2 with slits for the passage of the electrons are set up at right angles. Electrons from the accelerator pass through the slit S_1, traverse a 90° arc between the pole pieces of an electromagnet (required field intensity about 0.2 Wb/m²) and emerge through the slit S_2. After traversing a drift space 30 cm long they pass through a further slit S_3 to a collector. The current obtained from the collector is proportional to the number of electrons that leave the accelerator with an energy corresponding to the chosen value of the analyzing magnetic field.

The current pulse is amplified by a screened preamplifier close to the collector and then by a high-gain balanced amplifier; the output pulse is displayed on a high speed oscilloscope. In this way the duration of the electron pulse and the time distribution of electron energies occurring in the pulse may also be measured.

Since the electrons traverse a path of the order of 1 m length between the accelerator output window

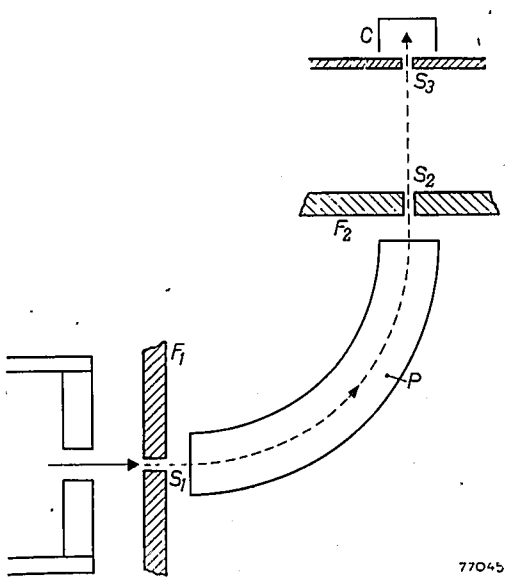


Fig. 25. Set-up for measuring the electron energy. The electrons emerging from the exit window of the accelerator pass through the slit S_1 and enter the magnetic field between the pole pieces P of an electromagnet. The field is adjusted so that the electrons describe a 90° arc and pass through the slit S_2 to collector C. (Of course, the magnetic field need be established only along the actual electron path, so that the pole pieces take the form of 90° arcs.)

and the collector, the scattering — if the path were at atmospheric pressure — would be considerable and the current pulse at the collector would be very small indeed. To avoid this the electrons pass through a similar beryllium copper foil window at S_1 and enter a tubular vacuum system which extends right up to the collector. The resultant scattering in the two foils and in the 1 inch path in air between the foils is not very serious since the pulse current is large and the absolute current value does not matter in these measurements.

Of course it would have been possible to avoid all scattering by joining the tubular vacuum system directly to the main vacuum system, which at the same time would eliminate the necessity of separate pumping. This scheme was not adopted, however, because of the physical difficulty of aligning accurately the considerable mass of the magnet and its associated trolley to the vacuum envelope. Moreover, it would have made it impossible to measure both the electron energy and beam power inside a short time interval (see below).

Another possible source of error has to be considered in the energy measurement. The mild steel plates F_1 and F_2 are used to eliminate the effects of the stray field of the analyzing magnet on the electron beam when it is near the target end of the corrugated waveguide. After passing through the slit S_1, however, an electron will be under the influence of the stray field of the magnet since the steel plates must be mounted some few inches away from the edges of the pole pieces to avoid affecting the uniformity of the analyzing magnetic field inside the pole gap. In order to allow for the slight deflection of the electrons in the stray field the magnetic field was measured along the path between S_1 and S_2 and the actual path of electrons entering and leaving with path normal to F_1 and F_2 was determined by numerical integration. The required correction can be expressed by an effective radius of curvature which is greater than the mean radius of the magnet poles.

The value of the mean *beam current* is derived from the measurements of the electron energy and the total beam power. The pulse current may then be calculated from the measured pulse duration. In order to obtain the correct current value it is necessary to ensure that conditions of operation do not change between the measurements of energy and beam power. The general procedure adopted therefore was to have as short a time delay as possible between these measurements. Since it requires quite some time to align the analyzing magnet to the system, the energy measurement was carried out first, followed as quickly as possible by the power measurement.

Measurements of the diameter of the electron beam emerging from the accelerator can be made by placing a piece of photographic paper in the beam. The paper quickly darkens under electron bombardment and gives a black central portion where the electron density is highest, surrounded by a gradually decreasing coloration. This method is quite useful for centering the beam and indicates that the electron density is high inside a diameter of 6 to 8 mm, but of course it gives only a qualitative estimate of the *radial electron distribution.* The system adopted to get a quantitative estimate is shown in *fig. 26.* To the end plate of the accelerator is attached a series of copper discs in which progressively smaller holes of $^3/_4$, $^3/_8$, $^1/_4$ and $^1/_8$ inch diameter are drilled. These discs divide the electron beam into a central core of $^1/_8''$ diameter and coaxial tubes with inner and outer diameters of $^1/_8$ and $^1/_4''$, $^1/_4$ and $^3/_8''$; and so on. A single turn cooling spiral is soldered to each disc and allows the relative powers in the various sections of the beam to be measured. The radial power distribution is also the radial current distribution, since there is no radial variation of electron energy.

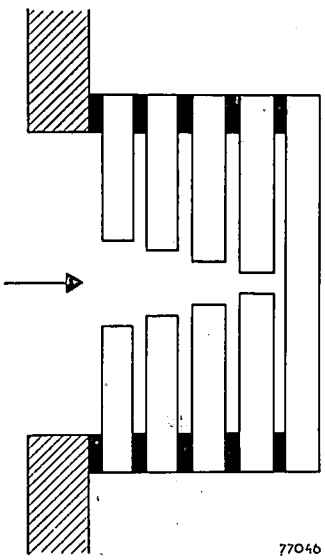


Fig. 26. Equipment for measuring the radial variation of electron current density in the beam. The beam strikes on a series of copper discs attached to the accelerator end plate, each thermally insulated from its neighbour and with an O-ring vacuum seal between adjacent discs. Progressively smaller holes are drilled in the discs, dividing the beam into a number of coaxial "tubes". The relative power in each "tube" is measured by means of a single turn cooling spiral.

Finally measurements of the *X-ray intensity* from a 4 mm platinum target were made using a concentric cylinder ionization chamber with 200 volts across a 1 mm gap [18]).

[18]) These measurements were made by Mr. G. S. Innes of St. Bartholomews Hospital, London.

Performance

The maximum energy obtained from the accelerator at low beam currents lay always between 14 and 15 MeV with a tendency towards the lower figure. The actual output achieved under any given conditions depends of course to a very marked extent on the frequency of operation, the magnetron power and the care with which the feedback loop phase shifters and the phase shifter between the two loops have been adjusted.

at the base and 1.1 μsec at the top giving a mean value of 1.3 μsec. The measured length of the current pulse through the magnetron was 1.85 μsec and the radio frequency power pulse will not be very much less. This appears to indicate that the time required to build up energy in the corrugated waveguide is of the order of 0.5 μsec.

The electrons emerging from the accelerator are not strictly mono-energetic but the spread of energy is reasonably small. The spread varies during

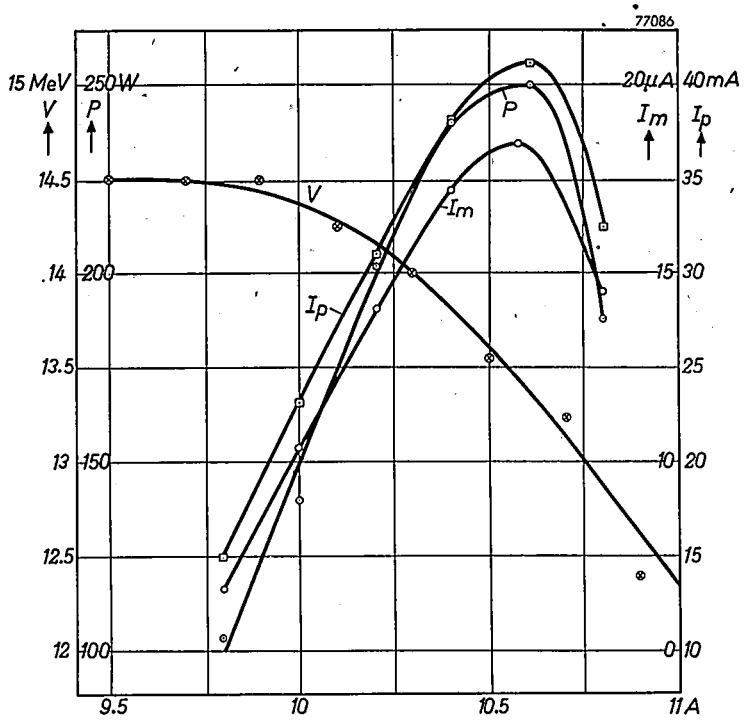


Fig. 27. Variation of electron energy V, beam power P and beam current (mean value I_m, pulse value I_p) with gun filament current (i.c. with beam loading) as measured in experiments with the 15 MeV linear accelerator. The magnetron pulse power was slightly more than 1.8 MW, frequency 2999 Mc/s, electron pulse duration 1.3 μsec, pulse repetition frequency 350 per sec. The measurements were obtained after reasonably careful adjustment of the phase shifters. It is seen that the actual performance falls somewhat short of the predicted values [19]) (15 MeV at 25 mA).

Fig. 27 shows graphically a typical set of measurements obtained after reasonably careful adjustments and with a magnetron which delivered slightly more than 1.8 MW in the pulse. It will be noted that with increasing gun filament current the beam power increases to a maximum and then falls off rapidly. This is due to the beam loading and consequent decrease of peak accelerating field which probably causes electrons to be slowed down to such an extent that they fall back behind the wave crest (region of peak accelerating field) and are lost.

The length of the electron pulse as measured on the oscilloscope display was found to be 1.5 μsec

[19]) Cf. also: C. W. Miller, Nature **171**, 297-298, 1953 (Febr. 14).

the pulse and is approximately 7% (total width of spectrum) when the current has attained half its peak value. The spread is mainly accounted for by the occurrence of two quite distinct energy levels which are separated approximately 0.25 MeV in energy. Electrons of the lower energy level occur mainly in the first part of the pulse whereas electrons of the higher level are concentrated in the latter part. This is clearly demonstrated by *fig. 28* showing the appearance of the current pulse as the analyzing magnetic field is varied. The exact extent of the dip in the centre current pulse is difficult to determine owing to the finite resolving power of the analyzing system. The effect has also been noted by other workers in the field on other

travelling wave accelerators. The authors are unable to advance any satisfactory explanation.

The radial variation of the current as measured by the series of copper discs and the calculated radial variation of current density are plotted in *fig. 29*. It will be seen that just under 60% of the current is contained inside a 5 mm radius and 50% inside a 4 mm radius.

The variation of X-ray intensity (dose rate) with gun filament current is shown in *fig. 30*. The beam current estimated from the power dissipated in the X-ray target is also given in this figure, to show the close correlation between current and X-ray

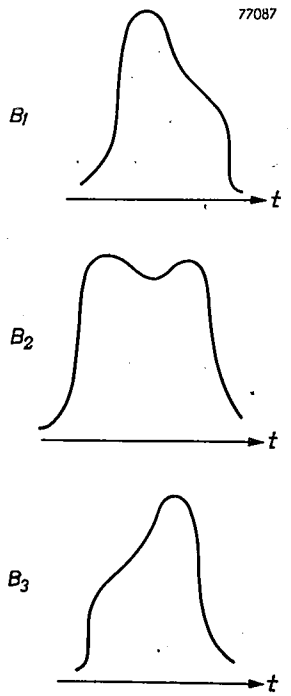


Fig. 28. Appearance of current pulse as obtained at the collector of the energy analyzer (fig. 25), for different values of the analyzing magnetic field ($B_3 > B_2 > B_1$). The oscillograms demonstrate the occurrence of two quite distinct energy levels, differing by about 0.25 MeV (as derived from the difference between B_3 and B_1).

intensity; the values of the current should not be taken as absolute values since there is considerably more heat loss associated with this measurement than with the water calorimeter. The efficiency of X-ray production in the heavy metal target will also be considerably greater than in the water target, i.e. more of the electron beam power will be in the X-rays emitted. In fact the current values appear to be about 20% low.

It will be seen from the graph that the radiation intensity peaks just before the current. This is as would be expected since the radiation intensity is proportional to the current and proportional to approximately the 2.7th power of the electron

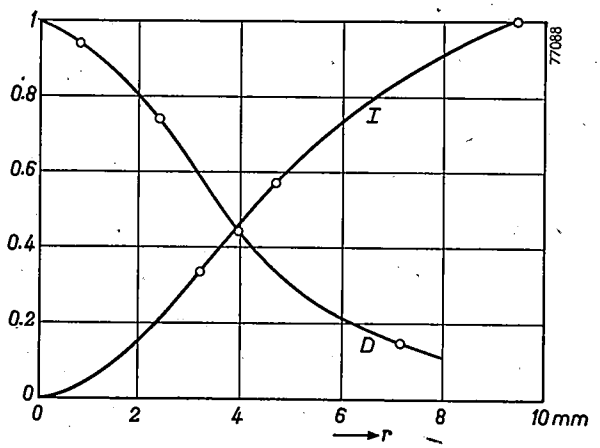


Fig. 29. Radial distribution of electrons in the beam. Curve *I*: relative current inside radius *r*. Curve *D*: relative current density at radius *r*.

energy. The energy is falling quite rapidly in the region of the beam current maximum.

All the measurements described above were made at a repetition frequency of 350 pulses per second. At Harwell the normal operating frequency is 400 pulses per second and in view of the danger of damage to the glass window of the magnetron, this must be considered a maximum for running periods in excess of 30 minutes or so. For short periods the magnetron may be operated (in this accelerator) at 500 pulses per second, when the maximum X-ray intensity is in the region of 2250 röntgen/minute at 1 metre from the target.

Acknowledgements

The thanks of the authors are due to the Director of the Atomic Energy Research Establishment and

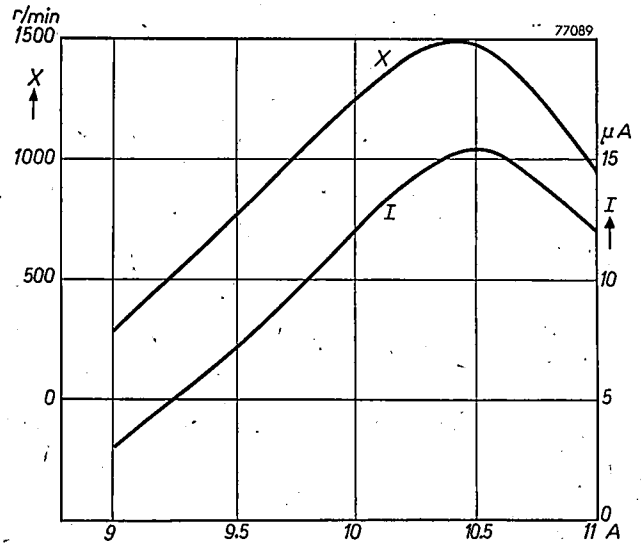


Fig. 30. Variation of the *X*-ray output (in röntgen/min at 1 m distance from a 4 mm platinum target) with gun filament current (curve *X*). The curve *I* represents the variation of beam current estimated from the power dissipated in the target cooling water.

to the Directors of the Mullard Radio Valve Company for permission to publish this article.

The authors also wish to thank friends at A.E.R.E. for their valuable cooperation, especially E.R. Wiblin and B. L. Tozer who assisted greatly during the construction and installation of the accelerator, and lastly their colleagues at Mullard Research Laboratories without whose untiring efforts this project could not have been completed in two years.

Summary of parts A, B, C. The linear electron accelerator described in this article was destined for nuclear physics work at the Atomic Energy Research Establishment, Harwell. Its design was based on a performance of 15 MeV electron energy and 25 mA beam pulse current, with a corrugated waveguide of length about 6 metres, and using a magnetron operating at 3000 Mc/s with 1.8 MW pulse output. The electron energy to be expected in a system with feedback was computed for different values of the guide parameter a/λ, using a formula which takes into account the heavy beam loading. As a result it was decided to divide the waveguide into two sections each of 3 metre length, each with its own feedback loop, and to use a value $a/\lambda = 0.168$, the electrons travelling at a stable phase 35° in front of the wave crest. In this way, since it is possible to adjust the phase between wave and electron after the first waveguide section, the frequency sensitivity of the accelerator is reduced to a workable value, having regard to the magnetron stability and practical machining tolerances.

To ensure easy assembly, dismantling and re-assembly without re-adjustment, the corrugated waveguide with its associated parts is mounted on six three-wheeled trolleys running on rails. The trolleys if necessary can be supported from above, thereby facilitating a conversion of the design to a machine for medical use. After a brief survey of the characteristics of the main components of the complete installation, a number of components are described more fully. This description includes the corrugated waveguide which consists of 316 cells, and whose machining and watercooling are discussed; — the electron gun with vacuum valve, and with gun filament supply stabilized by means of a "magnetic diode" circuit; — the focusing coil system surrounding the guide and divided into six sections, fed from six independent power supply units whose stabilized output current may be adjusted to give maximum electron beam current; — the vacuum system, consisting of a diffusion pump of 1500 litres/sec unbaffled pumping speed, two rotary backing (or roughing) pumps controlled by magnetically operated valves, and an extensive gauging and protective system; — the radar type modulator placed on top of the concrete shelter in which the accelerator is housed and feeding 40-50 kV pulses of 2 μsec duration with a maximum jitter of 0.1 μsec to the magnetron and to the electron gun; — the rectangular waveguide system evacuated to a low pressure to prevent sparking and incorporating the two feedback bridges (circular magic-Tee's, that are easily adaptable to the power addition ratios other than unity required in this system), the phase shifters which are adjusted by remote control using a specially developed magnetic coupling in order to avoid vacuum seals on moving members, and the power monitoring waveguide thermocouples required for adjusting the phase to give maximum power flux; — finally the general lay-out, the details of the protective system and the control and measuring equipment on the control desk.

The last part contains a description of the methods used for measuring the electron energy, the beam power (the beam current being derived from these two), the energy spectrum, the X-ray output and the radial variation of electron density in the beam. The energy obtained with an electron beam current of 25 mA when the phases in the feedback loops are carefully adjusted is between 14 and 15 MeV. 50% of the beam power is concentrated on the target within a circle of 8 mm diameter. The maximum X-ray output measured at 1 m distance from the target is 2250 röntgen per minute.

A LARGE-SCREEN TELEVISION PROJECTOR

by J. HAANTJES and C. J. van LOON. 621.397.62:535.88

Television pictures suitable for presentation in cinemas have to be about 10× larger than those given by domestic receivers. The problems of producing such a large image have been surmounted, as will be seen from the following description of the large projector designed by Philips. Prototypes of this equipment have for some years been available for demonstration to visitors to the Philips Research Laboratory and the E.L.A. Studio (a studio for the recording and reproduction of sound and for film projection). The projector has now been in actual production for some time.

In some countries television has made great strides, notably in the United States where the number of receivers in use is estimated at 20 million; in Great Britain there are some 2 million sets in use. There is little doubt that rapid expansion will follow in other countries and that, there too, television is due to play a growing role in the field of entertainment.

There are no longer any technical reasons why the use of television should be restricted to private entertainment in the home. The fact that it is possible to produce images several yards in width offers considerable scope for the extension of television applications. One such application is in the cinema. In the United States many cinemas are now equipped for showing television pictures [1]), but further developments in this direction depend on non-technical factors; these will not be discussed here.

An entirely different use for television arises from events which take place in locations where the public cannot be admitted, but which large numbers of people should nevertheless have an opportunity of seeing. In such cases the television camera and a large projector can be employed in a "closed circuit". A typical example of this field of application is in college hospitals, where the medical students are thus enabled to see on the screen in the lecture room an operation being performed just as though they were standing beside the surgeon. In 1949 Philips had already demonstrated this method at the University at Leiden [2]).

The size of the picture in the ordinary domestic television receiver has increased considerably since 1948. In that year the screen of the largest tube then in use was 30 cm in diameter, but there was much demand for a larger screen. However, as in other spheres (the automobile industry for example), European techniques have not met the demand along the same lines as in the United States. There, the most popular screen diameter is 50 cm, the largest being 75 cm. In Europe the largest "direct vision" tube normally obtainable is 43 cm. For pictures larger than this, projection television is normally used. In this system only a small tube is used, the very bright image thus obtained being enlarged to a considerable size by optical means. As far back as 1948 Philips had brought this system to a high state of perfection [3]) and, for private viewing, both systems are now in use. It is not yet certain which of the two systems will become the more popular.

Television for large audiences is another matter; a picture several feet in width is then required and projection is the obvious solution; many difficulties had to be surmounted, however, before satisfactory results were obtained.

The "Mammoth" projector, designed by Philips for use in halls (type No. EL 5750) gives a picture 4 × 3 metres in size using a tube only slightly larger than the kind found in home projection sets, in conjunction with an optical system with wide aperture.

The main difficulty in this system is that of obtaining sufficient light; this question is therefore considered first.

The necessary luminous flux

In making an estimate of the luminous flux required from the cathode-ray tube to give satis-

[1]) According to a recent estimate this number has increased from 1 in 1949 to 75 in 1952, divided among 37 cities. See N. L. Halpern, Theater television progress, J. Soc. Mot. Pict. Telev. Engrs. **59**, 140-143, Aug. 1952.

[2]) Philips tech. Rev. **11**, 42, 1949-50.

[3]) See the series of articles: A Projection Television Receiver, Philips tech. Rev. **10**, 1948-49:
I. pp. 69-78, P. M. van Alphen and H. Rinia, The optical system for the projection.
II. pp. 97-104, J. de Gier, The cathode-ray tube.
III. pp. 125-134, G. J. Siezen and F. Kerkhof, The 25 kV anode voltage supply unit.
IV. pp. 307-317, J. Haantjes and F. Kerkhof, The circuits for deflecting the electron beam.
V. pp. 364-370, J. Haantjes and F. Kerkhof, The synchronisation.

factory results in a darkened hall we cannot do better than assume the operating conditions in cinemas. A film projector with shutter rotating, but without film, is considered to give satisfactory screen luminance [4]) when this is equal to 34 nit [4]). To ascertain the luminance when the film is being exhibited this value must be multiplied by the transmission factor of the film; at the most transparent parts this is seldom more [5]) than 0.5. A television image having a maximum luminance of 17 nit can in this respect be regarded as on a par with cinema pictures.

Let us denote the area of the projection screen in square metres by S. If the screen reflects wholly in accordance with Lambert's law, that is to say if it diffuses incident light in accordance with the cosine law within a solid angle of 2π, a luminous flux of $17\pi\ S$ would have to reach it to produce a luminance of 17 nt.

It is preferable to work with average instead of i.e. one which reflects more in the direction of the normal and less towards the wings than would be in accordance with Lambert's law. This process cannot be carried too far, however, as the solid angle within which observers see the image sufficiently brightly is then too small.

In our own experience optimum selectivity is that which produces a gain in luminance by a factor of 3 in the direction perpendicular to the screen and, with a screen of this kind, an incident luminous flux of $^1/_3 \times 0.3 \times 17\pi\ S \approx 5\ S$ lumens is sufficient.

This luminous flux represents only a part of the flux that has to be emitted by the cathode-ray tube, as some of it is lost in the optical system used for the projection. Even with a system having a very large aperture such as the Schmidt system, about which we shall speak presently, only some 30 % of the emitted light reaches the screen. Hence the cathode-ray tube must be capable of yielding $5\ S/0.3 \approx 17\ S$ lumens.

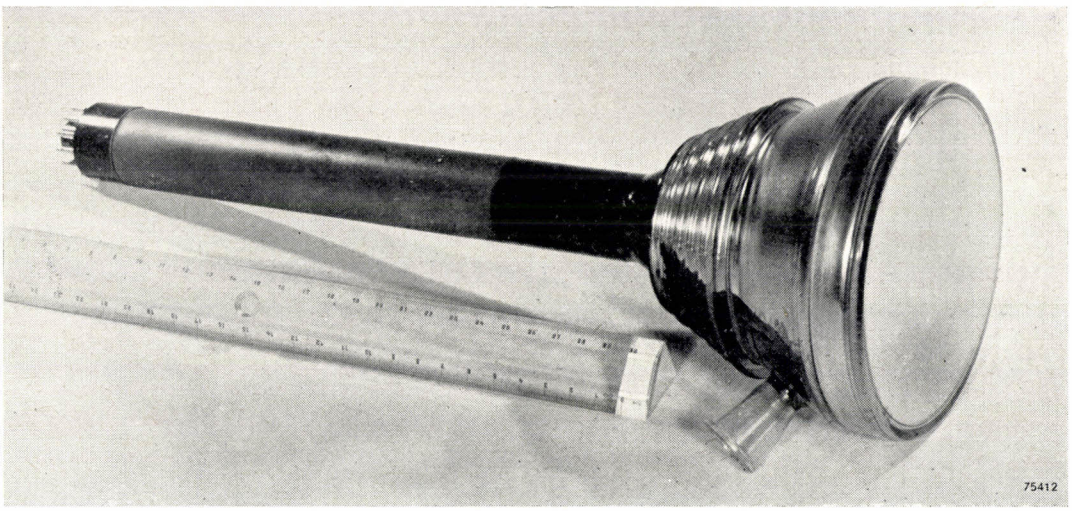


Fig. 1. Cathode-ray tube type MW 13-16 for large projection of television pictures. Screen dia. 12 cm. The H.T. terminal is seen on the lower right: the glass funnel and the ribbed cone of the tube serve to increase the creepage paths.

maximum values, as these are more easily measured. The average luminance can be assumed to be about 30% of the maximum, so that the average luminous flux would have to be $0.3 \times 17\ \pi\ S$.

Appreciably less is needed if use be made of a screen having a selective reflection characteristic,

[4]) The Commission Internationale de l'Eclairage recommends the use of the term "luminance" instead of brightness, and as units of luminance the nit and the stilb (abbrev. nt and sb): 1 nt = 10^4 sb = 1 candela (international candle) per m^2 = 1 lumen per m^2 per unit solid angle. The relationship between this and the much used units apostilb and footlambert is: 1 nt = π apostilb = 0.292 footlamberts. Hence 34 nt = 107 apostilb = 10 footlamberts.

[5]) A. Cazalas, Ann. Télécomm. **5**, 298-306, 1950.

The projection tube

The cathode-ray tube with which the Mammoth projector is equipped is the MW 13-16 (*fig. 1*), of which the effective screen diameter is 12 cm, the working voltage being 50 kV. The beam current may reach a peak value of 2 mA, with an average of 0.5 mA; hence the rated average load on the fluorescent screen is a maximum of 25 W. Forced air-cooling is necessary to prevent the screen from cracking.

The fluorescent screen consists of silicates, all of which fluoresce in different colours but which, when mixed in certain proportions, emit white light

(colour temperature 6500 °K). A layer of aluminium is applied to the inside of the screen ("metal backing") to reflect the fluorescent light emitted rearwards, thus ensuring that this light is not lost [6]). With the metal backing the efficiency of the particular silicate mixture used is 2.5 candela/W, so that the MW 13-16 tube is capable of supplying a luminous flux of $2.5 \times 25 \times \pi \approx 200$ lm.

According to the estimate in the previous paragraph a screen area S is required to give $17S$ lumens, and this tube therefore emits enough light for a projection screen of area $S = 200/17 \approx 12$ m², i.e. 4×3 metres.

At the brighter parts of the image the beam current momentarily exceeds the average of 0.5 mA, and the tube is then working at a point where the fluorescence exhibits saturation with any increase in current [7]), i.e. the light yield is no longer proportional to the increase in current. At the peak value of the current (2 mA) the luminous flux is 600 lm.

The focusing of the electron beam must be very precise, as will be appreciated when it is remembered that the image on the screen of the tube is only 72×96 mm in size. Deflection of the beam is accompanied by a certain amount of defocusing: it is essential, however, that the light spot be quite clearly defined at the edges of the picture. Efforts can be made to compensate for deflection defocusing by making the focusing current vary in a certain manner with the deflection current (both focusing and deflection are electromagnetic), but it is considered preferable to eliminate the defect at its source as far as possible. Deflection defocusing is dependent on the three-dimensional distribution of the deflection fields; for this reason much thought has been given to the design of the deflection coils so that they will produce the best possible field distribution. Furthermore, deflection defocusing is reduced according as the electron beam is reduced in width at the points where the deflection fields operate. Hence the narrowest possible beam is employed.

In this way deflection defocusing has been reduced to a minimum and no compensation is necessary.

The optical system

The optical system must be capable of producing on the projection screen a highly magnified image of the picture on the tube; in the present case this represents a linear magnification of roughly 40×.

Special consideration must be given to the following points in the design of such a system.

1) The system must gather as much light as possible from the cathode-ray tube.
2) The quality of the image must be such that a television picture of 625 lines is not noticeably reduced in contrast or detail.
3) The dimensions of mirrors and correcting plates should be as small as possible, as the cost of these components rises very steeply with the size.

Reflecting system

A system which most nearly approaches these conditions is the Schmidt system, which has already been mentioned in numerous articles in this Review [8, 9]). It consists of a concave spherical mirror and an aspherical correcting plate mounted at the centre of curvature of the mirror (*fig. 2*). The main advantage of a mirror as compared with lenses as principal component is that mirrors are free from chromatic aberration and, in contrast with equivalent lens systems, exhibit less spherical aberration and are much less costly. Image defects such as coma and astigmatism are avoided by mounting a diaphragm at the centre of curvature [10]).

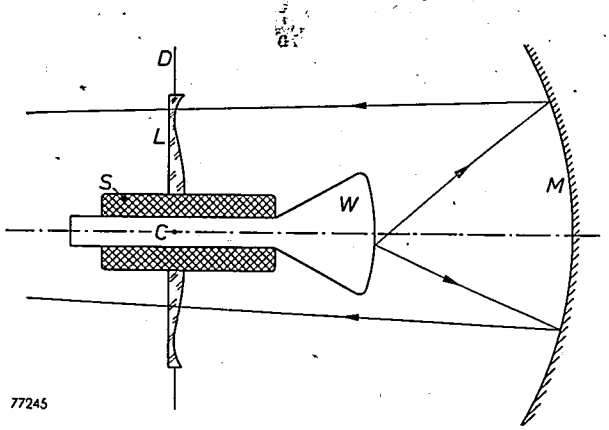


Fig. 2. Diagrammatic cross-section of the optical system. W cathode-ray tube with focusing and deflection coils S. M spherical mirror; L Schmidt correcting plate; D diaphragm; The planes of D and L pass through the centre of curvature C of the mirror.

The only defects of the third order then remaining are the spherical aberration and field curvature; the first is removed by using the correcting plate already

[6]) See article II, referred to in footnote [3]), in particular pp. 101-103.
[7]) A. Bril and F. A. Kröger, Philips tech. Rev. **12**, 120-128, 1950.
[8]) H. Rinia and P. M. van Alphen, Philips tech. Rev. **9**, 349-356, 1947-48.
[9]) See article I, referred to in footnote [3]).
[10]) See e.g. W. de Groot, Optical aberrations in lens and mirror systems, Philips tech. Rev. **9**, 301-308, 1947-48.

mentioned, and the second by curving the light-emitting surface, i.e. the screen of the cathode-ray tube, to a radius equal to the focal length of the system, and mounting it so as to be practically concentric with the mirror.

Optical systems made along these lines have a very wide aperture, they are relatively inexpensive to produce and, on plane surfaces, they give an image that is subject only to defects of the fifth and higher orders, and to an imperceptible chromatic aberration in the correcting plate.

Light-gathering power

We shall now determine the extent of the light-gathering power, that is, the ratio of the luminous flux entering the optical system to that emitted by the object, in this case the screen of the cathode-ray tube. For the moment we shall disregard losses of light due to absorption, imperfect reflection and masking effects: in other words we shall assume that the emitted luminous flux is equal to the incident flux.

To do this we make use of Abbe's law which states that when a luminous object is observed through a loss-free optical system, the luminance of the image as observed is equal to that of the object.

Let us denote the luminance of the image on the cathode-ray tube by L, the illumination on the projection screen by E, the diameter of the correcting plate by D and the distance from this plate to the projection screen by d. Then, in accordance with Abbe's law:

$$E = \frac{\frac{\pi}{4} D^2}{d^2} L.$$

The total luminous flux on the projection screen (area S_2) is therefore:

$$E S_2 = \frac{\frac{\pi}{4} D^2}{d^2} L S_2. \quad \ldots \quad (1)$$

According to simple laws of geometrical optics the following relationship exists between d_1, the focal length f and the linear magnification M:

$$d = (M-1)f, \quad \ldots \quad (2)$$

whilst

$$S_2 = M^2 S_1, \quad \ldots \quad (3)$$

where S_1 is the area of the image on the cathode-ray tube. Substituting (2) and (3) in (1), we then have:

$$E S_2 = \frac{\pi}{4} \left(\frac{D}{f}\right)^2 \left(\frac{M}{M-1}\right)^2 L S_1,$$

for which, applied to the present case, where $M \gg 1$, we may write:

$$E S_2 = \frac{\pi}{4} \left(\frac{D}{f}\right)^2 L S_1. \quad \ldots \quad (4)$$

It will be seen that, of the two parameters D and f which we have introduced in relation to the optical system, only their ratio, the so-called f-number D/f appears in equation (4). This equation states that, in order to obtain the optimum amount of light from the system, D/f must be as large as possible. At the same time the quality of the reproduction imposes a limit on D/f which experience shows to be roughly 1.4.

Another conclusion that may be drawn from (4) is that, since the luminous flux emitted by the cathode-ray tube is $\pi L S_1$, the fraction of this that reaches the optical system (i.e. the light-gathering power) will be $\frac{1}{4}(D/f)^2$. Taking D/f to be 1.4, this will be 0.5. (As explained in article I, referred to in footnote [3]), the light-gathering power is equal to the square of the numerical aperture, which in this case is 0.7.)

In practice the system is not free from losses; light is lost at the mirror, whose reflection coefficient is about 0.85; the cathode-ray tube itself intercepts some of the reflected light and the correcting plate reflects slightly. Consequently the luminous flux from the optical system is not more than 30% of the flux emitted by the cathode-ray tube, viz. the value which we have already employed in the preceding section.

Some of the light lost emerges from the system as scattered light, and this is a disadvantage as it reduces the contrasts in the image; light-scatter must therefore be avoided as much as possible, and this is done in the usual way by giving the interior of the optical system a dull black finish.

Dimensioning of the optical system

In order to ensure that the concave mirror will not limit the effective cone of light, the diameter of this mirror must be equal to at least $2f$. We have already seen that the diamter D of the correcting plate should be roughly $1.4f$.

The principal dimensions of the optical system are thus proportional to the focal length f. Hence to ensure compactness we shall want f to be as small as possible. On the other hand f cannot be made too small, as the image angle then becomes so large that the quality of the picture suffers, especially at the edges, where optical defects of the 5th and higher orders are mainly in evidence. In the Schmidt system the practical limit occurs at $f = 1.7$ times

the maximum dimensions of the object which, in our case, is the diameter of the fluorescent screen of the tube.

This diameter is itself subject to a lower limit. If the screen diameter — i.e. the size of the fluorescent image — is too small, one or more of the following difficulties is encountered:

1) Light saturation of the phosphor — even with the beam current below the maximum limit, the energy density on the light spot may be too high.
2) Heat dissipation — in a smaller area, simple air-cooling would not be adequate — the screen would deteriorate rapidly [11]) or even be in danger of cracking.
3) Size of light spot — to obtain a sufficiently small light spot, higher tube voltages would be necessary.

As a compromise between optical requirements and tube limitations a screen diameter of 12 cm was adopted.

To produce a picture diagonal of 5 m on the projection screen the magnification will be $M = 500/12 \approx 40\times$. From equation (2) it follows that the distance d from the optical system to the screen should then be about 8 m. Should the circumstances warrant a greater distance, the optical system would have to have a greater focal length. This would not affect the quality of the picture, but the equipment would have to be larger and more costly, without yielding the slightest gain in the luminance of the projected image, as will be seen from equation (4). (As will be shown presently, the design of the equipment is such that the need for a larger projection distance is unlikely to arise.)

The correcting plate

With a given correcting plate the Schmidt system gives good reproduction only when the magnification does not differ by more than 10% from the designed value. For greater variation of the enlargement different correcting plates are necessary. The Mammoth projector is supplied with a correcting plate adapted to the chosen projection distance; these plates are made from "Perspex" by turning them on a lathe.

Fig. 3 depicts the optical system in the projector,

[11]) A. Bril and H. A. Klasens, Intrinsic efficiencies of phosphors unter cathode-ray exitation, Philips Res. Rep. **7**, 401-420, 1952 (No. 6).

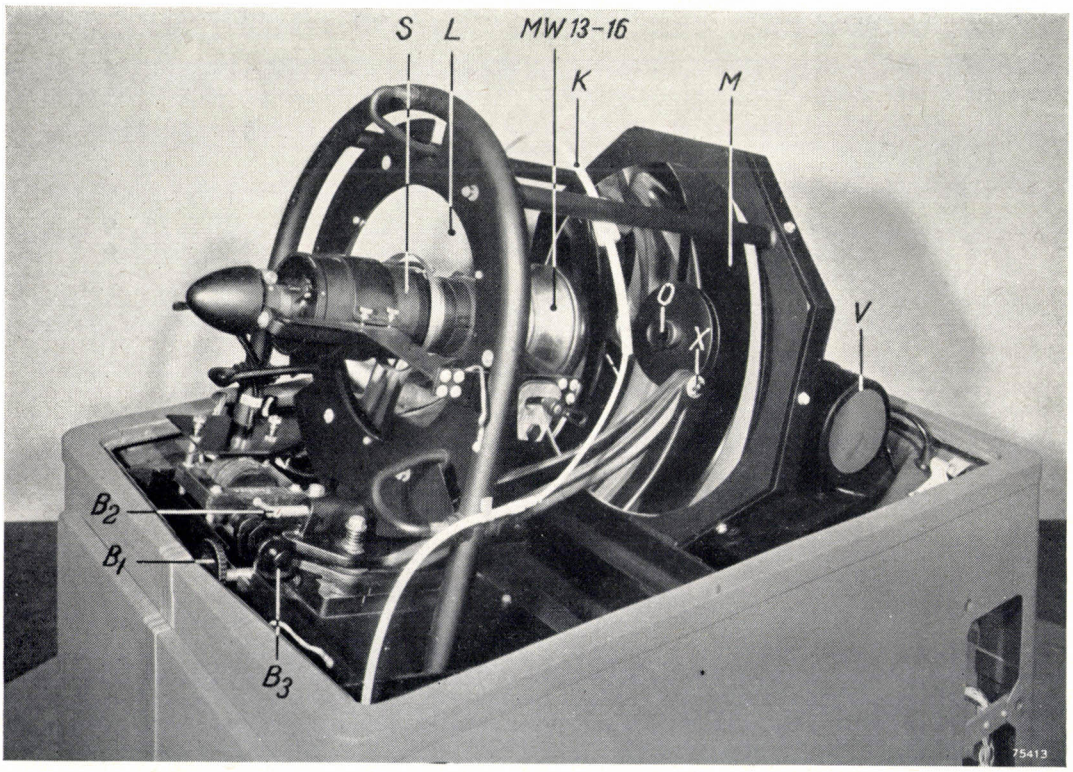


Fig. 3. The optical system in the Mammoth projector (with cover removed). S focusing and deflection coils on the cathode-ray tube MW 13-16. V motor driving blower at rear of the spherical mirror to blow cooling air through aperture O on to the screen of the tube. B_1 main adjustment for distance between tube screen and mirror. When a new tube is inserted, the axis of the tube must also be aligned with the principal axis of the mirror; hence the tube can be rotated on two axes, viz. with knob B_2 horizontally (X is one of the pivoting points) and with knob B_3 on a roughly vertical axis. K is the H.T. cable and L the correcting plate.

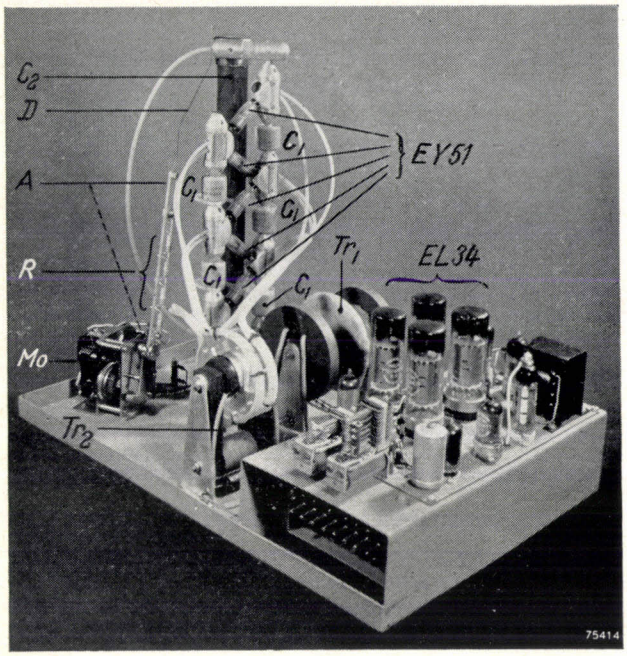


Fig. 4. Stabilized generator for 50 kV at an average current of 0.5 mA, as used in the Mammoth projector. In foreground: chassis with oscillator, control and protecting relay. The oscillator (frequency 20-25 kc/s) includes four valves type EL 34 and a transformer Tr_1. The oscillator voltage is rectified by six rectifier valves EY 51 with six capacitors C_1 in a cascade circuit (only 5 stages are visible in the photograph). A reservoir capacitor C_2 is included between the output terminals. Heater current for the rectifiers is obtained from a transformer Tr_2 which is fed from the oscillator. Tr_1 and Tr_2 both have Ferroxcube cores.

A is an insulated arm to which a series of resistors R and a contact wire D are attached. When the oscillator is switched off the arm assumes the position depicted, whereby D makes contact with the H.T. side of C_2. The capacitors are thus discharged across the resistors R, so as to eliminate all hazardous voltages. When the oscillator is switched on, a motor Mo moves this arm over to the position shown by the dotted line, and contact between D and C_2 is thus broken.

regarding which further details will be found in the subscript. Replacement of the cathode-ray tube necessitates only one or two simple operations, and adjusting knobs are provided for aligning the tube to its optimum position in relation to the mirror (the screen of the tube should be almost concentric with the spherical mirror, and the axis of the tube must coincide with the principal axis of the mirror). The whole operation of tube replacement and alignment takes only two minutes. The correcting plate is permanently mounted in the correct position with respect to the mirror.

The 50 kV D.C. supply

A number of methods of generating the high D.C. voltage needed for cathode-ray tubes have already been described in this Review [12]). In particular, mention was made of the generator designed for large television projection, and this generator is employed in the Mammoth projector. It consists of a valve oscillator operating at a frequency between 20 and 25 kc/s and delivering an A.C.voltage which is rectified to a high D.C. voltage by means of a circuit of rectifying valves and capacitors in cascade. The heaters of the rectifier valves are fed with high frequency current from the oscillator. A control system ensures a practically constant output voltage between no-load and full load, but with higher loads the voltage drops sharply, so that in the event of a short-circuit the current is very little more than on full load. In contrast to the method described in article [12]), the control potential in the present instance is in effect the difference between two D.C. voltages, one of which is proportional to the oscillator voltage, the other being a reference voltage maintained at a constant value by a stabilizer tube.

Should either the D.C. current or the voltage exceed a certain value owing to a defect in the control system, the generator is automatically switched off.

The H.T. generator of the projector consists of a single unit as depicted in *fig. 4*; further details are contained in the subscript.

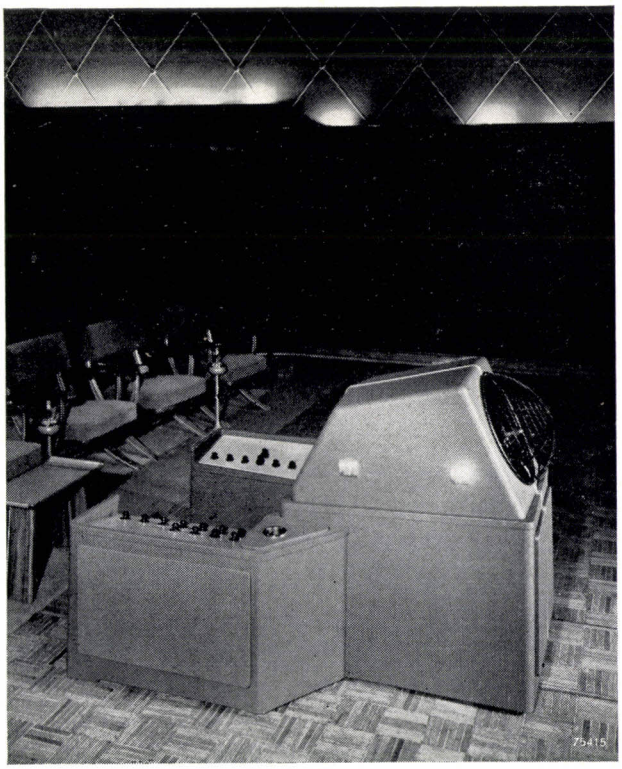


Fig. 5. The Mammoth projector (Type EL 5750) mounted in front of the first row of seats in a hall. The projectionist takes his place behind the main body, between the "wings". The unit is so low that projectionist and public have an unobstructed view of the screen.

[12]) J. J. P. Valeton, Philips tech. Rev. **14**, 21-32, 1952.

Construction details

In the design of the Mammoth projector the ultimate aim was to ensure facility in operation and a compact, self-contained unit; this takes the form of a main central part, with two "wings" (*figs.* 5 and 6) between which the projectionist is stationed.

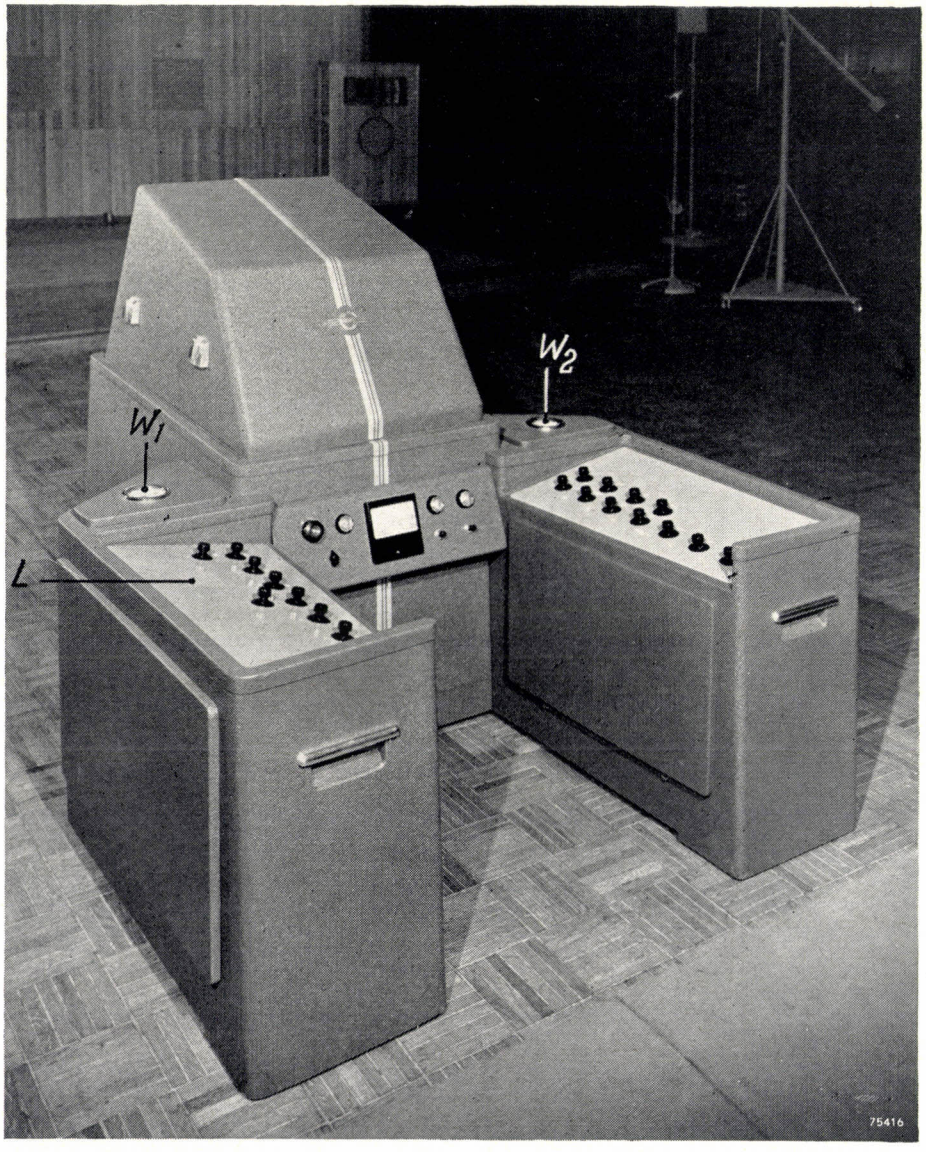


Fig. 6. Rear view of the projector. The left-hand wing contains a television receiver and an A.F. amplifier (25 W); the right-hand wing houses saw-tooth voltage generators for horizontal and vertical deflection, and also a video amplifier. The main body of the unit contains the optical system, H.T. generator and ancillary equipment. The meter indicates the beam current of the projection tube.

L monitor loudspeaker; W_1 and W_2 small cathode-ray tubes (type MW 6-2) for monitoring the picture as from the output of the receiver and video amplifier.

This equipment is complete in itself except for the aerial, the projection screen and the loudspeakers.

As already mentioned above, the distance from the projector to the screen must be 8 m for a magnification of 40×; this is simultaneously the minimum distance from which the audience should view the screen if the lines in the image are not to be perceptible to an uncomfortable degree. This minimum distance is based on the rule of thumb that it should be equal to twice the height of the picture (this rule holds for 625 lines; for other systems the distance is inversely proportional to the number of lines). The front row of the seating accomodation can accordingly be arranged on each side of, or just behind the projector, which has therefore been kept sufficiently low in design to prevent it from obstructing the view of the screen; the overall height of

the unit is not more than 1.15 m (3'-10"). The floor space occupied is only 1.35 × 1.65 m (4'-5" × 5'-5") (*fig. 7*).

As is seen in fig. 3 the upper part of the main body of the projector houses the cathode-ray tube and the optical system. Below these are the generator (fig. 4) and a stabilized rectifier unit for lower voltages. Mounted at the rear, within the immediate reach of the projectionist, is the panel which carries

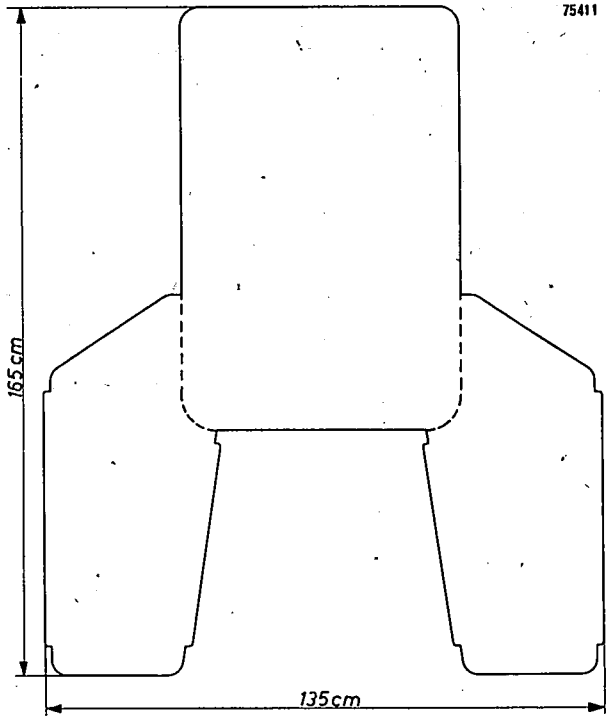


Fig. 7. Plan view of the projector (dimensions in cm). To facilitate transportation the wings can be detached from the main body (at the dotted lines).

the main switch, on/off buttons for the H.T. generator, a meter to indicate the tube current, and several indicator lamps (fig. 6).

The left-hand wing (fig. 6) contains a television receiver and sound amplifier, the latter having an output of 25 W. Both the sound and the picture can be monitored by the projectionist by means of a small loudspeaker and a cathode-ray tube which are also mounted in the left-hand wing.

In the right-hand wing are the video amplifier and the time base generators for the tube, and these also operate a small monitor tube; the projectionist does not switch on the projection tube until he is sure that the image on the monitor tube is satisfactory, particularly as regards synchronisation, thus sparing the audience the sight of confused and imperfectly synchronised images.

The MW 6-2 is used for both monitor tubes, this being the standard tube in projection receivers for home use [13]); in this case it is employed as a direct-vision tube and 9 kV is sufficient to operate it.

The video and A.F. amplifiers can if desired be connected to a television camera and microphone in "closed circuit", instead of to the receivers; the amplifiers are fully loaded for a video signal of 1.5 V across 75 Ω and an audio signal of 0.8 V across 500 Ω (in both cases peak to peak).

Needless to say, the necessary precautions have been taken to prevent contact with live parts of the equipment. For example the opening through which the light is emitted is fitted with a wire guard, so that the hands cannot be inserted. In the event of a fault developing in one of the generators, the beam current of the projection tube is suppressed to avoid damage to the fluorescent screen

The equipment can be easily moved about, as it is fitted with castors. With a view to transportation the wings are easily detachable; the electrical connections are completed by means of plugs and sockets, and the mechanical assembly with bolts and wing nuts. When detached, each of the three parts can be rolled along the floor.

As will be apparent from this description, the equipment as illustrated in figs 5 and 6, is self-contained (with the exception of the aerial, projection screen and loudspeakers).

The acknowledgements of the authors of this article are due to a large circle of co-workers.

[13]) See article II, referred to in footnote [3]).

Summary. A description is given of the "Mammoth" television-projector (type EL 5750), which is capable of giving pictures 4 × 3 m in size with a brightness (luminance) equal to that of cinema pictures (peak, 17 nit). The projector contains a 12 cm cathode-ray tube (type MW 13-16) operating on 50 kV, and a Schmidt optical system (concave spherical mirror with correcting plate). The dissipation at the fluorescent screen of the tube is 25 W and the screen is cooled by a jet of air. All the components, viz. the television receiver, video and A.F. amplifiers and H.T. generator are combined in this equipment, which can be operated by one person. The projector is mounted about 8 m from the projection screen, in front of the audience; it occupies a floor space of 1.35 × 1.65 m and is only 1.15 m in height. All the controls are within easy reach of the projectionist, and two small tubes and a loudspeaker are included for monitoring both picture and sound.

ABSTRACTS OF RECENT SCIENTIFIC PUBLICATIONS OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN

Reprints of these papers not marked with an asterisk * can be obtained free of charge upon application to the Administration of the Research Laboratory, Kastanjelaan, Eindhoven, Netherlands.

2015: K. H. Klaassens, J. Bakker, C. J. Schoot, J. Goorrissen: A simple method of measuring the moment of disappearance of soap micelles during polymerisation in emulsion (J. Polymer Sci. **7**, 457-461, 1951, No. 5).

A simple method is described for determining the moment of disappearance of soap micelles during the emulsion polymerisation of styrene by measuring the electrolytic conductivity during the polymerisation (with the potassium salt of stearic acid as a soap and potassium persulfate as an initiator). The results have been verified using the method of surface tension.

It is generally assumed that in the initial stages of an emulsion polymerisation soap in the aqueous phase is largely present in the form of soap micelles but rapidly transforms into monolayers surrounding the polymer-monomer particles formed by the polymerisation reaction.

2016: H. Bremmer: The jumps of discontinuous solutions of the wave equation (Comm. pure and appl. Math. **4**, 419-426, 1951, No. 4).

The connections existing between steady-state solutions and pulse solutions (in Luneberg's sence) of Maxwell's equations are studied. Luneberg's theory is based on the integration of Maxwell's equations over four-dimensional space-time domains. The present simplification of this theory is based on the use of the Heaviside unit function and the Dirac delta function. Equations for the eiconal function and for discontinuities on wave fronts in scalar wave propagation are treated in detail.

2017: E. J. W. Verwey, P. B. Braun, E. W. Gorter, F. C. Romeijn, J. H. van Santen: Die Verteilung der Metallionen im Spinellgitter und deren Einfluß auf die physikalischen Eigenschaften (Z. phys. Chem. **198**, 6-22, Oct. 1951). (The distribution of metallic ions in the spinel lattice and its influence on the physical properties; in German.)

After dealing with the distribution of the cation distributions in several oxygen spinels, as determined by X-ray analysis; the cases in which normal or inverted spinels are to be expected are discussed, using a heteropolar model. In the case of inverted spinels special attention is paid to the problem of order. The electrical problems of the ferrite spinels are in agreement with a theory exposed in anterior publications. The difference between Fe_3O_4 and Mn_3O_4 is discussed. The ferromagnetic saturation of the simple ferrites and their mixed crystals with zinc ferrite can be explained from Néel's theory. The indirect exchange interaction between magnetic ions is discussed. The influence of the angle Fe-O-Fe and its variation are in qualitative agreement with Anderson's theory. See Philips tech. Rev. **9**, 185-190, 231-248, 1947-48 and **13**, 194-208, 1952, No. 7.

2018: C. J. Schoot, J. Bakker and K. H. Klaassens: Remark on the article by S. R. Shunmickham, V. C. Hallenbeck and R. L. Guile, "Emulsion polymerisation of styrene, II. Effect of agitation" (J. Polymer Sci. **7**, 657, 1951, No. 6).

The effect of agitation on emulsion polymerisation, as found by Shunmickham and others, is ascribed to contamination of the nitrogen gas used in the experiments with traces of oxygen.

2019: H. P. J. Wijn and J. J. Went: The magnetisation process in ferrites (Physica **17**, 976-992, 1951, No. 11/12).

Initial magnetization curves of ferrites have been measured as a function of frequency up to 2 Mc/s. It has been found that the magnetization of sintered ceramic ferrites with a high permeability is brought about by at least two processes, one of which, in the frequency range covered, is independent of frequency and determines the initial permeability. The other process has a relaxation frequency of about 200 kc/s and is responsible for the irreversible processes during magnetization. From measurements on samples of sintered ferrites fired at different temperatures, it has been concluded that the frequency-dependent magnetization is caused by irreversible Bloch-wall displacements, whilst the initial permeability is caused by a reversible rotation of the magnetization in Weiss domains in the direction of the external magnetic field (in contrast to what is believed to be the case in cast ferromagnetic metals). A discussion shows that neither eddy current effects nor any inertia effects

so far known are responsible for the relaxation frequency of the Bloch wall at about 200 kc/s.

2020: M. E. Wise: Free electrons, traps and glow in crystal phosphors (Physica **17**, 1011-1032, 1951, No. 11/12).

The well known monomolecular and bimolecular theories of afterglow and thermal glow are summarized. Observed results on phosphors of ZnS type are discussed, particularly the phosphors with long afterglow in which the intensity-time laws are hyperbolic. These have been explained by many workers in this field by a monomolecular theory with a distribution of depths of electron traps. It is shown that certain mathematical approximations that must be used in relating the theory to observations of afterglow and thermal glow are not sufficiently accurate. A new and simple law is given that describes the same observations more exactly. A new physical theory for the free and trapped electrons is outlined, and this is discussed in relation to F and F′ centres in alkali halides and to an unsolved problem of semi-conductor theory.

R 179: J. L. H. Jonker: On the theory of secondary electron emission (Philips Res. Rep. 7, 1-20, 1952, No. 1).

Starting from (1) Whiddington's law concerning the velocity loss of electrons penetrating into a solid substance, (2) the experimental absorption law and (3) the assumption that the distribution of the secondary electrons within matter is isotropic, the author calculates the behaviour of secondary electrons as a function of various parameters. Good agreement is found with results obtained experimentally.

R 180: J. Volger: The specific heat of barium titanate between 100 °K and 410 °K (Philips Res. Rep. 7, 21-27, 1952, No. 1).

The specific heat of polycrystalline $BaTiO_3$ has been measured in a Nernst calorimeter. The lattice contribution cannot be described accurately with a single Debye function. In three transition regions the anomalous peaks in the specific heat have been determined.

R 181: W. J. van de Lindt: Application of multihole coupling to the design of a variable and calibrated waveguide attenuator and impedance (Philips Res. Rep. 7, 28-35, 1952, No. 1).

A short introduction describes the behaviour of two parallel waveguides, mutually coupled by n equidistant identical directional elements. A new type of calibrated variable attenuator and a variable impedance, capable of changing independently the amplitude and the phase of the reflection coefficient, are discussed.

R 182: G. Diemer and K. Rodenhuis: Optimum geometry of microwave amplifier valves (Philips Res. Rep. 7, 36-44, 1952, No. 1).

On the basis of Van der Ziel and Knol's theory of feedback amplifiers it is shown that for u.h.f. amplifier valves the maximum frequency limit for the amplification is reached if the electrode areas are so chosen that the useful capacitance is equal to the unavoidable stray capacitance. For optimum value of the product (gain × bandwidth), the useful capacitance should be somewhat higher.

R 183: G. Diemer and H. Dijkgraaf: Langmuir's ξ, η tables for the exponential region of the I_a-V_a characteristic (Philips Res. Rep. 7, 45-53, 1952, No. 1).

The inter-electrode distances of modern microwave diodes and triodes are often so small that the normal operating point lies in the exponential part of the characteristic. A set of ξ,η tables with the voltage gradient at the anode as parameter is given, from which the potential distributions in such cases can be derived. (See **R 7** and **R 8.**)

R 184: J. B. de Boer, V. Oñate and A. Oostrijck: Practical methods for measuring and calculating the luminance of road surfaces (Philips Res. Rep. 7, 54-76, 1952, No. 1).

In defining the qualities of an installation for road illumination the luminance distribution of the road surface is a more important factor than the illumination. However, owing to the lack of precise and quick methods of measuring and computing luminance, illumination values have mainly been used. In this paper methods are described which now enable the public lighting engineer to use road luminance in practice. The reflection properties of one particular road surface (dry asphalt) have been measured. Finally an example of the calculation of luminance is given in the appendix.

VOL. 15 No. 2, pp. 37-72 AUGUST 1953

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
THE PHILIPS INDUSTRIES

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, NETHERLANDS

AN ELECTRONIC HEARING AID

by P. BLOM.

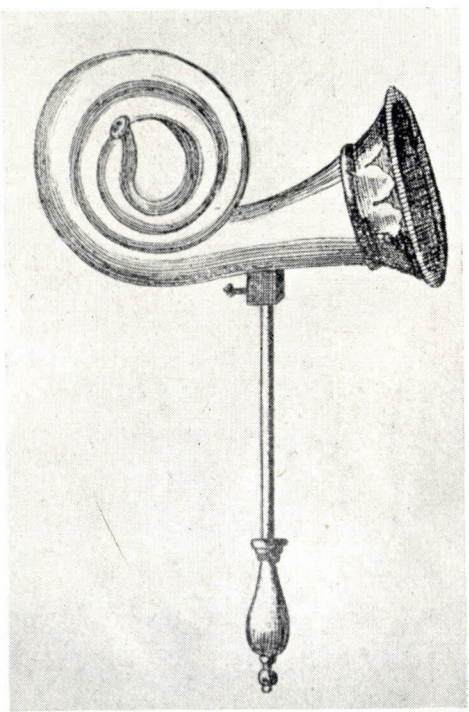
621.395.92:534.773.2

Means of compensating auditory deficiencies have been known for some centuries, but until recently these were large and clumsy appliances. The modern electronic hearing aid is small, light and unobtrusive, and it is much wider in its scope. For an electronic hearing aid to have the widest possible field of application it should be easily adaptable to suit the many different forms of deafness that exist. This point is given special consideration in the following description of Philips hearing aids.

The object of a hearing aid is to compensate, in some measure, the auditory deficiency in persons who are hard of hearing. The hearing aid is thus an amplifier of sounds, designed for one specific purpose. Such appliances have been known for some centuries, and everyone will remember the large trumpets which were still in use not so very long ago; their origin, of course, goes back much further than the seventeenth century ear trumpet shown in *fig. 1*. It is only within the last two decades that acoustic amplification has been achieved by electronic means. The advantages of electronic hearing aids are manifold; they are smaller, lighter and less obtrusive.

In this article a new hearing aid is described (marketed in three models, *fig. 2*), which differs in certain respects from earlier types. Every effort has been made to ensure not only compactness and low current consumption, but also the widest possible range of application. The design of the instrument provides for automatic volume compression based on several control levels, and a wide choice of response curves. A special characteristic of the design is the method of construction and assembly, which allows for the easy replacement of components and gives a high degree of resistance to tropical conditions.

A brief discussion of the human ear and the defects of hearing will first be given as an



75182

Fig. 1. Design for a hearing aid by A. Kirchner, 1684. (From G. v. Bekesy and W. A. Rosenblith, Early history of hearing: Observations and theories, J. Acoust. Soc. Amer. **20**, 727-748, 1948.)

introduction to the subject. A description follows of the technical features of the hearing aid and the miniature components which were specially designed for this instrument.

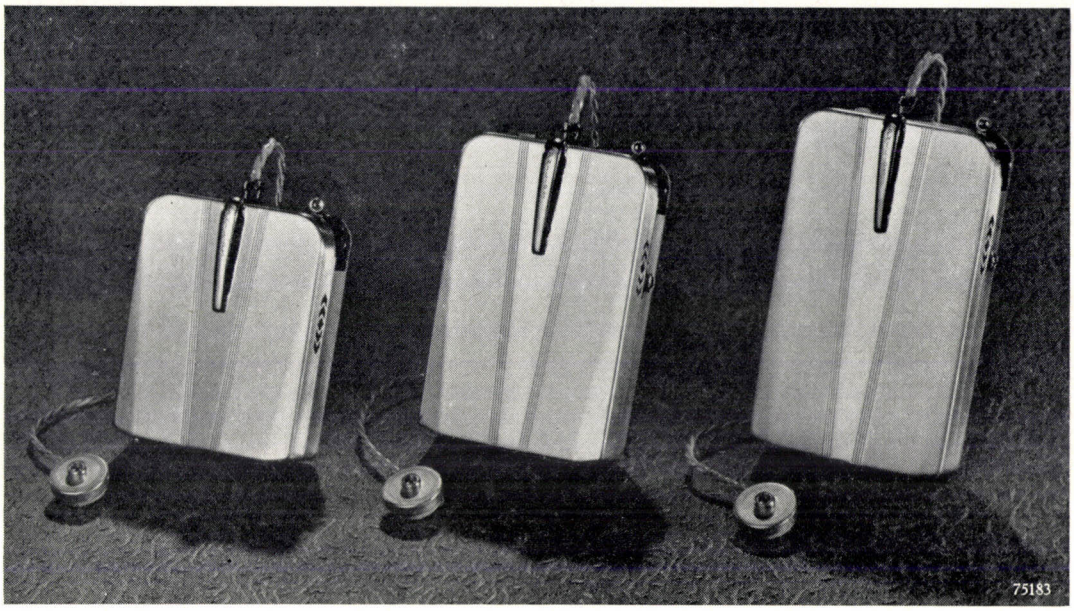


Fig. 2. The three models of new Philips hearing aid. The microphone slots are seen on the right-hand side, and the volume/tone control in the top right-hand corner. The switch next to the microphone slots operates the listening coil.

The organ of hearing

The human ear consists in the main of three parts, each of which can be regarded as quite distinct from the others (*fig. 3*). Starting from the outside we have first the outer ear which comprises the auricle or shell of the ear, and the auditory passage. The air vibrations of which sounds consist are collected by the auricle and are directed through the auditory passage to the middle ear, which is separated from the outer ear by the ear drum. The middle ear comprises the three well-known bones or ossicles of the ear. These ossicles serve to transfer the vibrations of the ear drum, as produced by the sound, to the membrane of the fenestra ovalis, an oval aperture which forms the connection with the inner ear. The inner ear communicates with the cavity of the nose and throat through the Eustachian tube and so prevents differences of pressure from occurring in the vestibule of the inner ear. Within the inner ear there is a spirally coiled duct of about $2^1/_2$ turns, narrowing towards the upper end, known as the cochlea. Now, the cochlea is divided lengthwise in two parts by a membrane, one half of which communicates with the fenestra ovalis and the other with a second aperture in the wall of the inner ear, the fenestra rotunda. The membrane is not continued right to the end of the cochlea and the two halves are therefore in communication with each other. The cochlea is filled with a viscous fluid, and the membrane, which is also hollow and filled with fluid, contains the so-called Corti's organ, to which the vibrations are transferred from the fenestra ovalis by the fluid. The cross-section of the cochlea with Corti's organ is not the same throughout its length, and it is due to this fact that each part of this organ responds to a different frequency; that part which is nearest to the fenestra ovalis responds to

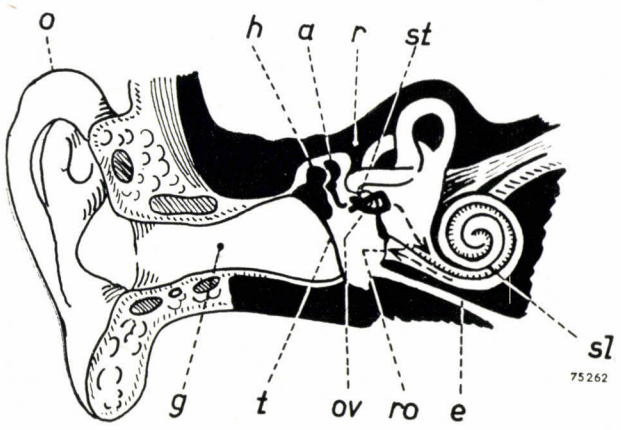


Fig. 3. Diagram of the human ear. *O* auricle; *g* auditory passage; *h* hammer; *a* anvil; *st* stirrup; *t* ear drum; *ov* fenestra ovalis; *ro* fenestra rotunda; *sl* cochlea; *e* Eustachian tube; *r* petrosal bone.

the very high frequencies, whereas the top of the cochlea is sensitive to very low frequencies. This mechanical selection of sounds within the cochlea is quite rough, but it is followed by a much finer, but as yet little understood process of selection by the auditory nerve, with which the cells of Corti's organ are in communication. The auditory nerve passes the sound stimuli to the brain, where the sound "signal", analyzed into its various component frequencies, is received; in this way we are enabled to form conclusions as to the quality and origin of a sound.

So far only air-conducted sounds have been mentioned, but sound can also reach Corti's organ by conduction through the bones of the skull (bone-conduction); by this means, however, it is mainly the lower frequencies that are perceived.

Defects of the ear may result in varying degrees of deafness and can be placed in three categories, viz. deficiencies in the conduction of the sound, causing conduction deafness; defects of the cochlea or auditory nerve, resulting in perception deafness; and defects of the brain. The last-mentioned will not be considered, as this cannot be made good by means of a hearing aid.

Conduction deafness may be caused by wax blocking the auditory passage, or it may occur as a result of inflammation. The ear drum can also be a source of conduction deafness, e.g. when wax is deposited on it. Another frequent cause is inflammation of the middle ear, possibly with rupture of the eardrum. Usually this kind of deafness is only temporary, but it can also be permanent (chronic inflammation).

A stoppage in the Eustachian tube, resulting in pressures below atmospheric in the cavity of the middle ear will also be accompanied by deafness.

Otosclerosis, a disease of the middle ear manifested in a morbid growth of bone, will sometimes hamper the movement of the auditory ossicles; in many cases this can be overcome by an ingenious operation known as fenestration.

In all these instances it is mainly the air conduction that is impeded or ceases to function altogether.

Perception deafness is frequently caused by a gradual degeneration of Corti's organ; usually this commences at the point where the higher frequencies are registered, and gradually progresses into the region of the lower frequencies. Curiously enough, however, at high sound *intensities*, high tones can still often be heard, and at their full strength too. There is then an abrupt transition between deafness and hearing (*fig. 4*). Since the threshold of pain (the maximum sound intensity which may reach the ear without causing actual pain) usually remains at the same level, the range or span of sound intensities that the sufferer can perceive is greatly reduced. This effect is known as regression.

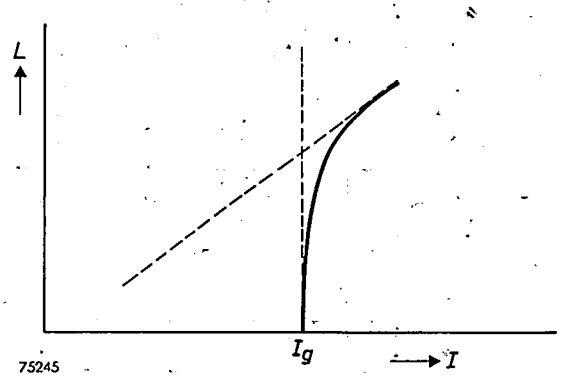


Fig. 4. The effect of regression. Loudness L as heard by the patient plotted against the sound intensity I in the auditory passage. Below a certain intensity I_g the patient hears little or no sound; above that point the hearing is comparable with that of a person with normal hearing.

In order to diagnose the nature of a deficiency in hearing, a chart is plotted for the auditory loss of the sufferer, for both air-conduction and bone-conduction, as a function of the frequency of the submitted sound. The auditory deficiency is defined as the ratio of intensities that can just be perceived by the patient, to those in the case of a person of normal hearing; this is expressed in decibels. The curve thus obtained is known as an audiogram and it is plotted with the aid of an instrument called an audiometer [1]). A number of audiograms are depicted in *fig. 5*, and from these it is seen that in serious cases of conduction deafness the sound pressure required to produce a sensation that is only just perceptible can be up to 60 dB, i.e. roughly 1000 times stronger than for persons with normal hearing. A pronounced dependence on frequency is not usually apparent. Perception deafness is distinguished from conduction deafness by the fact that hearing is impaired in both air-conduction and bone-conduction; this can be seen from the audiograms — as a rule, in conduction deafness, bone-conduction suffers no attenuation (2, fig. 5). As already mentioned, perception deafness is usually more evident in the higher frequencies, and in that region it may be absolute, i.e. more than 100 dB.

Many cases of deafness are of a mixed type, as will be seen from fig. 5. Regression cannot be

[1]) An instrument of this type is described in an earlier number of this Review; see L. Blok and H. J. Köster, Philips tech. Rev. **6**, 234-239, 1941.

diagnosed from a single audiogram; more extensive tests are essential when an examination is to be made.

The task of the hearing aid

The main function of a hearing aid is to so amplify sounds that speech will be intelligible and that local noises will give the user an impression of what is going on around him. The reproduction of music takes second place.

From the foregoing brief review of the major forms of deafness it will be clear that the individual requirements to be imposed on a hearing aid must differ very considerably between one case and another. In addition, it has not yet been found possible to establish any well-defined relationship between the nature of a deficiency in hearing as characterized by the audiogram, and the optimum characteristics required of the hearing aid by the user. For example, chart *3* in fig. 5 would indicate the need, on physical grounds, for a hearing aid with a rising amplification factor (rising frequency response curve), but very often, if not invariably, the patient will give preference to a flatter characteristic.

One of the difficulties in determining the most desirable characteristic for a hearing aid is the fact that the hearing of a deaf person using a hearing aid is monaural. In monaural hearing several curious effects become apparent. The usual directional effect largely disappears, although with some practice this can be recovered. Turning of the microphone produces differences in the quality of the signal as received, since the response curve of the instrument depends to a certain extent on the direction from which the sound is picked up [2]). Another effect of monaural hearing is the apparent reception of all local sounds from one and the same plane; in other words, the user of the hearing aid is not well able to determine the distance from the sound source, although here, too, a certain amount of judgment is restored in most cases, after some practice.

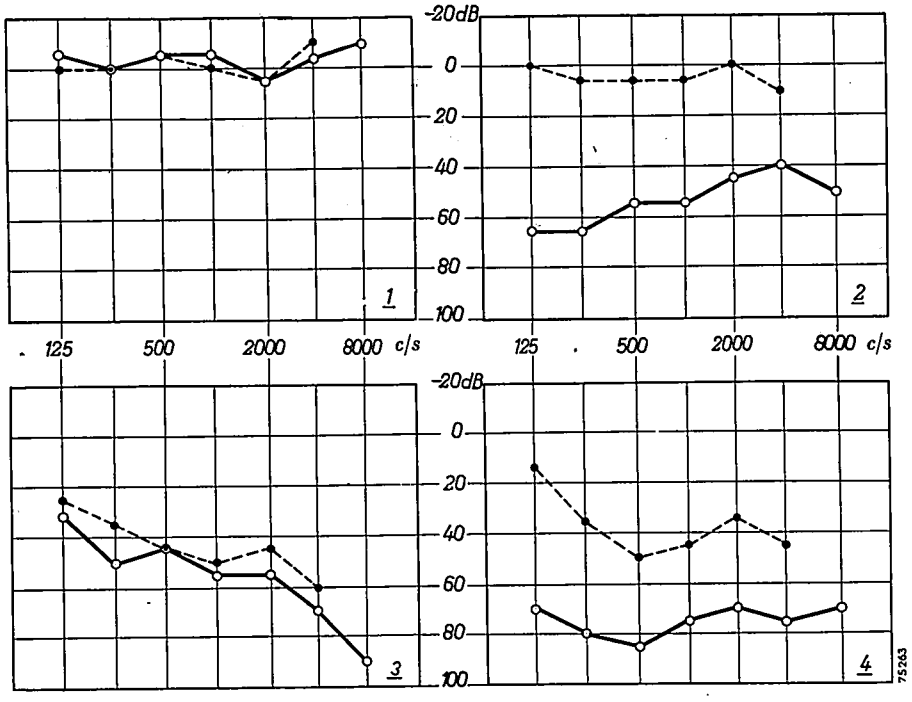


Fig. 5. Audiograms. Auditory deficiency in dB (vertical) compared with the average sensitivity of the human ear, as a function of the frequency of the sound in c/s (horizontal). Chart *1* represents normal hearing; *2* a person with conduction deafness; *3* a person with perception deafness and *4* a sufferer from a mixed form of deafness. The full lines refer to air-conducted sound and the broken lines to bone conduction. This figure consistsex of amples only: each particular form of deafness yields its own audiograms.

[2]) See also K. de Boer and A. Th. van Urk, Some particulars of directional hearing, Philips tech. Rev. **6**, 359-364, 1941.

As it is only the patient himself who can be the ultimate judge of the most desirable characteristics for the hearing aid, it is most essential that the characteristic should be variable within wide limits. This applies not only to the response curve, but also to other characteristics. Manufacturers would be able to meet the need for variable characteristics very easily were it not for the fact that they are tied in two respects. First, from the point of view of the user, certain other requirements have to be met, which are also quite important, e.g. lowest possible current consumption, small dimensions, lightness, robustness etc. In view of this it is usual to limit amplification to a range of frequencies of from 100 to 5000 c/s, but even then it is very difficult to adequately fulfil all the requirements mentioned.

Secondly, it is important to the manufacturers themselves to be able to meet the needs of the greatest possible number of users with only one type of hearing aid. In other words, their aim is to cover the various needs with a single type of instrument, capable of simple manual adjustment. This facilitates mass production and reduces costs.

Some details now follow of the manner in which these several requirements have been met in the Philips hearing aid.

Scope of the hearing aid

The basic requirement for the response of a hearing aid may be said to be an amplification of 50 dB in a frequency range of 200-4000 c/s (which includes speech and local sounds). Care must be taken, however, that the sound of loud shouting (input sound intensity about 30 dB above normal speech intensity), or of other very loud sounds, is not distorted, and that the user does not receive such a volume of sound that the threshold of pain is passed. The need for such control on the volume is particularly evident among sufferers from regression deafness, for their margin between the thresholds of hearing and pain is usually smaller than the range of intensities of the sounds which are to be rendered audible.

Control on the amplification can be either abrupt or continuous; the former method is known as signal cut-off, which implies that the signal is prevented from exceeding a certain value at the output. *Fig. 6* illustrates a signal subjected to cut-off; it will be seen that considerable distortion is introduced, and this affects the clarity of reproduction. In the Philips instrument, therefore, the continuous method has been adopted, viz. automatic volume compression, as this avoids distortion. In addition, the operating level of the control is variable in seven stages, to meet the need for versatility explained above.

Modification of the response curve of the Philips hearing aid can be effected in the first place by means of a tone switch. The user is thus given the facility of selection from three response characteristics to suit his particular need, the difference between the characteristics lying mainly in the limits of the frequency range. Of even greater importance is the fact that when purchasing his hearing aid the user is given a choice of seven earphones, each having an entirely different response curve. In all, therefore, a selection can be made from $3 \times 7 = 21$ overall characteristics.

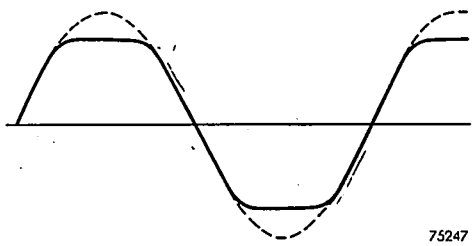


Fig. 6. Illustrating the distortion arising from abrupt cut-off of the output signal.

The microphone, pick-up coil and earphone

The microphone in the Philips hearing aid is of the crystal type, all the components being so assembled as to form a complete removable unit. A diagram of the microphone is given in *fig. 7*. The "capsule" of the microphone consists of a round housing in which the acoustically sensitive element, a square crystal of Rochelle salt (sodium potassium tartrate) is mounted. The capsule is mounted flexibly in a rectangular housing which also carries contact pins for connections to the amplifier. Sound vibrations are picked up by a small conical diaphragm of aluminium foil which is connected to one corner of the crystal. The electrical connections through the round and rectangular housings are insulated

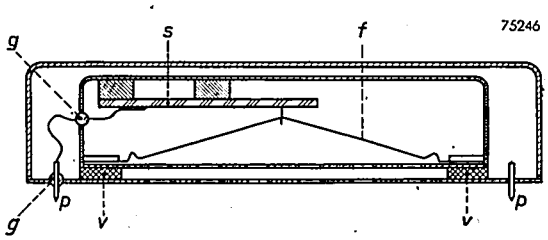


Fig. 7. Cross section of the microphone in Philips hearing aid. *s* Rochelle salt crystal; *f* aluminium foil; *v* resilient suspension of housing with crystal and foil; *g* glass beads; *p* contact pins.

with glass "bead" bushings which also serve to seal the casing against moisture.

A large part of the background noise that users of hearing aids always hear is produced by friction between the clothing and the microphone or neighbouring parts of the instrument ("case noise"). In order to suppress such noise as much as possible in the new hearing aid, the slots for the microphone are situated in the side of the case.

Rotation of a switch enables the user to connect the microphone to a so-called "pick-up coil". By means of this coil signals may be picked up inductively from the field of an exterior coil, such as that of a telephone receiver. The pick-up coil has a large number of turns, so that the stray field from a telephone receiver induces a signal voltage which is sufficiently large for the hearing-aid amplifier. In this way the acoustical link between the telephone and the hearing aid, which is responsible for considerable distortion, is eliminated. In addition, the magnetic transfer of the signal effectively eliminates all airborne sound, so that background noise [3]) — unusually distracting in a telephone conversation due to monaural hearing — is greatly reduced.

The pick-up coil is also being used to an increasing extent at concerts and in theatres and churches. A large loop of wire is laid round the walls of the building and connected to the output of an audio-frequency amplifier; the A.F. magnetic field of the loop induces voltages in the pick-up coil. Those who are hard of hearing need thus no longer be restricted to the front row or other specific places, but are free to move about while listening to reproduced sound which is almost wholly free from auditorium noises and reverberation. This system can also be used to great advantage by persons with normal hearing, for example in conducted parties in museums ("radiophonic" address) (*fig. 8*), for communication with personnel in large factory bays where the noise level is high or, again, in television or film studios.

The earphone of the hearing aid is attached to the auditory passage by means of a plastic ear-piece (*fig.9*). The closure should be as nearly as possible airtight in order to avoid loss of sound and distortion.

As already pointed out, the purchaser of the hearing aid has a choice of seven different earphones. Five of these are of the magnetic type with an impedance of 100 Ω (at 1000 c/s) and different response curves; the other two are crystal earphones of 100 000 Ω impedance, giving more output in the higher tones. The crystal receiver is also lighter and smaller than the magnetic one. All these receivers are fitted with a standard press-stud, by means of which the earpiece is attached.

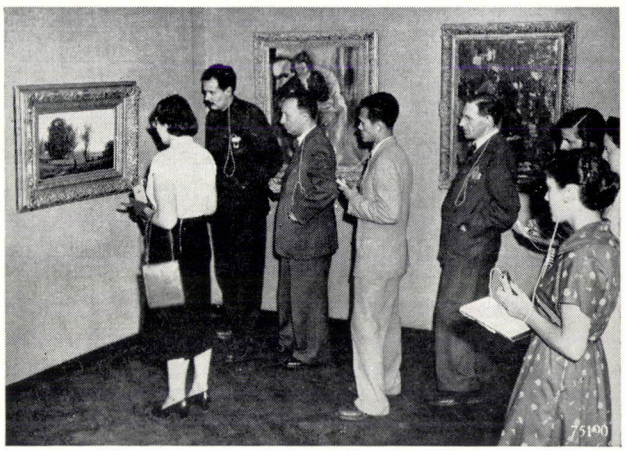


Fig. 8. Visitors to the Municipal Art Gallery, Amsterdam, listening to the "guide" with the Philips hearing aid and pick-up coil.

The amplifier

The amplifier in the Philips hearing aid is of the conventional 3-stage audio-frequency type. The maximum electrical gain is 72 dB, or an amplification of about 4000, which can be reduced manually by means of a volume control situated after the first valve. The average amplitude of the alternating grid voltage as delivered by the microphone is 1 mV for a normal sound level (speech); the A.C. voltage supplied by the output transformer is then about 5 V. Measurement of the sound pressure of the receiver placed in an artificial ear [4]) shows that

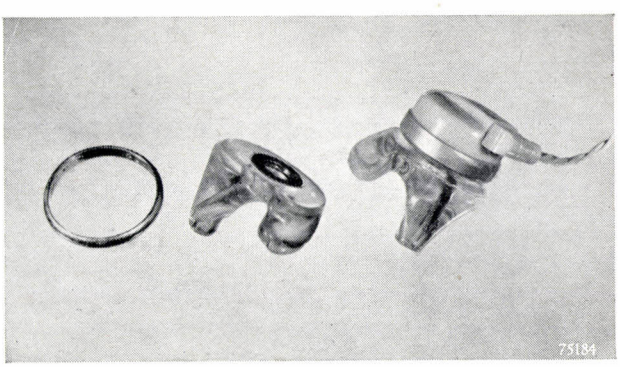


Fig. 9. Plastic ear-piece with press-stud by means of which the receiver is attached to the ear; on the right it is shown with the earphone attached. A wedding ring is shown for comparison.

[3]) Local noises can also be excluded by using a throat microphone, see Philips tech. Rev. **5**, 6-14, 1940.

[4]) For a description of the artificial ear see K. de Boer and R. Vermeulen, On the improvement of defective hearing, Philips tech. Rev. **4**, 316-319, 1939.

the acoustic amplification of the complete hearing aid is 52 dB at 1000 c/s.

The secondary side of the output transformer has an impedance of 100 Ω, matching the impedance of the magnetic earphone. By reversing a plug in its socket the user can increase the impedance at the output terminals of the amplifier to 100 000 Ω, thus providing a matching for a crystal earphone. *Fig. 10* shows the circuit diagram of the complete amplifier.

The automatic volume compression mentioned earlier is provided by that part of the circuit shown by heavy lines in the figure. The alternating anode voltage from the output valve B_3 is taken through R_1 and C_1 to a selenium rectifier *Se*, which rectifies it, and the resultant D.C. voltage is applied to the grid of the first valve B_1. In this way the bias of the input valve is automatically reduced and the amplification limited when the input voltage rises too much.

It is of course essential that strong sound pulses should be attenuated very quickly, as otherwise the initial components of a word might be distorted (the amplifier would be overloaded). The "attack time" of the A.V.C. is governed mainly by the product of the values of R_4 and C_2, which function as a filter for the control voltage from the selenium rectifier to the input valve. On the other hand, this time should not be too short, as this tends to make the amplifier unstable (motor-boating). This can be avoided, however, by coupling elements of suitable values. In practice an attack time of 150 msec has been found satisfactory.

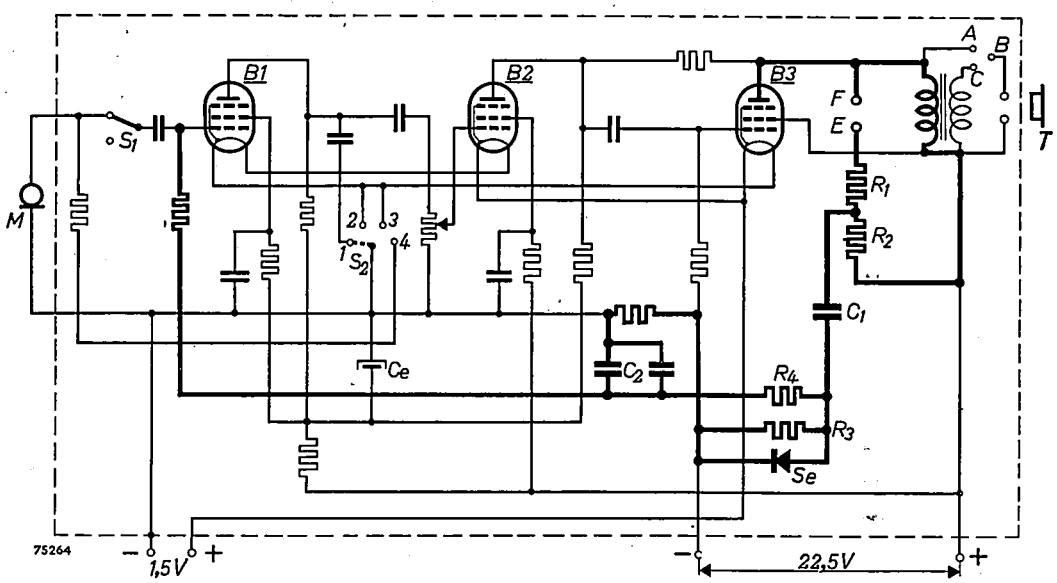


Fig. 10. Circuit diagram of the amplifier. Only the principal components are listed: *M* microphone; S_1 switch, microphone/pick-up coil (the leads to the pick-up coil are not shown). B_1, B_2, B_3 valves; S_2 tone switch: in position 1 the hearing aid is inoperative. C_e electrolytic decoupling capacitor. *A*, *B*, *C*, *E* and *F* contact sockets, strapped by means of a plug. *T* earphone receiver. The automatic volume compression circuit is indicated by the heavy lines. The alternating anode voltage from the output valve B_3 is fed through R_1 and C_1 to a selenium rectifier *Se*, and the rectified voltage is applied across R_4 to the grid of the input valve B_1. C_2 and R_4 determine the "attack" time of the compression, and C_2, R_4 and R_3 the "release" time. The control level is determined by the value of resistor R_2.

When the sound reaching the microphone is no longer abnormally loud, the amplification should return to its original level, and the time in which this takes place is the "release time"; in order to avoid a noisy background as a result of the A.V.C. becoming inoperative during the pauses in a sentence, this time must not be too short; it is determined by the time constant of the system C_2, R_4, R_3 in fig. 10, and lies somewhere between $^1/_2$ and 1 second.

Fig. 11 shows the action of the A.V.C. with acoustic pulses applied to the hearing aid. With

the control at maximum, the amplifier is overloaded for a very short time and therefore gives a distorted signal only for a few milliseconds.

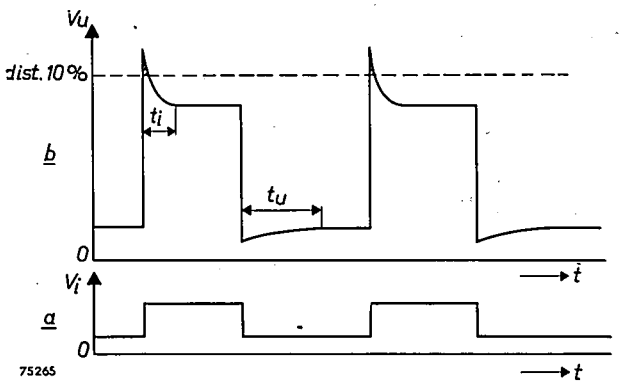


Fig. 11. The action of the automatic volume compression. *a*) A voltage of square wave-form (V_i) is applied to the input of the amplifier, the amplitude being such that, without the A.V.C., the distortion of the output voltage would be more than 10%.
b) The output voltage V_u, with A.V.C. operating, exhibits distortion of more than 10% for only a very short period. The significance of the attack and release times of the A.V.C. (t_i and t_u respectively) is clearly seen from the figure. The scale of t_i is exaggerated in the diagram.

As mentioned above, the control level can be adapted to the needs of the user. This individual matching is accomplished by means of a resistance (R_2, fig. 10) placed in series with the receiver. It is not necessary, nor is it desirable, that the user should have control over all the variables in the instrument: he will obviously never require to adjust the instrument to respond to forms of deafness other than his own. For this reason the output level is pre-set at the time of purchase, while the A.V.C. resistor is provided in the reversible plug to which reference has already been made. The patient, assisted by the doctor, has a choice of 2 × 8 different plugs, which will give him either uncontrolled output, or control in one of seven different ranges, each differing by 5 dB, either with the crystal or the magnetic earphone.

Some response curves for the complete hearing aid are reproduced in *fig. 12;* these show the ratio of output sound pressure to input sound pressure in dB, as a function of the frequency. The characteristics depicted are examples of what can be achieved with the seven-times-three combinations of receiver and tone control. This choice of characteristics, together with the eight ranges of A.V.C., provide a very wide range of application of the instrument, capable of meeting different forms of deafness.

Construction

The construction of the Philips hearing aid differs in some respects from conventional types. Apart from the necessity for small dimensions, robustness, resistance to tropical conditions etc., easy replacement of the components was made an essential feature of the design. If components can be easily replaced, repair work is greatly simplified. The dealer is then in a position to remedy the more usual faults himself with the aid of a manual only, no soldering or measuring instruments being necessary.

The advantages of the interchangeability of the receiver as well as the reversible plug are obvious. The amplifier is composed of units linked mechanically and electrically by contact pins; the batteries form another unit in themselves. Contact pins and springs complete the circuits of the microphone, tone switch etc., which are also separate units. This method of construction has proved a very satisfactory answer to the problem of the simple replacement of components.

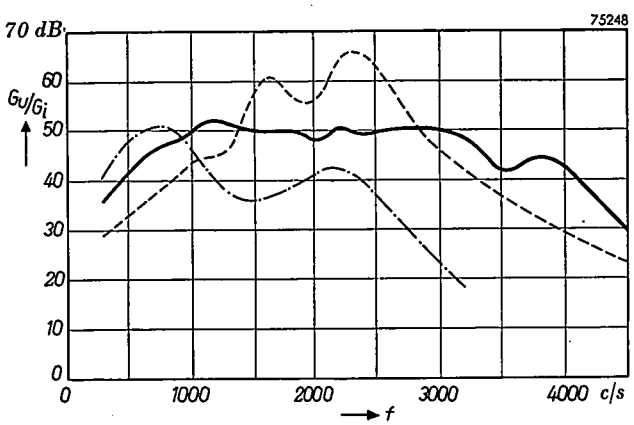


Fig. 12. Response curve of the complete hearing aid for three of the 21 combinations of earphone and tone control setting. The vertical axis represents the acoustic amplification, i.e. the ratio of output sound pressure G_u to input sound pressure G_i, expressed in dB.

Fig. 13 is an exploded view of the various components and shows how they are grouped into their various units. A photograph of the interior of the hearing aid is reproduced in *fig. 14*, which shows the intermediate model of the range of three (see fig. 2 and below). No individual resistors etc. will be seen, either in the exploded diagram or in the photograph, this being due to the special method employed in incorporating them all in their particular unit. The resistors and capacitors, as well as the output transformer, are first laid out on a resin-bonded paper plate (*fig. 15*), all the connections being made on the back of the plate. There

is an almost complete absence of soldered joints in this hearing aid: nearly all the leads are spot welded in order to avoid a dangerous rise in temperature during assembly of the components, which are mounted in close proximity to each other. The entire unit is moulded into a block of polyester resin [5]) which is then polymerized; the block thus formed is denoted by 7 in fig. 13. Polyester resins have very excellent insulating, damp-resisting and mechanical properties, closely resembling those of polystyrene and, owing to their low surface tension before polymerization, permeate easily into the small spaces between the components.

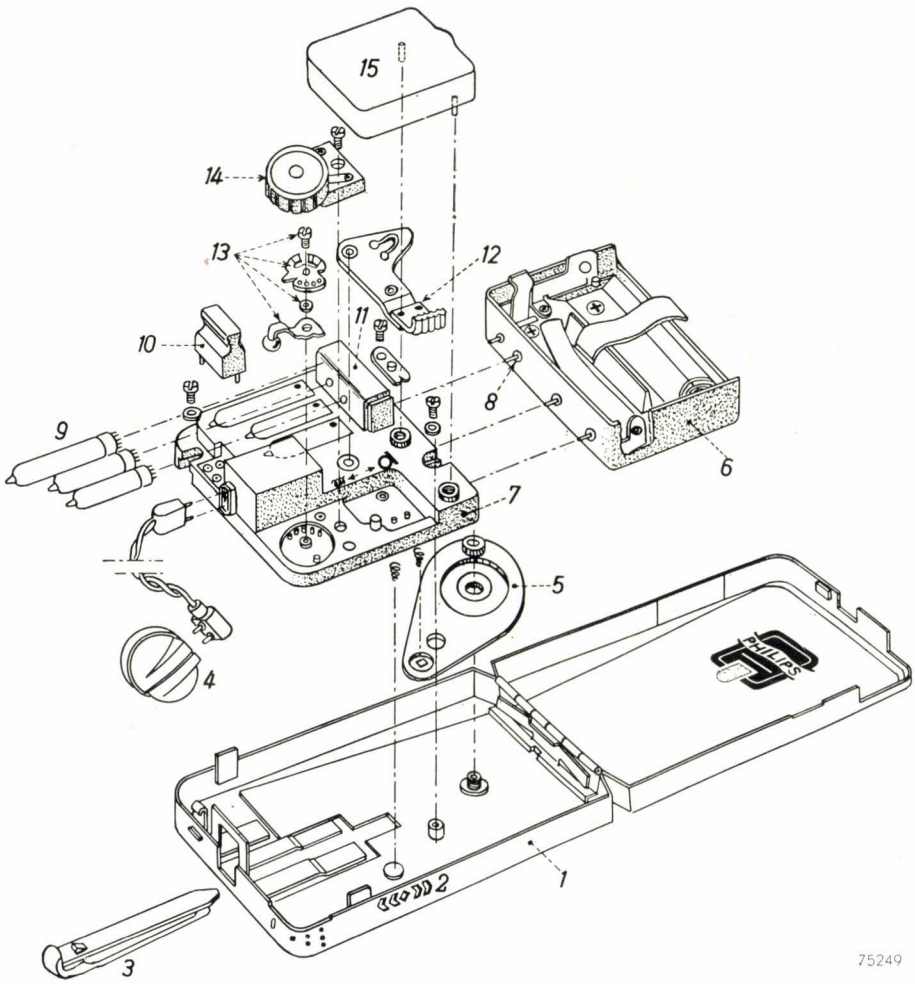


Fig. 13. Exploded view of the Philips hearing aid. The principal components are numbered: *1* case with microphone slots *2*, and clip *3* for attachment to the clothing; *4* earphone with cord; *5* pick-up coil; *6* battery case; *7* block of polyester resin containing the electrical components; *8* contact pins linking the components of the hearing aid electrically and mechanically; *9* valves; *10* plug: *11* valveholder; *12* switch for microphone/pick-up coil; *13* tone switch; *14* volume control. *15* microphone in screened housing.

The local "climate" in which hearing aids are generally used differs considerably from that in which most electronic equipment operates. The set is carried either in the breast pocket, or almost in direct contact with the body. In the latter case a relative humidity of 100% is fairly closely approximated. The temperature differs greatly according to the method of carrying and, when the case is worn close to the body, the possible effects of skin secretions must also be taken into account. All these factors are accentuated in tropical climates. In the design, therefore, allowance has been made for temperature variations between 0 and 45 °C and relative humidities between 0 and 100%. Thermal stability has been ensured by specially designing each component; this will be referred to again later. Sensitivity to humidity is avoided by moulding the moisture-sensitive components into the polyester resin block; judicious selection of the constituents of the resin has yielded a polymerization temperature of only 35 °C, so that damage to the components is avoided and only negligible

[5]) See J. P. Mudde, Polyester casting resins, Plastica **6**, 12-15, 1953 (No. 1).

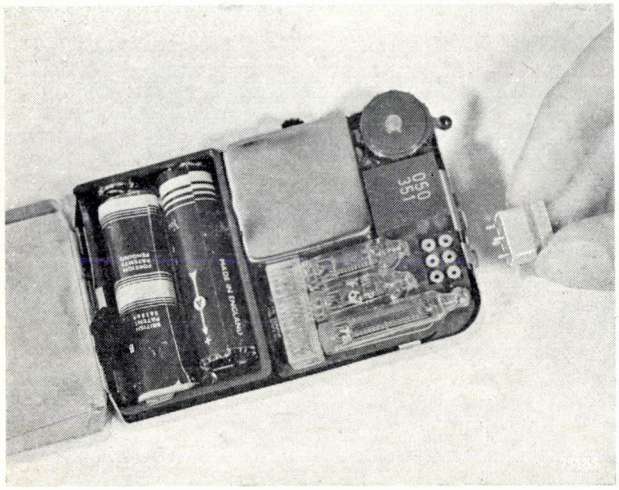


Fig. 14. Interior of Philips hearing aid. Only the screened microphone, the volume control, the valves and their holder, the plug and the batteries can be seen.

shrinkage stresses (on cooling) occur. In order to reduce still further the shrinkage and increase the physical strength, a quantity of ceramic powder ("Kersima" [6]) is added to the moulding powder. This ensures at the same time that the instrument shall be shockproof. Hearing aids belong to a class of equipment which, in contrast to radio receivers for example, are subjected to constant jolting and vibration; robustness, therefore, is an essential requirement.

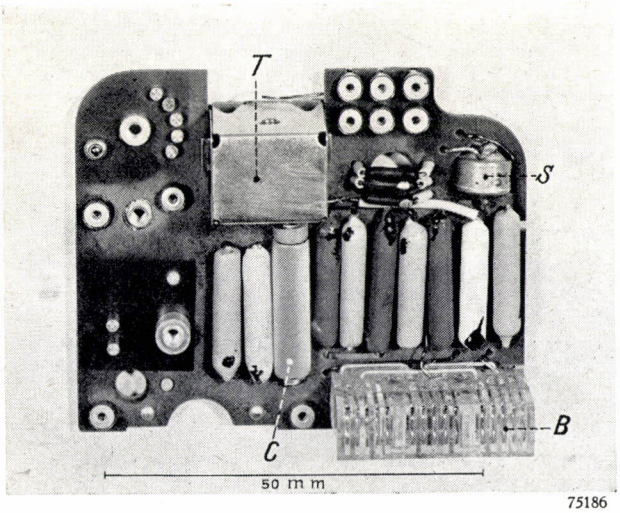


Fig. 15. Assembly of the electrical components on resin-bonded paper plate. The whole is later moulded solid in polyester resin. *T* screened output transformer. *C* electrolytic capacitor; *B* valveholder; *S* selenium rectifier.

[6]) For a description of different kinds of "Kersima", see R. A. Ijdens, Ceramics and their manufacture, Philips tech. Rev. **10**, 205-213, 1949; E. J. W. Verwey and R. D. Bügel, Ceramic materials with a high dielectric constant, Philips tech. Rev. **10**, 231-238, 1949.

The miniature components

The smallness of the hearing aid is due to the very compact design of the components. As standard components would require too much space, an entirely new range of components was designed for the new Philips hearing aid, very small in size, but with excellent electrical properties. The use of these components are not confined to hearing aids alone, for there are many other fields in electronics in which an important outlet for such components exists. The more important components of the hearing aid will now be discussed.

The valves (*fig. 16*) are of course battery valves, of the sub-miniature type. The DF 67 is used in the first and second stages; the output valve is the DL 67, except in the smallest model of the hearing aid which is fitted with the DL 66. The valves are mounted in a special valveholder.

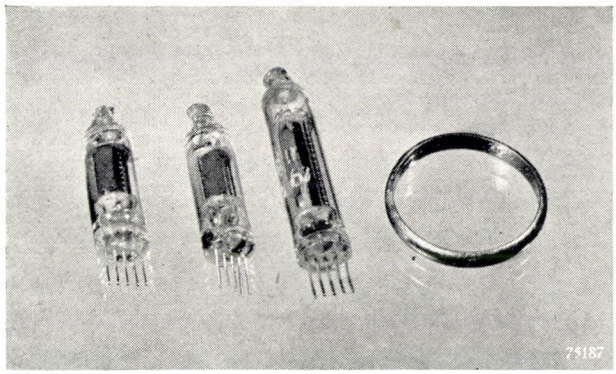


Fig. 16. The three amplifier valves, with a wedding ring for comparison of size. The valve at the right is the output valve DL 67.

The coupling resistors between the three stages of amplification are very small indeed, viz. $7 \times 1^1/_2$ mm, and are capable of carrying a load of 1/50 W (*fig. 17*). They consist of ceramic tubes coated with graphite suspension, the characteristics of which can be varied to yield resistance values of from 1000 Ω to 15 mΩ. The coupling capacitors, of which a specimen is also shown in fig. 17, are also manufactured from ceramic tube, in this case a grade of "Kersima" of which the dielectric constant is very high, namely 4000. With this material very small capacitors can be made in a range of values up to 6400 pF. The dimensions are $16 \times 3^1/_2$ mm. A special feature of these components is that the leads are attached inside the tube of the capacitor, leaving the exterior quite smooth. The capacitors are protected by a coating of insulating material and can therefore be mounted in close contact with each other.

An electrolytic capacitor (*Ce* in fig. 10) is connected in parallel with the anode battery to prevent the amplifier from oscillating (it decouples the anode circuit). There is always some risk of oscillation as the batteries get older; the internal resistance then rises and the coupling between the circuits becomes tighter. Hence the capacitor in the circuit enables the user of the hearing aid to obtain the longest possible service from his batteries, although there is of course a limit, imposed by the ultimate decrease in the amount of power available. The electrolytic capacitor is also quite small; it is housed in an aluminium tube 18 × 4 mm (fig. 17) and is made in two ratings, viz. 2 μF at 25 V and 10 μF at 2 V.

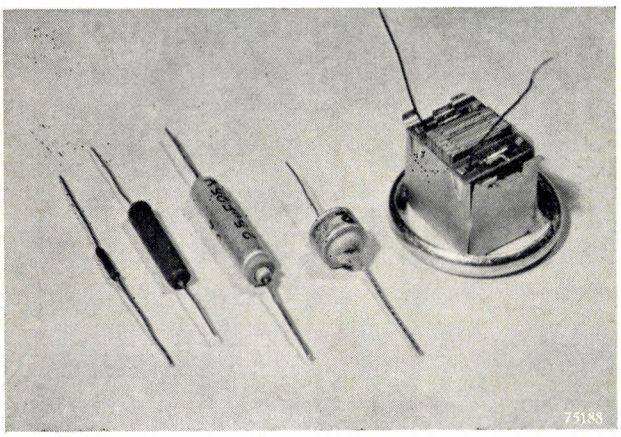


Fig. 17. Miniature components: left to right, carbon resistor, ceramic capacitor, electrolytic capacitor, selenium rectifier and A. F. transformer with part of the screening box removed.

The output valve of the amplifier is matched to the receiver by means of a special transformer (also shown in fig. 17) which provides the impedance matching of 100,000 or 100 Ω. When a crystal receiver is used, the coupling consists only of the self-inductance of the primary winding. Owing to the high impedance of the valve, it is necessary, for effective matching, to provide a high self-inductance; to realize this within small dimensions is made more difficult by the presence of a D.C. component of 200 μA which reduces the permeability of the core. The use of high alloy-content laminations and a large number of turns of very thin wire (30 micron) offers a satisfactory solution. At 1000 c/s the efficiency is 60%, which is high for such a small transformer, being of the same order as that of output transformers of standard dimensions.

As already mentioned, the transformer is moulded into the polyester resin block with the other components; the penetration of the resin guarantees the resistance of this component to tropical conditions.

The amplification of the hearing aid is controlled manually by means of a potentiometer (*fig. 18*) connected between the input and second valves. This is a rotary potentiometer and the physical requirements to be met by this component are fairly heavy. The user is continually adjusting it to give the required volume of sound, and a life of 20,000 operations on both directions is essential. To withstand this a robust spindle bearing is necessary, and fine tolerances on the other parts of the potentiometer. An improved carbon track with a moving carbon contact has practically eliminated potentiometer noise, even for resistances up to 5 MΩ. The amplification between the potentiometer and the output valve is some 400 × , so that it was further necessary to eliminate ripple voltages induced in the grid circuit of the second valve by the output valve. Such voltages could reach the potentiometer through body capacitance via the finger operating the control, and for this reason the potentiometer is provided with an internal screen.

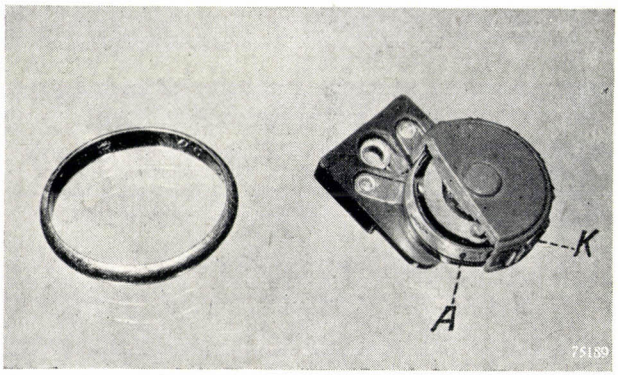


Fig. 18. Volume control with internal screening. The rotary cap *K* is shown cut away to reveal the screening *A*.

The need for such screening could have been avoided by placing the potentiometer between the second stage and the output valve, but strong input signals might then have overloaded the second valve.

To a large extent the dimensions of the hearing aid are dependent on the size of the batteries, two of which are required, viz. one of $1^1/_2$ V for the filaments and another, of 22 V, for the anodes. As sub-miniature valves are used, very little power is needed (40 mW from the filament battery and 5 mW from the other), so that for a reasonable life, only very small batteries are required (filament battery life about 20 hours; anode battery about 250 hours).

In order to leave the compromise between dimensions and battery life (also economy) to the user, the three models as illustrated were placed on the market, the only difference being in the size of the battery holder (the smallest model, moreover, has no listening coil and is fitted with an output valve of smaller rating). There is accordingly a choice of very small batteries with a correspondingly small case, and larger batteries with larger case, but lower running costs.

Summary. Defects in hearing can assume many different forms. The greater the adaptability of an electronic hearing aid in its applications, therefore, the larger the number of persons who can benefit from it.

This need for adaptability is emphasised in practice by the fact that it has been found impossible to define a unique relation between the desired characteristics of the hearing aid and the nature of the deafness in each individual case. In the new series of hearing aids adaptability is ensured by providing automatic volume compression in a range of seven different levels, from which a selection can be made by ordering the appropriate plug. Further, there is a choice of seven different earphones, five of which are magnetic, and two of the crystal type. These features, in conjunction with a 3-position tone switch offer a total range of 21 different characteristics. The microphone can be switched over to a pick-up coil whereby signals are received directly by induction, for example, from a telephone. In the design of these hearing aids every endeavour has been made to ensure that all the components shall be easily replaceable, so that repairs can be carried out quite simply without any soldering operations. The electrical components are of the miniature type, specially designed for the purpose. All the resistors and capacitors, as well as the output transformer, are moulded into a solid block of polyester resin, ensuring a high degree of robustness and resistance to tropical conditions, and low sensitivity to humidity. All the components, too, were designed with a view to long life and reliable service. Three models of the hearing aid are manufactured, the principal difference between them being in the rating of the batteries.

THREE METHODS OF MEASURING MAGNETIC FIELDS

I. MEASUREMENT BASED ON THE GENERATOR PRINCIPLE

by B. F. JÜRGENS.

II. MEASUREMENT OF THE FIELD ON THE AXIS OF MAGNETIC ELECTRON LENSES

by A. C. van DORSTEN and A. J. J. FRANKEN.

III. MEASUREMENT BY THE PROTON RESONANCE METHOD

by H. G. BELJERS.

621.317.42:538.082

The best method of measuring magnetic field strength depends on the characteristics of the field. Three methods currently employed in the Philips laboratories are described in this article, each of which has its own particular applications. The first method described is suitable for most magnetic circuits and is accurate to within two or three percent. The second was developed specially for measurements on magnetic electron lenses, while the third provides a very accurate absolute measurement of fairly strong homogeneous fields.

Introduction

A large number of methods have been devised for the measurement of magnetic fields and much is still being written on the subject *). When the necessity for such measurements arises in experimental work, special apparatus is often constructed to suit the particular conditions; the need for such special equipment may be due to the fact that the field is restricted within a very confined space, or that measurement is to be effected close to the walls. Examples of such conditions occur in the fields within the air-gaps of electrical machines or moving-coil instruments. The accessibility of the points of measurement will also affect the design of the equipment. For example, the air-gap in an iron circuit will usually be accessible only in a direction at right angles to the lines of force; on the other hand the aperture of a magnetic electron lens is accessible only in a direction parallel to the lines of force. The measurement of inhomogeneous fields also has its own particular problems; here, a certain average value of the magnetic induction over the space occupied by the measuring head is recorded, and there is some uncertainty as to the exact point where the measured value actually exists. Often this difficulty is partly overcome by employing a very small measuring head and, in the first of the methods described in this article, this is the solution adopted. At the same time, it is sometimes possible to determine local values of the field with precision by means of an instrument that is not small, by utilizing certain properties peculiar to the magnetic field. The second method described is an example of this **).

The discovery and study of new phenomena and their behaviour in magnetic fields have again and again led to the development of new kinds of measuring equipment, and this applies to the third method dealt with in this article. This is an incidental outcome of research into the magnetic moments of atomic nuclei. A distinguishing feature of the method is the unusually high degree of accuracy which it offers (within 0.01%).

*) A review of practical methods is given by L. W. McKeehan, The measurement of magnetic quantities, J. Opt. Soc. Amer. and Rev. sci. Instr. **19**, 213-242, 1929.

) Another example will be found in W. F. Brown and J. H. Sweer, The fluxball. A test coil for point measurements of inhomogeneous magnetic fields, Rev. sci. Instr. **16, 276-279, 1945.

I. MEASUREMENT BASED ON THE GENERATOR PRINCIPLE

Use is often made of the generator principle to measure magnetic induction [1]). A coil is rotated within the field, about an axis perpendicular to the axis of the coil; the alternating voltage thus generated is proportional to the speed of rotation and also to that component of the induction which is at right angles to the axis of rotation. If the speed be kept constant, the induced alternating voltage is a measure of this component. As a rule a synchro-

[1]) A. Kohaut, Ein Messgerät für magnetische Felder, Z. tech. Phys. **18**, 198-199, 1937. R. H. Cole, A magnetic field meter, Rev. sci. Instr. **9**, 215-217, 1938. O. Klemperer and H. Miller, Search-coil oscillator for measuring fields of magnetic electron lenses, J. sci. Instr. **16**, 121-123, 1939.

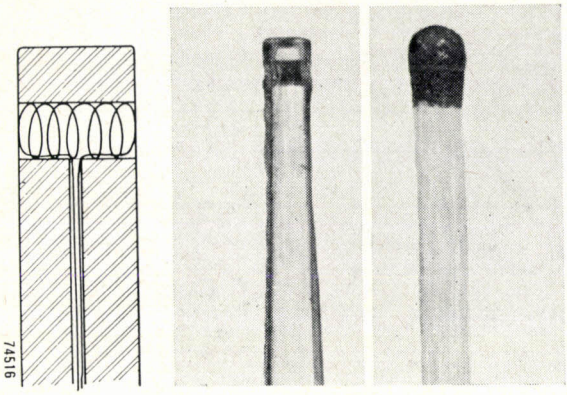


Fig. 1. For measurement of a magnetic field by the generator principle a coil is rotated at constant speed in the field. The coil is mounted in the extremity of a "Perspex" spindle fitted as extension piece to the shaft of a synchronous motor. On the left is an enlarged sketch and in the centre a photograph of the end of the actual spindle with coil. Right: a match by comparison on the same scale.

nous motor is used to drive the coil, the speed then being constant in so far as the mains frequency is constant.

In the instrument under review, the spindle of the small synchronous motor carries an extension piece of some easily machinable plastic such as "Perspex". A very small coil is mounted in this spindle, near the extremity *(fig. 1)*. The A.C. voltage generated when the coil is revolved in a magnetic field, is taken by means of bronze contact rings and silver springs *(fig. 2)* through a screened cable to an amplifier, where the signal is amplified, rectified, and measured by a moving-coil instrument. The deflection of this meter is a direct measure of the field component perpendicular to the spindle at the point where the coil is located. The equipment is shown in use in *fig. 3*.

The exploring coil used in this case is considerably smaller than anything made hitherto for equipment of this type. The inside and outside diameters are 0.5 and 1 mm respectively, and the length is 1.5 mm. These very small dimensions make it possible to measure local values in inhomogeneous fields without difficulty.

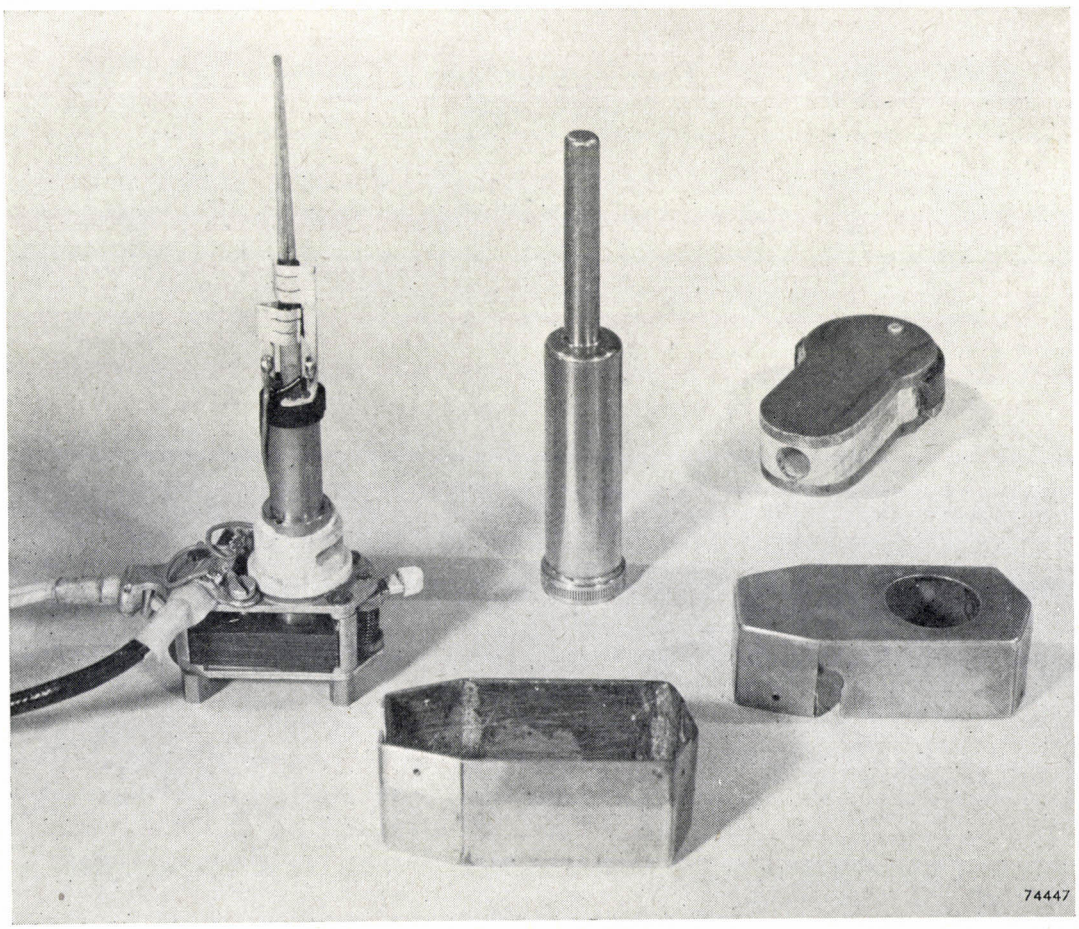


Fig. 2. The synchronous motor with "Perspex" extension piece. The silver contact springs for collecting the voltage can be clearly seen. In addition to the motor housings and the spindle cover, a magnet used for calibration of the instrument is shown in the photograph (background, right).

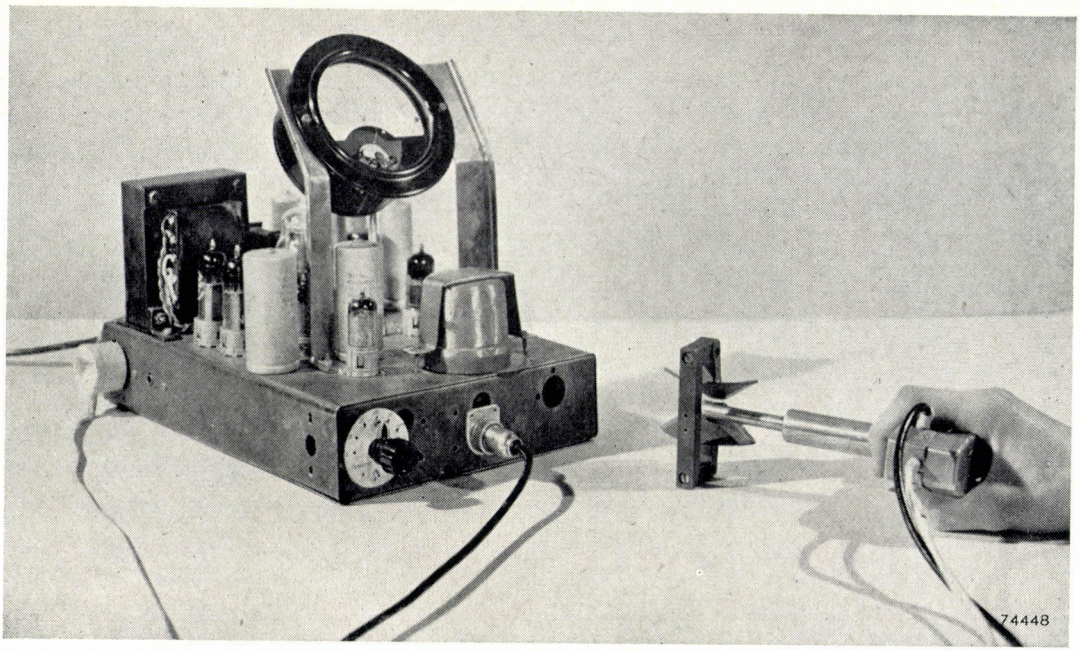


Fig. 3. The field-strength meter in operation. At right: the motor is held in the hand and the exploring coil in its cover is inserted between the pole pieces of the magnet to be measured. At the left is the amplifier and meter.

In such a small coil the flux per turn will of course be very small indeed; moreover, the number of turns cannot be very great, as there is a practical limit beyond which the thickness of the wire cannot be reduced. The sensitivity of the exploring coil drops sharply with the dimensions, but, owing to the amount of amplification provided, the meter gives full scale deflection (at max. amplification) for a field of 3×10^{-4} Wb/m^2 (3 gauss).

The coil is made of enamelled copper wire 25 microns in diameter, which, in the space available permits of 500 turns. With a view to the desired sensitivity a fairly high speed of rotation was adopted, viz. 50 revolutions per second (3000 r.p.m.) It would not be possible to go much beyond that, seeing that no extra bearings are provided for the "Perspex" extension on the motor spindle.

When the rotating probe is held with its axis parallel to the lines of force at the coil, the field component in the plane of rotation is zero and the meter gives no deflection. This, then, is a means of ascertaining the direction of the lines of force. The absolute value of the induction is measured by finding the position corresponding to maximum meter deflection. It is of course necessary that there is sufficient room for such positioning.

In order to ensure optimum transfer of energy from the exploring coil to the amplifier, it is necessary to match the impedance of the former with the input impedance of the latter. This impedance consists almost solely of the attenuator which is provided to permit variation of the measuring range, and is 500 000 Ω (see *fig. 4*). At a frequency of 50 c/s the self-inductance of the exploring coil is still determined only by the D.C. resistance, which is 70 Ω. It was accordingly found desirable to couple the coil to the amplifier by means of an input transformer. Now, the available voltage induced in the exploring coil is divided between the resistance of the coil itself and the inductive impedance of the transformer. When the frequency increases, this inductive impedance rises and, with it, the ratio of the voltage across the primary of the transformer to that induced in the exploring coil. In the region of 50 c/s this ratio is roughly proportional to the frequency [2]). As the voltage induced in the exploring coil itself is also proportional to the frequency (speed of rotation) the potential applied to the grid of the input valve of the amplifier is approximately proportional to the square of the mains frequency. Now, in order to avoid serious errors in measurement arising from this, in the event of fluctuations taking place in the mains frequency, the amplifier was designed to have a response curve such that the gain in the region of 50 c/s decreases as the square of the frequency. For this purpose two *RC* filters are employed, these being located between the first and second amplifier valves (fig. 4).

[2]) It is possible to design a transformer which, under the conditions described, will ensure a voltage ratio that is practically independent of frequency in the region of 50 c/s.

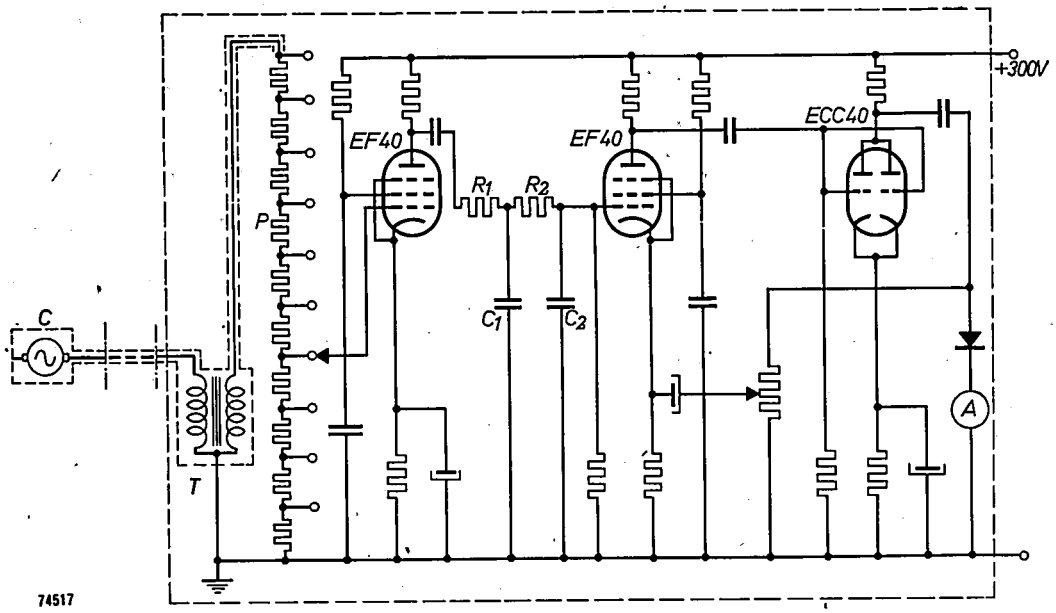


Fig. 4. The amplifier circuit. The voltage generated in the exploring coil C is taken across a magnetically screened input transformer T to the amplifier. The effect of the double RC filter (R_1-C_1 and R_2-C_2) between the first and second valves is to modify the response curve in such a way that the reading will be independent of the mains frequency. P, calibrated attenuator for the different measuring ranges. A, represents the meter.

The input transformer is provided with effective magnetic screening to avoid interference due to magnetic leakage-fields set up by the mains transformer (such fields are of the same frequency as that of the signals used); the screened transformer can be seen at the front of the chassis shown in fig. 3.

A stabilized supply unit is used for the amplifier, so that variations in the mains voltage shall not affect the amplification.

The complete unit is calibrated by placing the exploring coil in a magnetic field of known strength; the calibrating magnet supplying this field is depicted at the rear in fig. 2. The space between the pole pieces of the magnet is filled with wood in which a hole is provided that just accommodates the screening tube on the "Perspex" spindle. To give the desired reading of the known field strength on the meter, the amplification is adjusted by means of variable feed-back. Since the attenuator has a linear calibration with respect to the meter, the reading is then correct in all the other measuring ranges as well.

In the most sensitive position, a full-scale deflection corresponds to an induction of 3×10^{-4} Wb/m^2 (3 gauss), and in the least sensitive position to 3 Wb/m^2 (30 000 gauss). The error is roughly 3% of full scale, except in the lowest measuring ranges (viz. 3×10^{-4} Wb/m^2 and 10^{-3} Wb/m^2). With such low induction values, a certain amount of interference is experienced in the form of a meter deflection not due to the magnetic field, and corresponding to about 3×10^{-5} Wb/m^2.

II. MEASUREMENT OF THE FIELD ON THE AXIS OF MAGNETIC ELECTRON LENSES

In order to determine the focusing characteristics of magnetic electron lenses such as those used in electron microscopes, it is usually necessary to ascertain the form of the field along the axis of symmetry. When this is known a computation can be made of the field away from the axis.

This statement holds for a part of the space in which no electric currents flow and no magnetic materials occur, and this is the case in the vicinity of the axis of a magnetic electron lens. In such spaces the value of the magnetic induction can be derived from a scalar potential conforming to Laplace's equation. (The derivation is an application of potential theory.)

In essence, a magnetic electron lens consists of a coil enclosed within a soft iron housing (see sketch *fig. 1a)*. The curve representing the field along the axis is more or less bell-shaped. (fig. 1*b*). Approximate computation of this field is possible, but the accuracy is not always high enough for the purpose in view, especially if saturation phenomena occur in the iron. If, say, the spherical aberration of a lens is to be calculated with any precision, the error in the calculation for the field should not exceed 1%. It is better, therefore, to measure the field direct

from the lens itself, or from an enlarged or reduced scale model of it.

In the following a description is given of a method of measuring, rapidly and sufficiently accurately, the induction along the axis of a magnetic electron lens. This method is a modified version of the one employed by Van Ments and Le Poole for the same purpose [1]).

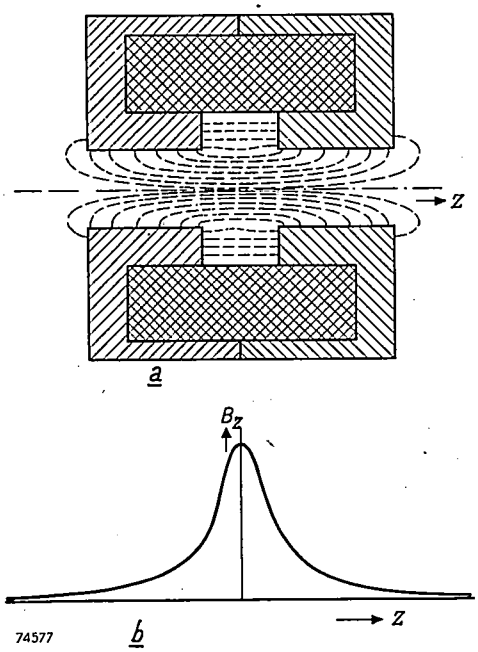


Fig. 1. *a*) Diagrammatic representation of a magnetic electron lens. The lens consists of a coil, almost wholly encased in soft iron; on the inner surface is a small air gap. The field is concentrated mainly inside the coil, near the gap in the iron. *b*) The curve of the magnetic induction B_z on the axis of such a lens is more or less bell-shaped.

Van Ments and Le Pool measured the static force exerted upon an exploring coil within the magnetic field, with a constant direct current flowing through the coil. The coil is specially shaped and consists of a long thin solenoid, the windings of which are reversed at the mid-point M (*fig. 2*). To make the measurement, the coil is positioned with its axis coincident with that of the electron lens. The length of the solenoid is such that the two ends are then situated at points where the field is practically zero.

The principle upon which the measurement is based may be described as follows. Each turn of the coil, as far as its magnetic properties are concerned, may theoretically be replaced by a flat plate, homogeneously magnetized in a direction perpendicular to its face. The plate would have the same thickness and periphery as one turn of wire; one side of it would carry a uniformly distributed positive magnetic charge and the other a similar negative charge, so that adjacent positive and negative charges of the successive plates neutralise each other. At the point where the winding direction changes, however, like charges face, and therefore reinforce each other. Magnetic charges would also occur at the end faces of the solenoid, but these are outside the field and therefore experience no force. The charge at M, on the other hand, encounters a force that is proportional to the induction at that point. The magnitude of the magnetic charge is proportional to the current flowing in the solenoid and, if this current is constant, the force on the solenoid must be a measure of the induction at M.

In fact, the force on the solenoid is not proportional to the induction along the axis at M, but to the flux Φ_M contained per turn at M. If we denote the axial components of the magnetic field strength and the magnetic induction by H_{ax} and B_{ax} respectively, and the (constant) density of the magnetic surface-charge at M by σ, the force on the solenoid will be:

$$K = \int \sigma H_{ax}\, dS = \frac{\sigma}{\mu_0} \int B_{ax}\, dS = \frac{\sigma}{\mu_0} \Phi_M,$$

where S is the cross-section of the solenoid at M, over which the integration must be performed, and μ_0 is the permeability of a vacuum. Φ_M is of course proportional to the average value of B_{ax} for the cross-section, and what is measured is, in effect, this average.

A correction can be made, however. If we consider the z co-ordinate along the axis of symmetry and denote the flux per turn (radius r) and the magnetic induction on the axis at the point z by Φ_z and B_z respectively, it follows from potential theory that:

$$\Phi_z = \pi r^2 \left\{ B_z - \frac{r^2}{8} \frac{d^2 B_z}{dz^2} + \frac{r^4}{192} \frac{d^4 B_z}{dz^4} - \ldots \right\}.$$

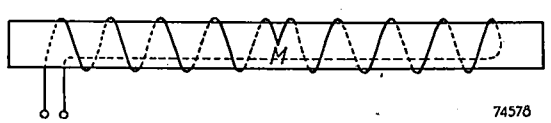


Fig. 2. For the measurement of the field of a magnetic electron lens a long thin solenoid ,the winding direction of which is reversed at the centre M, is used as the actual measuring head. The axis of the solenoid is made to coincide with that of the lens. The force acting on the solenoid is proportional to the magnetic induction at M and to the current flowing in the coil. The ratio of length to thickness is much greater than as shown in the figure.

Here, then, Φ_z is expressed as a series of powers of r. B_z is the term to be evaluated and Φ_z is the value measured. The derivatives d^2B_z/dz^2 etc., appearing in the corrective terms, are actually unknown, but we can, as an approximation, replace B_z by its average value $\Phi_z/\pi r^2$ for the area of one turn. Then:

$$B_z = \frac{1}{\pi r^2} \left\{ \Phi_z + \frac{1}{8} \frac{d^2 \Phi_z}{dz^2} r^2 - \frac{1}{192} \frac{d^4 \Phi_z}{dz^4} r^4 + \ldots \right\}. \quad (1)$$

Φ_z is measured as a function of z; hence the derivatives can also be ascertained as a function of z, enabling the corrections to be computed. It will be seen that the corrections are smaller according as r (the radius of the solenoid) is reduced.

[1]) M. van Ments and J. B. Le Poole, Numerical computation of the constants of magnetic electron lenses, Appl. sci. Res. **B 1**, 3-17, 1947.

The solenoid can be suspended horizontally on wires (*fig. 3*). When a direct current is passed through it, a deflection occurs owing to the force acting upon it, and this can be measured optically, for example. The deflection is a measure of the induction at the point of reversal in the direction of winding.

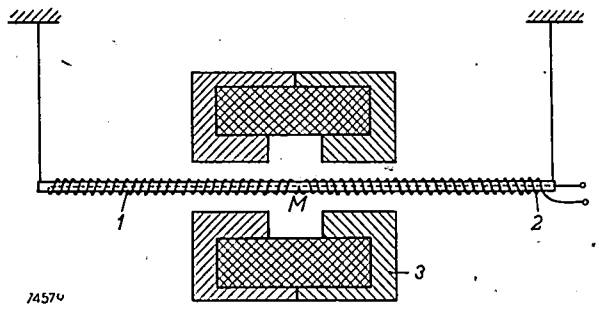
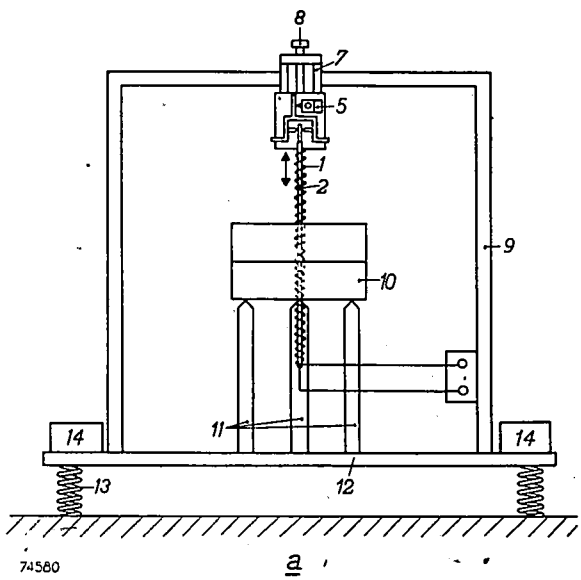


Fig. 3. Principle of system adopted by Van Ments and Le Poole [1]). The rod *1* on which the measuring solenoid *2* is wound, is suspended horizontally by means of thin wires, so as to be concentric with the axis of the electron lens *3*. A horizontal force is exerted on the solenoid in the field of the lens and the resultant displacement, which is measured optically, is a measure of the field strength at *M*.

In the present case, however, another method is adopted; the exploring coil is connected to a source of A.C. voltage, a variable resistor being used to adjust the current flowing in the solenoid to a certain value which is kept constant throughout the process of measurement. The solenoid is thus subjected to a periodic force proportional to the induction at the reversal point in the winding, and accordingly vibrates. The amplitude of this vibration is measured with the aid of an ordinary Philips gramophone pickup of the piezo-electric crystal type. The fact that the signal can be amplified means that measurement is at once simple and accurate.

The solenoid is suspended vertically from two flat springs (*fig. 4*) which also carry a vertical metal plate. This plate follows the vibration of the solenoid. The plate is in contact with the sapphire needle of the pickup and transmits the vibration to the needle by reason of the friction between plate and needle. Should very large amplitudes occur as a result of jarring etc, the needle slips on the plate and damage is prevented. The flat springs and pickup are mounted on a slide in a vertical track which actually forms part of the framework (*9*, fig. 4*a*). To effect the measurement, the electron lens is supported so that its axis coincides with that of the solenoid; the solenoid is then moved vertically along the slide so that the



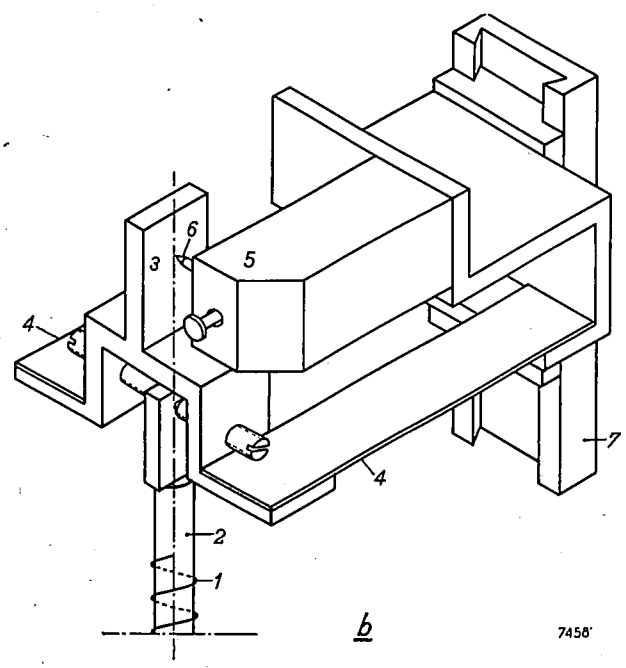


Fig. 4. Schematic view of the system with a vertically suspended solenoid. The coil *1* is wound on an aluminium rod *2* which also serves as return line for the current. An alternating current is passed through the coil, causing this, together with the rod *2* and plate *3*, to vibrate vertically. Vibration is made possible by mounting the system on two horizontal flat springs *4* (see detail in fig. *b*). The amplitude of the vibration is a measure of the magnetic induction to be determined, and is measured by means of a standard Philips gramophone pickup *5*. The sapphire needle *6* of the pickup rests against the vertical plate *3*. For normal measurements, where the amplitude is only a few hundredths of a millimetre, friction causes the needle to follow the motion of the plate. In the event of jarring, the needle slips and damage is avoided. The reversal point in the solenoid winding is positioned at the desired point of measurement by sliding the whole system vertically in a guide *7*; this is effected by means of a screw *8* (not shown in *b*). The guide forms part of the framework *9*. The magnetic lens *10* to be measured is supported on screws *11* for vertical adjustment. Baseplate *12* is mounted on four steel springs *13* to eliminate external vibration, and weights *14* serve to increase the vibration period of the whole system to 0.5 sec.

induction can be measured as a function of the position along the axis of the lens.

Apart from the vibration of the solenoid, the friction plate may very easily transmit undesirable movements to the pickup. e.g. external vibrations originating in the table on which the apparatus is set up. The effect of these vibrations is so marked that its elimination is essential to the effectiveness of the method. The whole apparatus, with the exception of the amplifier and measuring instrument, is therefore placed on a board supported by four vertical compression springs (fig. 4*a*). Weights placed on the plate, moreover, reduce the natural frequency of the whole to about $\nu_0 = 2$ c/s. As the lowest frequencies encountered in floors, walls and tables are usually somewhere between 12 and 20 c/s, the frequencies ν of such extraneous vibrations lie well above ν_0. They are thus transferred to the measuring apparatus attenuated in amplitude [2]) in the ratio $(\nu_0/\nu)^2$. The effect of the external vibration is accordingly reduced by a factor of more than 100.

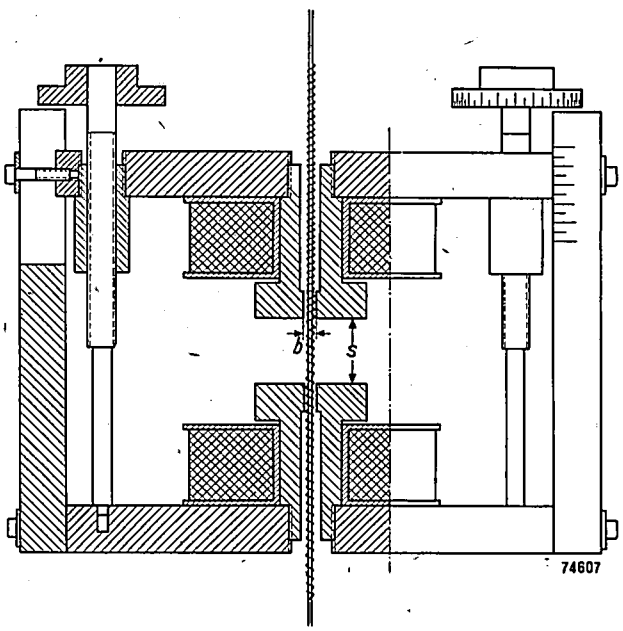


Fig. 5. Magnetic electron lens with bore b and variable pole-piece spacing s. This lens was specially designed in order to ascertain the effect of the ratio s/b on the characteristics of the lens. s/b is variable between 0 and 8.

[2]) J. A. Haringx, Philips tech. Rev. **9**, 16-23, 1947.

In the design of electron lenses the dependence of the focusing characteristics on the ratio of pole-spacing s to diameter of the bore b is of particular interest. A special lens has therefore been constructed,

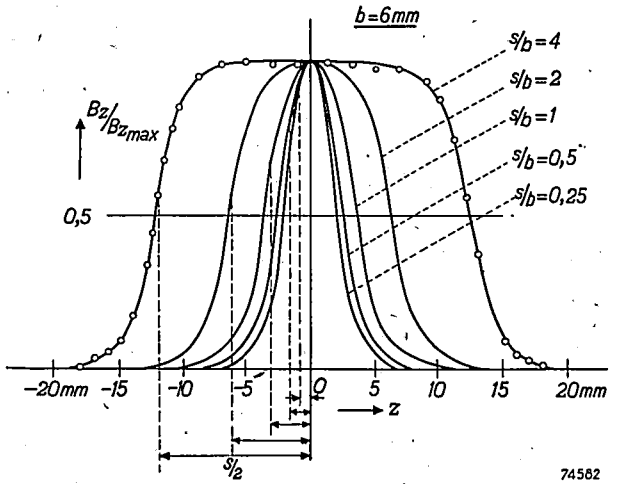


Fig. 6. Results of measurements taken from the lens shown in fig. 5. Relative value of the magnetic induction $B_z/B_{z\max}$ on the axis, for various values of s/b.

in which the ratio of s to b is variable between fairly wide limits, viz. 0 to 8 (*fig. 5*). The results of measurements taken from this lens are reproduced in *fig. 6*. These provide a check on the accuracy of the computed values as well as an insight into the effect of saturation in the iron.

The derivatives dB_z/dz and d^2B_z/dz^2, the second of which occurs in the correction terms for B_z in expression (1), can also be measured directly if a small modification be made in the apparatus described above. Instead of the long solenoid, a short coil is placed on an aluminium rod (*2* in fig. 4) suspended from the flat springs. The coil may be visualised as the equivalent of evenly distributed positive and negative magnetic charges at the two end faces of the coil. Since the spacing of these faces is a constant of the apparatus, the resultant force is proportional to dB_z/dz at the location of the coil.

If the short coil be made with a reversal of winding direction at the centre, the force to which it is subjected will be proportional to the difference between the forces on the two halves of the coil. The force is thus proportional to the *difference* in the values of dB_z/dz for the two halves of the coil and, hence, also proportional to d^2B_z/dz^2 for the location of the whole coil. In principle it would be possible to measure d^3B_z/dz^3 with a combination of two divided coils, and so on.

III. MEASUREMENT BY THE PROTON RESONANCE METHOD

The physical principles underlying the two methods already described are so well-known that the emphasis in sections I and II has been on the mode of application rather than on the principles themselves. The third method, however, which is characterized by exceptional accuracy, involves the application of new principles. It is an incidental result of new developments in the sphere of atomic

magnetism. A brief outline of these developments will now be given, in so far as they are necessary to an understanding of this method of measurement.

Nuclear spin magnetism and proton resonance

The principle of this method of measurement can be explained broadly and simply in terms of the quantum theory. The proton (nucleus of the hydrogen atom) has a constant mechanical angular-momentum known as the "spin" which is analogous to the spinning of a top. Furthermore, the proton has a constant magnetic moment μ_p. The vectors of the spin and magnetic moment always lie in the same direction, as will be obvious if the proton be imagined as a rotating electrical charge.

According to the quantum theory a proton in a magnetic field B_0 can assume only two orientations, viz. those in which the magnetic moment, and hence also the spin, are parallel, and antiparallel, to B_0. An amount of energy $\Delta E = 2\mu_p B_0$ is needed to change the orientation of a proton from the parallel to the antiparallel position [1]). Thus, spins in the antiparallel position possess an energy which is $2\mu_p B_0$ higher than the energy level of parallel spins. The distribution of the spins among these two energy levels will be in accordance with Boltzmann's law, which states that the concentration in an energy level is proportional to $e^{-E/kT}$, where E is the energy of the level, and k is Boltzmann's constant. The higher the energy the lower the concentration. Let us denote the number of nuclear spins at the lower and higher levels by N_1 and N_2 respectively; then:

$$\frac{N_2}{N_1} = e^{-\Delta E/kT} = e^{-2\mu_p B_0/kT}. \quad . \quad . \quad . \quad . \quad (1)$$

Since, normally, $2\mu_p B_0 \ll kT$, the surplus of protons in the lower level $(N_1 - N_2)$ is very small compared with the total number $(N_1 + N_2)$ of protons [2]).

Suppose that a weak alternating field $B_1 \cos 2\pi\nu t$, of frequency ν be established perpendicular to the field B_0: quantum theory states that energy can be derived from, or imparted to this field only in quantities which are integral multiples of the energy quantum $h\nu$, where h is Planck's constant. If the value of ν is such as will satisfy the equation $h\nu = 2\mu_p B_0$, one single energy quantum will be just enough to change one spin from the parallel to the antiparallel orientation. The chance of parallel spins absorbing energy quanta and becoming antiparallel spins is then greatly increased, i.e., resonance occurs. The frequency at which this takes place is called the resonance frequency ν_r. Hence:

$$B_0 = \frac{h}{2\mu_p}\,\nu_r. \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Measuring principle

The above considerations have led to the following method of measuring magnetic fields.

Protons are introduced into the field B_0 to be measured, in the form of hydrogen nuclei from substances rich in hydrogen. A weak alternating field is established perpendicular to the field B_0, the frequency being so adjusted that equation (2) is satisfied. The absorption of energy which occurs at the resonance frequency, is employed as an indicator for the adjustment, and the apparatus therefore includes a cathode-ray oscilloscope, from which the occurrence of this energy absorption can be noted visually. As the resonance frequency is related to B_0 (see eq. 2), the measurement of the magnetic field is now reduced to a measurement of frequency. Since high frequencies can be measured very accurately with little difficulty and, as the onset of resonance is very sharp, this method of measuring magnetic induction is much more precise than other, traditional methods.

For the comparison of magnetic fields it is only necessary to determine the resonance frequencies in these fields. Nevertheless it is also possible to effect absolute measurements, seeing that the exact value of $h/2\mu_p$ is known. With ν_r in cycles per second, we may write:

$$\begin{aligned} B_0 &= 2.3489 \times 10^{-8}\,\nu_r \text{ Wb/m}^2 \\ &= 2.3489 \times 10^{-4}\,\nu_r \text{ gauss}. \quad . \quad . \quad . \quad (3) \end{aligned}$$

A resonance frequency of 30 Mc/s therefore occurs in a field of 0.7047 Wb/m^2 = 7047 gauss.

---

[1]) B is written instead of H because of the definition of the magnetic moment. The proton is visualised as a very small circuital current, and the local magnetic field (in vacuo) is defined by the magnetic induction B. The magnetic moment of a plane circuital current i encircling a plane surface S is defined as $m = iS$. The couple imparted to such a circuit in a magnetic field B whose lines of force make an angle ϑ with the normal at S, is $iSB \sin\vartheta$. Sometimes the magnetic moment is defined as $m' = \mu_0 iS = \mu_0 m$. The expression for the resultant couple is then $m'B \sin\vartheta/\mu_0 = m'H \sin\vartheta$. ($H$ = magnetic field strength).
These last expressions for the magnetic moment and resultant couple are automatically arrived at if the magnetic moment be imagined as the outcome of a combination of equal but opposed magnetic charges. We shall, however, adhere to the first-mentioned definition (see for example Slater and Frank, Electromagnetism, McGraw Hill, New York 1947, p. 59).

[2]) $\mu_p = 1.410 \times 10^{-26}$ A.m^2 and $k = 1.375 \times 10^{-23}$ J/degree. With a field $B_0 = 1$ Wb/m^2 (10 000 gauss) and a temperature of 300 °K, $2\mu_p B_0/kT \approx 10^{-5}$ and N_2/N_1 is accordingly very little less than unity.

The value of $h/2\mu_p$ can be derived very exactly from measurements taken by Bloch and Jeffries [3]) who set out to determine the value of μ_p, making use of the proton resonance described. The principle on which their measurements were based is as follows. Since protons carry an electric charge e ($e = 1.6019 \times 10^{-19}$ coulombs) they will experience a force when moving in a magnetic field. When the movement is in a plane perpendicular to a *homogeneous* field (B_0, say) they will describe circular paths, which they traverse with a frequency ν_c (independent of velocity) given by:

$$\nu_c = \frac{eB_0}{2\pi M_p} \quad \ldots\ldots\ldots\ldots \quad (4)$$

(M_p = mass of the proton = 1.6722×10^{-27} kg).

Suppose we have a cyclotron with a magnetic field strength B_0; ν_c is then the frequency for the Dee voltage in order that the cyclotron may accelerate protons [4]). Bloch and Jeffries employed a small, specially constructed cyclotron and determined the frequency at which this functioned with protons; this was the frequency ν_c. In the same cyclotron field they then ascertained the proton resonance frequency ν_r which, in accordance with equation (2) is given by:

$$\nu_r = \frac{2\mu_p B_0}{h} \quad \ldots\ldots\ldots\ldots \quad (5)$$

From (4) and (5) it follows that:

$$\frac{h}{2\mu_p} = \frac{\nu_c}{\nu_r} \cdot \frac{2\pi M_p}{e}.$$

The interesting point is that the induction B_0, which cannot be independently determined with any great accuracy, is thus eliminated from the equations. According to the measurements of Bloch and Jeffries, $\nu_r/\nu_c = 2.7924 \pm 0.0002$. Using the above-mentioned values of e and M_p, we find that $h/2\mu_p = 2.3489 \times 10^{-8}$, which value appears in equation (3). (Since h is known, the value of μ_p itself can also be computed.)

Before we proceed to a description of the measuring equipment itself, it is necessary to discuss in greater detail two further points, viz. the magnitude of the energy absorption, and the "line width" of this absorption.

Magnitude of the energy absorption

As already mentioned, there is, in respect of every parallel spin, a certain chance that it will absorb an energy quantum from the field and thus change to the antiparallel orientation. On further consideration, however, this is in itself not enough to explain the energy absorption observed. There is, in effect, an equally good chance that the alternating field may induce a proton with antiparallel spin to release an energy quantum and thus assume the parallel position, this process being known as "induced emission". Both chances are proportional to the concentration of energy quanta of the alternating field, that is, to the energy density of this field, i.e. to B_1^2. Now, the number of parallel spins N_1 is greater than that of the antiparallel N_2 (Boltzmann distribution; equation (1)), so that, from the start, more parallel spins change to antiparallel than vice versa; consequently N_2 increase at the expense of N_1, but, as N_2 and N_1 approach each other, this process slows down. In the absence of any other influences, a situation would finally arise in which $N_2 = N_1$, with just as much energy absorbed from the alternating field as is restored to it; the net absorption would then be zero. Energy absorption would then take place only during such time as N_2 were approaching N_1, and the total absorbed energy would be so small as to be undetectable.

There is, however, another mechanism that causes reversal of the spins, and this is associated with the thermal motion. An interchange of energy occurs between the magnetic energy of the protons in the magnetic field B_0, and the kinetic energy of their thermal motion; this process tends to restore the original Boltzmann distribution (1) of the parallel and antiparallel spins. On average, therefore, the spins release energy towards the thermal motion, and an equilibrium is set up in which the spins take up just as much energy per second from the alternating field as they impart to the thermal motion. If this release of energy is sufficiently great, then, a considerable amount of energy can be absorbed continuously from the alternating field. It is this absorption, as already mentioned, which serves as an indicator for the attainment of the resonance frequency.

The protons, which are brought into a state of resonance in the apparatus about to described, are actually those forming the nuclei of hydrogen atoms in a small quantity of water. With pure water, however, the continuous energy absorption from the alternating field is too small to yield a clear indication, the reason for this being the very slow transfer of the magnetic energy of the protons to the thermal motion. This transfer is enhanced by adding paramagnetic ions to the water, e.g. by dissolving ferric or manganese salts in it. Because of the pronounced interaction between the magnetic fields of the protons and the strong fields of the ions, the exchange of energy is much greater and so also is the amount of energy absorbed from the alternating field. The resonance indication is then clear enough for a practical measurement.

---

[3]) F. Bloch and C. D. Jeffries, A direct determination of the magnetic moment of the proton in nuclear magnetons, Phys. Rev. **80**, 305-306, 1950.

[4]) See for example W. de Groot, Cyclotron and synchrocyclotron, Philips tech. Rev. **12**, 65-72, 1950.

Line width; sharpness of the resonance indication

The field of the protons and that of the paramagnetic ions, alternating in time and space, are superimposed on the external field B_0; the protons accordingly exist in fields that exhibit a certain variation about the value of the external field B_0. There is therefore not one specific resonance frequency, but a certain range of frequencies in which resonance occurs, and this range lies on and around the frequency corresponding to B_0. The resonance thus possesses a finite "line width". If the rate of absorption of energy, for a constant value of B_0, is plotted as a function of the frequency, the absorption spectrum reveals the peak shown in *fig. 1a*,

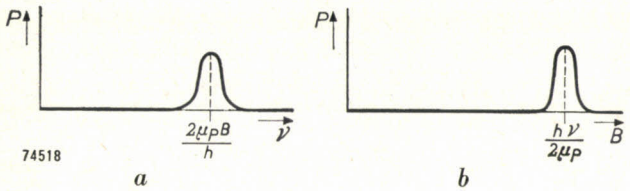


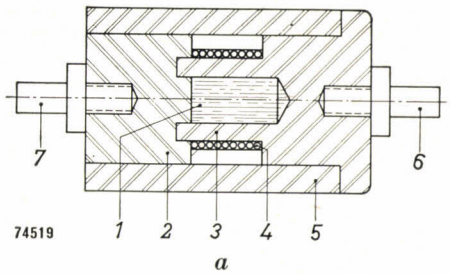
Fig. 1. "Line width" of proton resonance in a magnetic field B with a superposed perpendicular H.F. field (frequency ν). Resonance can be attained in two ways, viz. 1) with constant external field B, by varying the frequency of the H.F. field; 2) with constant frequency, by varying the external field B. For measuring purposes the two methods are combined.
a) With constant external field B, resonance occurs in a small range of frequencies about the frequency value relevant to the field B, in accordance with equation (2). P is the absorbed power.
b) With constant frequency, resonance occurs in a range of magnetic induction values about the value relevant to that frequency, in accordance with equation (2). The line width can therefore also be expressed in Wb/m².

rather than a clearly defined line at $\nu_r = 2\mu_p B_0/h$ (see Eq. 2). Similarly, when the rate of absorption of energy is plotted against the external field B for constant frequency, a straight absorption line is not obtained for $B = h\nu_r/2\mu_p = B_0$, but again a peak of finite width about B_0 (fig. 1*b*). According to whether we are considering fig. 1*a* or 1*b*, the line width can be expressed in c/s or Wb/m². The relationship between the two alternatives follows at once from equation (3): 1 c/s corresponds to 2.3489×10^{-8} Wb/m².

For pure water the line width is roughly 10^{-6} Wb/m² (10^{-2} gauss). The addition of paramagnetic ions increases the width, as is to be expected, since the magnetic fields of these ions increase the "spread" of the local fields around the value of B_0. At the same time, even when pure water is replaced by a solution of ferric or manganese salts, the line remains so narrow (in our case 10^{-5} Wb/m² = 0.1 gauss) that the definition of the resonance indication is not appreciably affected. These figures demonstrate the sharpness of the resonance indication. The inhomogeneity of strong magnetic fields (of the order of 1 Wb/m²) within the space occupied by the measuring head (about 30 mm³), even in the most uniform fields, is usually of the same order of magnitude, so that greater precision in the indication of resonance would in any case be superfluous.

Description of the proton resonance equipment

The equipment about to described was constructed along the lines of a design previously employed by Pound and Knight [5]). A coil, wound round a small "Perspex" container, is contained in a flat coil-holder (*fig. 2*), the axis of the coil lying in the plane of the holder. The holder is inserted in the magnetic field to be measured with its flat face at right angles to the lines of force. An alternating current passed through the coil produces the weak alternating field perpendicular to the field to be measured. The "Perspex" chamber is filled with an aqueous solution of manganese sulphate or ferric nitrate, to provide a certain concentration of manganese or ferric ions. The hydrogen atoms of the water



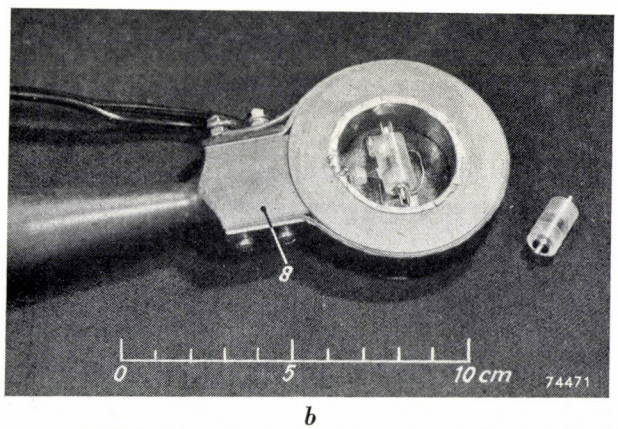


Fig. 2. *a*) Construction of the oscillator coil. The cavity *1* (internal diam. 3 mm, length 4 mm) contains a solution of a ferric or manganese salt. The walls are formed by the "Perspex" components *2* and *3*. *4* is the oscillator coil proper. *5* is a "Perspex" bush. *6* and *7* are brass terminals for the coil.
b) The oscillator coil is fitted in the centre of a flat coil holder *8*. The lead for connection to the capacitor in the oscillator circuit is taken through the handle of the holder; on the left will be seen leads for an auxiliary coil which produces a ripple on the field to be measured. At right: oscillator coil for a different measuring range.

[5]) R. V. Pound and W. D. Knight, A radio frequency spectrograph and magnetic field meter, Rev. sci. Instr. **21**, 219-225, 1950.

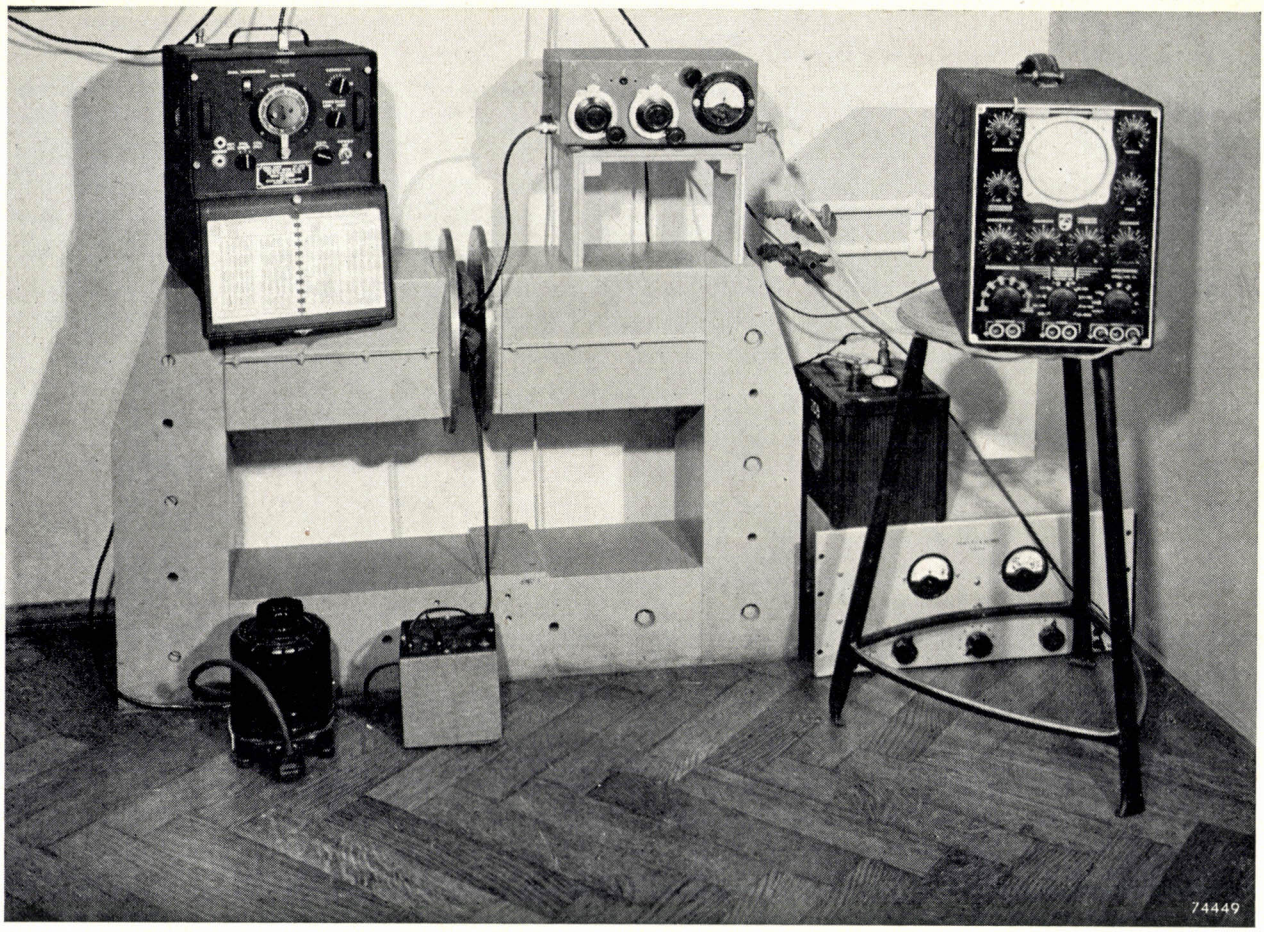


Fig. 3. Measuring the magnetic field of a large permanent magnet (weight 600 kg) by means of the proton resonance equipment. The coil holder is inserted between the pole pieces of the magnet. The electronic equipment is standing on the right-hand pole piece; on the right of the illustration is the oscilloscope for the resonance indication. At right below, high tension unit and battery supplying heater current for the valves. Standing on the left hand pole piece of the magnet is the wavemeter used for measuring the resonance frequency. On the ground at left: a variable transformer to supply the ripple added to the measured magnetic field.

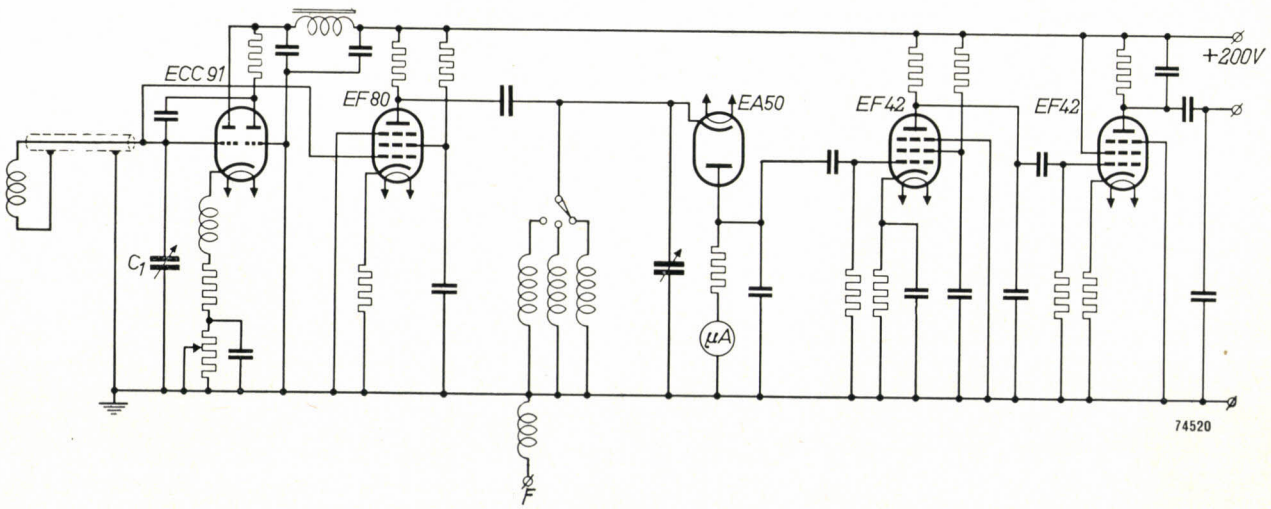


Fig. 4. Circuit diagram of the proton resonance measuring equipment. On the left is the exploring coil connected to capacitor C_1, and a double triode operated as negative resistance to form an oscillator. On the right-hand side of this valve is the circuit for amplification and detection of the voltage modulation in the oscillator circuit. F connecting point for wavemeter.

contain the protons which are to be brought into resonance in the measured field, and the paramagnetic manganese or ferric ions ensure an effective transfer of the energy absorbed at resonance to the thermal motion.

A low-capacitance coaxial cable connects the coil to the electronic equipment (*fig. 3*). The coil, together with a capacitor C_1 and the cable capacitance, constitute a tuneable oscillator circuit (*fig. 4*); this is coupled to a negative resistance circuit formed by a double triode ECC 91 to complete the oscillator circuit. The frequency of this circuit is tuned to the resonance frequency of the protons by adjusting C_1. In order to cover a wide frequency range, the coil is interchangeable (fig. 2). There are in all three coils, each with a different number of turns.

By means of the biasing resistor common to the two triodes, the oscillator is adjusted to the threshold of oscillation. Under these conditions the voltage on the circuit is very sensitive to the small increase in damping that occurs with proton resonance. In the rim of the coil-holder there is a flat auxiliary coil, through which a 50 c/s current is passed during measurement; this causes the field at the location of the exploring coil to alternate between $B_0 + B_s$ and $B_0 - B_s$. B_0 represents the value of the field to be measured and B_s the amplitude of the field of the auxiliary coil, at the "Perspex" container. Proton resonance and the accompanying absorption of energy will occur twice per cycle of 1/50 sec when the oscillator is tuned to a frequency ν such that (see *fig. 5*):

$$B_0 + B_s \geqq \frac{h\nu}{2\mu_p} \geqq B_0 - B_s.$$

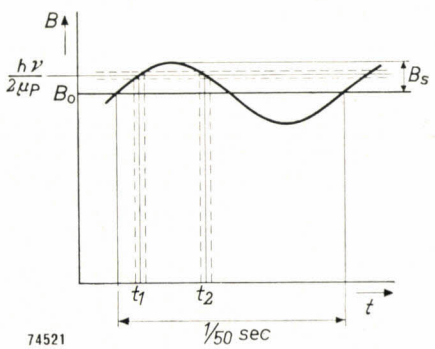


Fig. 5. The auxiliary coil in the coil holder superimposes a ripple of amplitude B_s on the measured field B_0. At a given frequency ν of the H.F. field, resonance occurs when the value of B lies between the two dotted horizontal lines. This occurs twice per cycle of 1/50 sec., viz. about the times t_1 and t_2. The oscilloscope thus reveals absorption peaks at the points corresponding to these times. If, now, the ripple amplitude is reduced under the conditions as shown, where $h\nu/2\mu_p > B_0$, t_1 and t_2 move towards each other. The oscilloscope shows the absorption peaks approaching each other, coinciding and disappearing.

Twice per cycle, then, the resonance requirement $\nu = 2\mu_p B/h = \nu_r$ (see Eq. 2) will be met, and at these moments (t_1 and t_2 in fig. 5), the voltage of the oscillator circuit will vary by a small amount. Owing to the finite width of the resonance range (shown by the horizontal dotted lines), the instants of time are increased to short time intervals. The high-frequency signal across the circuit is modulated by these dips in voltage. The fundamental frequency of the modulation is thus 50 c/s.

The signal is first amplified by an EF 80 valve, the anode circuit of which is tuned by means of a variable capacitor and one of three coils operated by a switch, corresponding to the exploring coil used in the coil holder. Next the signal is rectified and smoothed by a combination of capacitor and resistor with low time constant, to avoid attenuation of the peaks, which occurs with a frequency of 50 c/s. Finally, the signal, which now contains only

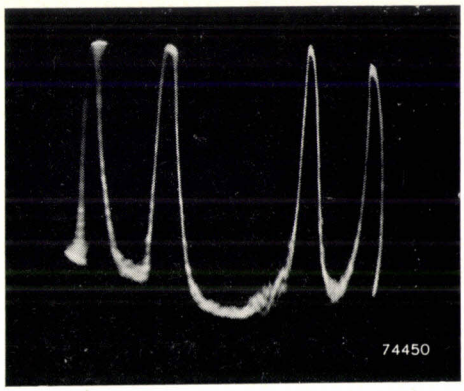


Fig. 6. Oscillogram of modulation of the oscillator amplitude. The frequency of the time base in this case was 25 c/s: hence the two pairs of peaks.

the absorption peaks, is again amplified and applied to a cathode-ray oscilloscope. The time base is synchronised with the 50 c/s voltage on the auxiliary coil. When resonance conditions are obtained, two absorption peaks are seen on the screen, well above the noise level (*fig. 6* shows two pairs of peaks, this being due to the use of a 25 c/s time base).

To measure a magnetic field B_0, the current in the auxiliary coil is gradually reduced. In the case of fig. 5, when the frequency is too high, the traces of the two peaks will be seen to come together, coincide and then disappear; when the frequency attains the value $\nu_r = 2\mu_p B_0/h$ appropriate to B_0, the peaks remain stationary (*fig. 7*). (Because of the finite width of the resonance zone, the peaks become wider according as B_s is reduced. When the ripple on B_0 lies wholly within the resonance zone, the resonance is continuous and, with it, the absorption, in which case the H.F. voltage on the

oscillator circuit is no longer modulated, but has a constant value. As only the modulation is visible, the peaks then disappear from the screen.) The resonance adjustment, then, consists of tuning the oscillator till the separation of the peaks becomes

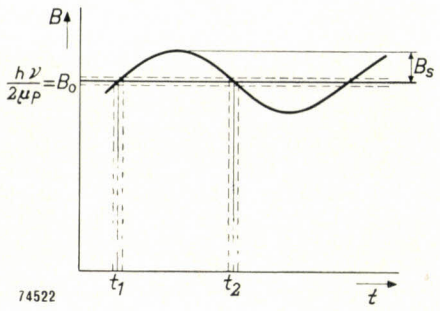


Fig. 7. The frequency ν being so adjusted that $h\nu/2\mu_p = B_0$, the absorption peaks remain stationary when the ripple amplitude is reduced. This is the indication that the frequency is correct.

insensitive to changing the current in the auxiliary coil. The frequency is then carefully measured, e.g. by means of a calibrated auxiliary oscillator beating with the signal. Finally, the value of the field is obtained with the aid of equation (3).

With little trouble the frequency can be measured with a relative error of less than 10^{-4} but, if a smaller accuracy is sufficient, use can be made of a previously prepared chart showing the frequency plotted as a function of the reading of the capacitor C_1.

The amplitude of the alternating field

One or two points will now be mentioned concerning the choice of amplitude B_1 of the high-frequency alternating field. Any increase in B_1 results in an increase in the concentration of energy in this field, and more energy quanta are then available for raising protons to the higher energy level. In the resultant new condition of equilibrium, the difference in the concentration of the two possible levels would then be smaller, which means that the departure from the Boltzmann distribution would be greater and the release of energy to the thermal motion accordingly increased. Since this release of energy is equal to the absorption of energy from the alternating field, the absorption will likewise increase. It will be clear, however, that, according as B_1 is increased further, the departure from the Boltzmann distribution approaches a limit, viz. the point corresponding to equal concentrations in the two energy levels (see page 57). Simultaneously the release of energy to the thermal motion would also approach a limiting value, with a corresponding limit to the energy absorption from the alternating field. An indefinite increase in the absorption peaks seen on the oscilloscope cannot therefore be expected as B_1 is increased. In practice it is observed that while the peaks first increase in size, they subsequently *collapse* with further increase in B_1. That this must be so is easily demonstrated. In the description of the equipment it was mentioned that the oscillator must be adjusted to the threshold of oscillation, as it is then most sensitive to variations in the damping. In order to increase B_1 the oscillator must be made to oscillate more strongly and the sensitivity to variations in the damping then drops; the increase in damping and the drop in sensitivity thus counteract each other. When the absorption approaches its maximum value, the decrease in sensitivity gains the upper hand and the absorption peaks shrink in size.

Classical considerations

Mention should be made of the fact that the subject of proton resonance can be viewed in the light of the laws of classical mechanics [6]). This is a fairly complicated procedure however and gives a less straightforward picture than the simplified explanations given here on the basis of the quantum theory. In the classical theory the proton is regarded as a kind of top, which executes a precessional motion under the influence of magnetic fields [7]). The transfer of energy to the thermal motion takes place with a certain delay which is expressed as a certain relaxation time. This relaxation time is considerably reduced by the addition of paramagnetic ions.

Apparatus for fields up to 1.4 Wb/m² (14 000 gauss)

The measuring range of the equipment described above is roughly 0.16 to 0.82 Wb/m² (1600 to 8200 gauss), corresponding to a frequency range for the oscillator of 6.8 to 35 Mc/s. Another unit has now been made up, which measures field strengths up to 1.4 Wb/m² (14,000 gauss) (*fig. 8*). The oscillator

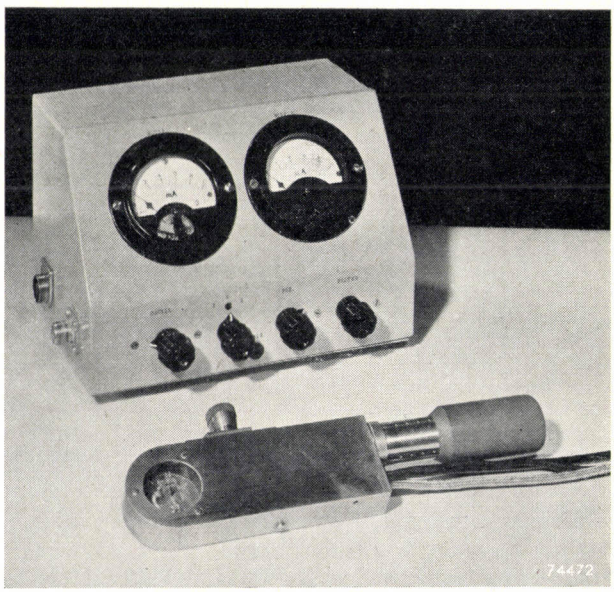


Fig. 8. The proton resonance measuring unit for fields of from 0.5 to 1.4 Wb/m² (5000 to 14 000 gauss). The oscillator frequency is variable from 21 to 60 Mc/s, this being possible because the whole oscillator is housed in the coil holder, thus eliminating the capacitance of the cable between valve and oscillator coil. The double triode is in the handle, that is, as far as possible from the magnetic field to be measured. In strong magnetic fields the triodes will not function. Behind the coil holder will be seen the unit for amplification and detection of the signal. (The lead from the coil holder is shown disconnected.)

[6]) F. Bloch, Nuclear induction, Phys. Rev. **70**, 460-485, 1946.
[7]) See also H. G. Beljers and J. L. Snoek, Gyromagnetic phenomena occurring with ferrites, Philips tech. Rev. **11**, 313-322, 1950.

frequency range in this case is extended to 60 Mc/s, this having been made possible by mounting the whole oscillator in the coil holder. This dispenses with the lead between the coil and the oscillator valve, with a corresponding reduction in the circuit capacitance. It is probable that at the upper limit the measuring range could be extended to 2.0 Wb/m^2 (20 000 gauss). A unit for lower field strengths (lower limit approximately 10^{-2} Wb/m^2 (100 gauss) is now in course of development.

Applications of the method

The apparatus described — which is also very suitable for demonstrating the phenomenon of nuclear spin resonance — can be used for checking the variation of induction between the poles of permanent magnets during the course of time. It is also useful for ascertaining the effect of certain external influences such as mechanical impact or vibration of the magnet, temperature variation and so on. As an example, the table below gives details of the field strength at the centre between the pole pieces of a large permanent magnet of "Ticonal" (see Fig. 3). This magnet was fabricated in the Philips laboratories and is used mainly for calibrating coils for fluxmeters.

Table. Induction between poles of large "Ticonal" magnet.

Date	Treatment	Induction in Wb/m^2 as measured after treatment
25-11-51		0.4351
3-12-51		0.4350
16-1-52		0.43325
16-1-52	Heating from 20 to 49 °C	0.4301
16-1-52	Cooling down to 19 °C	0.4326
18-1-52		0.4326
18-1-52	Heating from 20 to 68 °C	0.4284
18-1-52	Cooling down to 20 °C	0.4325

From the measurements it is apparent that a rising temperature cycle results in a drop in the induction, at least, when this occurs for the first time; this process may be regarded as a sort of ageing. It appears from both the temperature cycles that the temperature coefficient of the magnet is roughly 0.85×10^{-4} Wb/m^2 per °C.

Summary of I, II and III. Three instruments for the measurement of magnetic fields, based on three entirely different principles, are described. The first of these makes use of an extremely small coil (outside diam. 1 mm, length 1.5 mm), mounted on the shaft of a small synchronous motor and rotated with constant angular velocity in the field to be measured. The voltage generated is collected by contact rings and springs and is taken to an amplifier and then rectified. The resultant D.C. voltage is a measure of the magnetic induction. The influence of the mains frequency on the induced signal voltage is compensated by a suitable response curve of the amplifier. The measuring range is very wide, viz. 10^{-4} to 3 Wb/m^2 (1 to 30,000 gauss), and the accuracy is 3%.

The second instrument was developed especially for the measurement of fields along the axis of magnetic electron lenses. A long thin solenoid, the winding direction of which is reversed at the centre, is suspended vertically by two flat springs and is mounted so as to coincide with the axis of the electron lens. When a constant 50-cycle alternating current is passed through the solenoid, an alternating force acts upon the coil, this being proportional to the magnetic induction at the point where the winding changes direction. In consequence, the solenoid vibrates vertically, and this vibration is communicated to the sapphire needle of a gramophone pickup. The A.C. voltage thus obtained, which is amplified, is a measure of the induction. In practice, it is essential that the apparatus be so mounted as to be free from external vibration. By this method the field at any point on the axis of a magnetic electron lens can be measured simply and quickly with an error of less than 1%.

In the third method, a small container of water is placed in the field to be measured. A weak alternating field set up at right angles to this field, brings the hydrogen nuclei of the water molecules into resonance. The frequency at which this proton resonance occurs is proportional to the strength of the field to be measured (30 Mc/s in a field of 0.7047 Wb/m^2). Measurement of the magnetic induction is thus reduced to measurement of this resonance frequency, which can be carried out with a high degree of accuracy. The method is very suitable for the accurate, absolute measurement of strong homogeneous magnetic fields (error less than 0.01 %). The measuring range is 0.16 to 0.82 Wb/m^2. Another model has a range of 0.5 to 1.4 Wb/m^2. This method has been employed in the Philips laboratories at Eindhoven to check the stability of a large permanent magnet used for calibration purposes.

THE EFFICIENCY OF FLUORESCENCE IN CATHODE-RAY TUBES

by A. BRIL and H. A. KLASENS.

535.376:621.385.832

In the manufacture of cathode-ray tubes it is desirable to have quantitative data on the fluorescent efficiency of the screen, i.e. the efficiency of conversion of electron energy to luminous energy. Special measurements and theoretical considerations make it possible to determine the conditions for optimum screen brightness with a given electron energy, in particular for screens coated with a layer of aluminium.

Introduction.

The fact that certain minerals fluoresce when subjected to an electrical discharge in a rarefied gas was noted by E. Becquerel about 1860. After the discovery and subsequent study of cathode rays by Hittorf (1869) and Crookes (1879), a special investigation of this kind of fluorescence was made by the latter in such substances as CaS, BaS, diamond, ruby etc. Lecoq de Boisbaudran (1883) made considerable additions to the list of substances that exhibit this phenomenon, which became known as cathodoluminescence.

Lenard (1903), who had done a great deal of pioneer work in the fields of both cathode rays and luminescence, carried out a quantitative investigation into the phenomenon of cathodoluminescence. According to him, the relationship between the brightness L of a screen exposed to cathode rays (electrons), the current density j and the acceleration voltage U, can be formulated as:

$$L = kj(U - U_0) \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where k and U_0 are constants which differ between one case and another. The quantity U_0, i.e. the value of the acceleration voltage below which no noticeable light appears, was given as 300 V for the phosphor CaS-Bi, 1700 V for ZnS and 6000 V for uranium glass.

Other research workers have since carried out more accurate investigations into cathodoluminescence at low acceleration voltages[1]) and have arrived at formulae of the following kind:

$$L = k' j U^n \quad (n > 1) \quad . \quad . \quad . \quad . \quad (2)$$

The efficiency η of the luminescence, i.e., the ratio the radiation emitted per unit area of the screen to the energy per unit area of the electron beam, jU, is thus defined by the equation $\eta = k'' U^{n-1}$.

It will be clear that a formula such as this, in which the efficiency increases steadily with U, can be valid only for a limited range of values of U, since η can of course never be > 1.

The efficiency of cathodoluminescence is nowadays of great practical importance in many technical applications—e.g. for oscilloscopes, television tubes, electron microscopes, radar. Renewed investigations in this field have therefore been made in the Philips Laboratories at Eindhoven. The work was undertaken with a view to obtaining results of practical value but having also a sufficient generality to enable them to be applied independent of arbitrary factors such as the grain size of the phosphor or the thickness of the screen.

The mechanism of cathodoluminescence; conditions for measurement of efficiency.

The screen of a cathode-ray tube is usually built up of one or more layers of crystals. With ZnS, for example, the size of the crystal may vary between one and some tens of microns; other substances such as willemite (Zn_2SiO_4—Mn) have smaller crystals. When a crystal of this kind is struck by an electron, the electron may leave the crystal without loss of energy (elastic collision); more frequently, however, the electron will impart some, or the whole, of its energy to the crystal (inelastic collision) and the energy thus released is transferred mainly to the electrons present in the crystal. Within the crystal, then, a swarm of "excited" electrons is formed, which possess extra energy of the order of 10 eV. These excited or secondary electrons will, if the substance is not a "phosphor", ultimately yield up this energy to the crystal lattice without radiation, in the form of heat. With phosphors, on the other hand, some of the excited electrons transfer their extra energy directly or indirectly to so-called luminescence centres, which can then emit a quantum of light[2]). Other secondary electrons will leave the crystal immediately (secondary emission), and this

[1]) For details and bibliography, notably for the period since 1925 see A. Bril and H. A. Klasens, Philips Res. Rep. **7**, 401-420, 1952 (No. 6).

[2]) See also A. Bril and F. A. Kröger. Saturation of fluorescence in television tubes. Philips Tech. Rev. **12**, 120-128, 1950 (No. 4).

is important from the point of view of the charging of the screen.

The physicist is interested in all the elements of this process i.e. the number and velocity-distribution of the secondary electrons and the distribution of the radiative and non-radiative energy released in the crystal. Technically, however, the number of emitted light-quanta per primary electron is the main interest. The object of the measurements described in this article and dealt with in greater detail in the article referred to in footnote [1]) is to determine this last-mentioned quantity as accurately as possible.

It has been pointed out that, according to the earlier findings of Lenard as well as to later measurements, the efficiency of cathodoluminescence at low values of the electron energy is very low, but that it rises sharply when the energy is increased. This has been attributed to the small depth of penetration of the electrons when their energy is low. These electrons are for the greater part slowed down in the outer atomic layers of the crystal where, in consequence of lattice defects or impurities, they produce but little luminescence.

To obtain quantitative results which are independent of such surface effects, the investigation of these phenomena is restricted to high electron velocities, corresponding for example to acceleration potentials of 10 to 50 kV. This range of potentials, whereby electrons will penetrate crystals such as ZnS to a depth of 1 to 25 μ, is also the most interesting from the point of view of technical applications.

To yield reliable data of practical value it is desirable that measurements be made with on a comparatively thick phosphor layer (e.g. 0.3 mm). The reason for this is as follows.

As already mentioned, the screen of a cathode-ray tube usually consists of one or more layers of crystals. Consider first a single crystal. Secondary electrons are produced along the path of the primary electrons within the crystal, and some of these secondaries will be responsible for the emission of light. Before the light finally leaves the crystal it will generally be internally reflected a number of times from the faces of the crystal (total reflection), so that the whole crystal, as it were, is filled with light. Hence the crystal may be regarded as a light source, embedded in the screen and emitting light in all directions. Once the light has left the crystal it falls upon neighbouring crystals, which also produce internal reflections and thus pass the light in turn to others. This transfer of the radiation from crystal to crystal is a cause of scattering of the light in the screen which, in itself, does not represent a loss of energy. Apart from the scattering, however, absorption takes place during the passage of the light through the crystals, and this of course is accompanied by a loss of energy.

Consider now the light emitted from both sides of the screen. The two sides do not emit equal quantities of light, the distribution of the total available luminous energy between the two being determined mainly by the scatter and the absorption. These two factors of course depend on the thickness and nature of the screen (crystal size and absorption of the phosphor), as well as on the penetration depth of the primary electrons. If the thickness of the screen be increased, no light will ultimately penetrate to the front of the screen (i.e. the side remote from the electron beam); apart from absorption, all the light will then be emitted from the rear face. This situation would obviously be of little use for practical purposes, but it does lend itself admirably to measurements of the efficiency of the fluorescence.

The measurements are carried out as follows. A beam of electrons of constant current density is allowed to strike a thick screen from a certain angle and the flux emitted per unit area and per unit solid angle is recorded. Since it is permissible to assume that the radiation from a thick screen conforms to Lambert's law (brightness independent of the direction of observation of the screen), it is only necessary to multiply by π the measured specific flux emitted in the direction of the normal in order to obtain the total energy per unit area. If we now divide this by the product of current density and acceleration voltage, that is, by the power per unit area of the primary electron beam, we shall know the efficiency. This efficiency value is to a certain extent influenced by the absorption losses and, if necessary, a correction may be made, based on the fact that the light is actually produced by a very thin surface layer of the phosphor. One half of this light is emitted directly from the back of the screen without appreciable absorption, and the remainder only after diffusion in the phosphor crystals. Owing to absorption, this remainder undergoes a reduction corresponding to the reflectivity R_∞ of an infinitely thick layer of crystals. If we denote the measured efficiency by η_m, and the efficiency thus corrected for absorption losses — i.e. the "intrinsic" efficiency — by η_i, it follows that:

$$\eta_m = (1/2 + 1/2\, R_\infty)\, \eta_i \quad . \quad . \quad . \quad . \quad (3)$$

or:

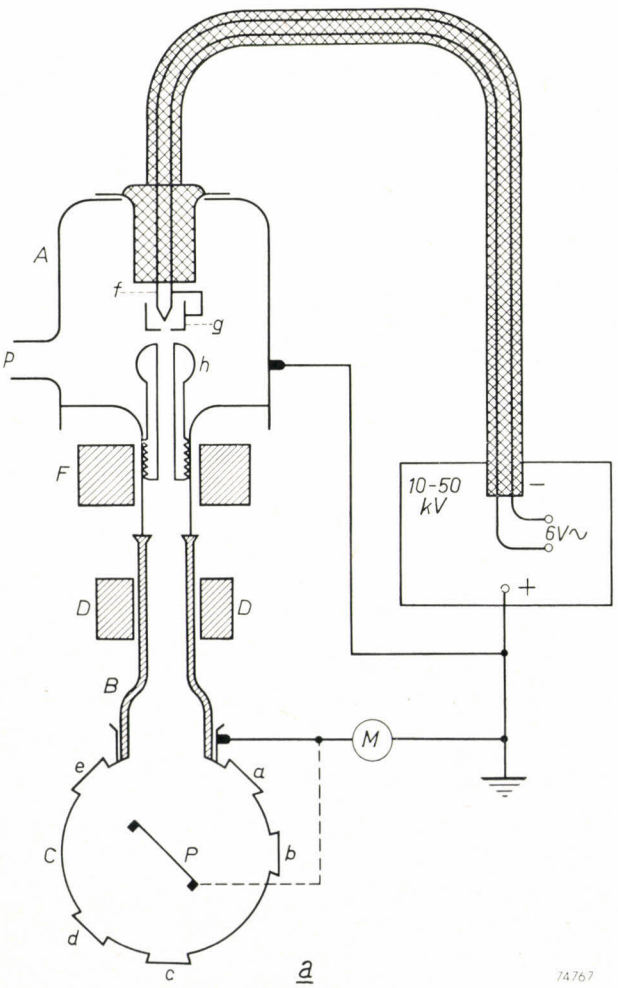
$$\eta_i = \frac{2}{1 + R_\infty}\, \eta_m \quad . \quad . \quad . \quad . \quad (4)$$

The measuring equipment

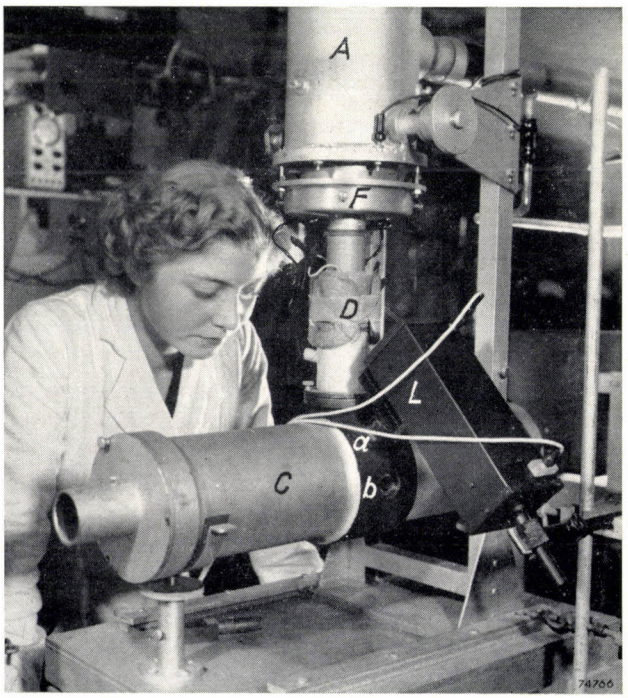
The apparatus consisted of a demountable cathode-ray tube permanently connected to an oil diffusion pump backed by a rotary oil pump. This tube (*fig.* 1 *a* and *b*) comprises 3 sections, *A*, *B* and *C*. Section *A*, which is of the same design as that used in the Philips' electron microscope, contains the electron gun with an oxide cathode *f* having a core of nichrome wire; this cathode is enclosed in a cap *g* having in it an aperture through which the electrons stream. The space between the cap *g* and the hollow anode *h*, which is attached to the metal wall of *A*, is adjustable. The anode and the shield *A* are earthed, whereas the heater *f* and the cap *g* can be given a high negative potential by means of a transformer and rectifier, this potential being varied from the primary side of the transformer. The leads to the cathode are taken into the tube through a porcelain insulator.

Section *C* is a metal cylinder in which an aluminium plate *P* is mounted at an angle of 45°; this plate has in it a number of shallow recesses which can be filled with various phosphors. The plate can slide along the axis of the cylinder, so that the various phosphors in turn can be subjected to electronic bombardment under identical conditions. The fluorescent radiation is observed, and measured, perpendicular to the plane *P*, through a quartz window *a*. Window *d* serves for the measurement, if desired, of the light emitted from the front of a screen; in this case the solid plate *P* is replaced by a plate with openings carrying glass plates with the phosphors applied to them. Windows *b*, *c* and *e* enable the phosphors to be observed during measurement; in particular, window *b* can also be used for measuring the fluorescent light from directions other than the normal, and also for determining its spectral composition.

A focusing coil *F* is used for concentrating the electron beam on a particular point of the test plate, and two deflection coils *D* enable a limited area of the layer of phosphor (1.5 × 2 cm) to be scanned in two perpendicular directions through this point. In this way the electron beam can be made to trace a raster on plate *P*, this being visible through window *a*. The beam current density is obviously higher than the average current density on the screen. There is some risk, especially



a



b

Fig. 1 *a*) Diagram of apparatus for the measurement of fluorescent efficiency. *A* metal chamber, *h* anode, *f* filament, *g* cap, *p* pump connection, *F* focusing coil, *D* deflection coils, *B* glass tube, *C* metal cylinder, *P* target plate, *a*, *b*, *c* and *d* quartz windows. *M* microammeter.
b) Photograph of the equipment. *L* box containing photoelectric cell.

[3]) See Philips tech. Rev. **12**, 33-51, 1950.

with sulphide phosphors, that this may result in saturation of the fluorescence (see note[2])), but current density and focus are so adjusted that saturation cannot occur, or is so small that a correction can easily be made. The plate P and the wall of the tube C are connected together and are earthed across a microammeter M. Electrons leaving the layer of phosphor ultimately all arrive either on the plate P or on the cylinder C. At the commencement of the measuring operation electrons will also strike section B of the tube wall; this part is made of glass, but it has on the inside a layer of metal which is insulated from A and C. The layer is not earthed and therefore quickly acquires a sufficiently high negative potential to prevent secondary electrons from reaching this part of the tube. The current, as read from M, is then the primary electron current flowing to the plate P. The metal plate prevents the phosphor itself from acquiring a negative potential.

For observations on both sides of the screen, i.e. with the phosphors applied to glass plates, potentials of up to 10 kV were employed. No trouble was experienced as a result of charges on the screen.

The radiant energy of the fluorescence in the range of wavelengths 2000 Å — 7500 Å was measured with the aid of photo-electric cells which had been calibrated by comparison with a previously calibrated thermopile, using certain phosphors. Two vacuum photo-electric cells were used, a caesium antimony cell (having nearly constant sensitivity between 2500 and 5000 Å) and a rubidium cell (practically constant sensitivity between 5000 and 7000 Å). The first of these was employed for measurements on phosphors emitting blue or ultraviolet radiations, and the other for green, yellow or red fluorescence. In the latter instance, a separate test was necessary to ensure that no infra-red was present in the radiation, e.g. by repeating the measurement, this time with a filter passing only radiations of $>$ 7500 Å. In some instances use was also made of a selenium barrier-layer cell. *Fig. 2* shows the relative spectral sensitivities of the photocells employed.

Most of the measurements were made with an acceleration voltage of 20 kV. The current density of the beam was about 1 μA/cm^2.

For reasons which will be mentioned later, it was found desirable to know the effect of the temperature of the phosphor on the efficiency, and for this purpose a tube was used of the kind depicted in *Fig. 3*. This tube contains a demountable cathode and, although provided with a focusing coil, has no deflection coils. The stationary beam is slightly converged, but is not sharply focused on the plate P. The temperature of the plate is raised by means of a heating element mounted underneath it and is measured with the aid of a chromel-alumel thermocouple attached to the plate. Secondary electrons

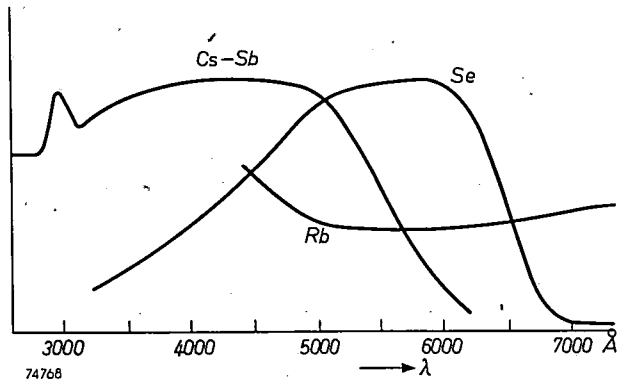


Fig. 2. Relative spectral sensitivities of the photo-electric cells used.

either return to the plate P or are caught by the walls of the tube C, which are silver-plated, both being earthed through a microammeter. In this case, a tube B, internally silvered and insulated from the wall of C, has the same purpose as section B in fig. 1. A voltage of —100 V is applied to this electrode. The beam current employed was in general between 2 and 5 μA with an acceleration voltage of 8 kV.

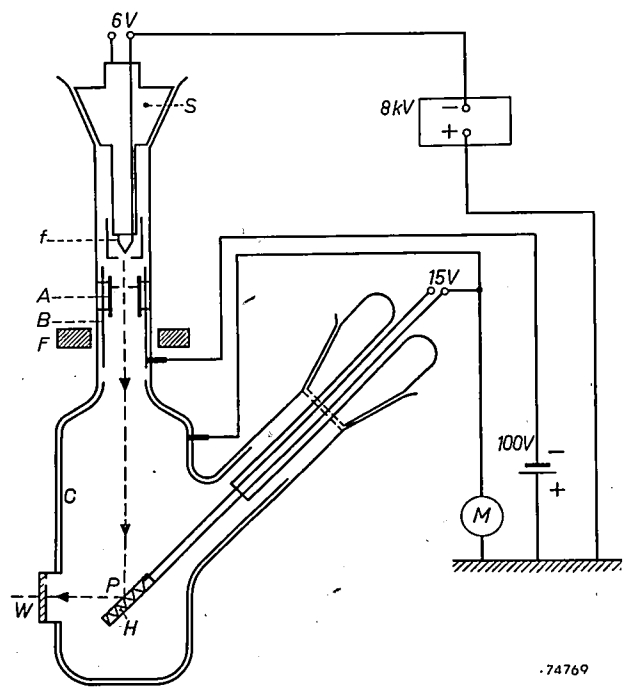


Fig. 3. Diagram of apparatus for the measurement of the efficiency as a function of the temperature. S stopper with interchangeable cathode f. A anode, F focusing coil, B and C internally silvered glass vessels, W quartz window, P target plate. H heating element, M microammeter. The temperature of the phosphor on the target plate is measured with a thermo-couple (not shown).

[4]) The lower limit of the wavelength range (2000 Å) was fixed by the transmissivity of the quartz windows.

Measurement of the efficiency

The absolute measurements of phosphor efficiencies are made with the aid of a calibrated thermopile, placed in front of window *a* (fig. 1), so that the light output in the direction normal to the phosphor layer is measured.

Denoting the area of the thermopile by S, the sensitivity by AS (watt/volt) and the voltage measured at the thermopile by v, the energy reaching the thermopile per second is thus ASv. This energy is emitted by the phosphor within a solid angle S/r^2 (r = distance thermopile to phosphor). The total energy radiated by the phosphor per second, having regard to Lambert's law and making a correction for the transmissivity τ of the window (a), is:

$$I = \pi r^2 Av/\tau. \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

Let i_p represent the primary electron current and U the acceleration voltage; the energy I_p of the primary electron current flowing to P per second will then be $i_p U$, and the efficiency:

$$\eta_m = \frac{I}{I_p} = \frac{\pi r^2 Av}{i_p U \tau}. \quad . \quad . \quad . \quad . \quad (6)$$

We have already seen above, how the intrinsic efficiency η_i can be derived from this measured efficiency η_m.

For a few phosphors the absolute efficiencies were determined in this way. The efficiencies of other phosphors were measured relative to these with the aid of the calibrated photocells described in the previous section.

The results of a number of measurements are shown in *Table I*, in which the values in the first column are those of the measured efficiency η_m under the conditions mentioned. The second column shows the reflectivity R_∞ for light of the same spectral composition as the fluorescent light, and the third column the intrinsic efficiency η_i derived from the first and the second column. The fourth column contains the equivalent luminosity [5]) (lm/W_r) of the fluorescence, calculated from the spectral distribution, and the fifth column gives the product of the values in columns 1 and 4, i.e., the emitted luminous flux in lumens — per watt of the primary electrons (lm/W_e). The sixth column gives an indication of the colour of the fluorescent light.

Table I. Efficiencies of the fluorescence of a number of phosphors under electronic bombardment at 20 kV.

Phosphor	η_m %	R_∞ %	η_i %	$\frac{lm}{W_r}$	$\frac{lm}{W_e}$	Colour
ZnS—10^{-4}Ag—Cl	21			37	7	blue
ZnS—5×10^{-4}Ag —5×10^{-4}Al	23	83	25	35	8	blue
(0.50 Zn, 0.50 Cd)S— 5×10^{-5}Ag—Cl	19.5			500	98	yellow
ZnS—10^{-5}Cu—10^{-4}Al	23	83	25	400	92	green
ZnS—0.015Mn	4	87	4.5	460	18.5	yellow
ZnO—Zn	7	85	7.5	340	24	green
$CaWO_4$	3	80	3.5	70	2	blue
$MgWO_4$	2	97	2	250	5	bluish
$CaO{\cdot}MgO{\cdot}SiO_2$—0.03Ti	7	91	7.5	70	5	blue
Zn_2SiO_4—0.004Mn	8.5	99	8.5	475	40	green
$2(88ZnO,12BeO)SiO_2$— 0.02Mn	7	97	7	450	31	yellow
$Cd_2B_2O_5$—0.002Mn	5.5	96	5.5	240	13	orange
$Cd_5Cl(PO_4)_3$—0,05Mn	7	95	7	400	28	orange
$Zn_3(PO_4)_2$—0.3Mn	5	100	5	170	8.5	red
$2CaO.Al_2O_3.SiO_2$—0.04Ce	4	90	4			blue

The sulphide phosphors ZnS-Ag-Cl and (ZnCd)S-Ag-Cl are generally employed for direct-vision television tubes; (ZnCd)S-Ag-Cl, (with less Cd than indicated in Table 1, giving greener fluorescence) is also used for the screen of the electron-microscope. Phosphors $CaO.MgO.SiO_2$-Ti and $2(ZnO.BeO)SiO_2$-Mn are used for the screens of projection television tubes. $CaWO_4$ is encountered in X-ray techniques.

Phosphors ZnO-Zn and $2CaO.Al_2O_3.SiO_2$-Ce are remarkable for their short persistence (10^{-6} and 10^{-7} sec respectively); these are used for the televising of films by the scanning method ("flying spot scanner") Zn_2SiO_4-Mn (willemite) is employed in oscilloscope tubes, and this (green luminescing) phosphor in conjunction with $Zn_3(PO_4)_2$-Mn (red) and $CaO.MgO.SiO_2$-Ti (blue) is of importance in certain colour-television systems.

The table also includes a number of other phosphors whose importance will become evident in the next section.

Effect of temperature on efficiency

It has already been seen that the bombardment of the screen by primary electrons results in only a part (η_i) of the energy being converted into radiation. The remainder is ultimately liberated at the screen in the form of heat. In the kind of tube used for large-screen projection television, in which the primary image is very small, the screen gets very hot and air-cooling is essential to prevent the temperature from becoming dangerously high. It is therefore useful to know how the efficiency of the fluorescence varies as a function of temperature.

[5]) This has a value of 660 lumen/watt at the region of maximum visual sensitivity (5500 Å).

The temperature dependence of the fluorescence has been investigated in a number of different phosphors by means of the apparatus depicted in fig. 3. For the blue-fluorescing phosphors the

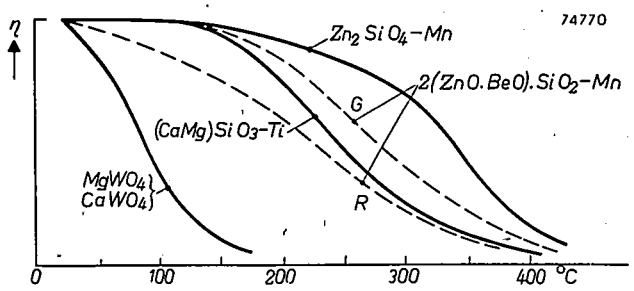


Fig. 4. Effect of temperature on the efficiency η of a number of phosphors. Curve *G* refers to the green, and *R* to the red band of the ZnBe phosphor.

measurements were made with the Cs-Sb vacuum photo-cell; for the phosphors the selenium cell was employed [6]).

As a rule a rapid decline in efficiency occurs above a certain temperature, but in some cases the drop is not so marked, or there may even be an initial increase.

A number of the results are shown graphically in *figs. 4* and *5*, in which the relative efficiency, i.e. $100\ \eta/\eta_{max}$, is plotted against the temperature in °C (η_{max} represents the maximum observed efficiency under prevailing conditions).

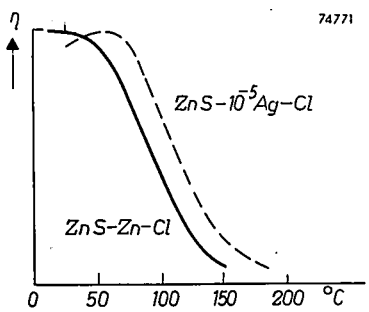


Fig. 5. Effect of temperature on the efficiency η of two sulphide phosphors.

Effect of the energy of the primary electrons on the efficiency

In addition to the above measurements, the variation in efficiency of certain phosphors with the energy of the primary electrons, or acceleration voltage, was determined. The results are reproduced in *fig. 6*, which shows the relative efficiency as a percentage ($100\ \eta/\eta_{max}$) plotted as a function of the acceleration voltage in kV. (Here η_{max} represents the maximum measured efficiency under these conditions).

This maximum efficiency was in most cases attained with a voltage of 25 to 30 kV. For some phosphors the efficiency remains constant above that voltage, whereas with others it drops slightly on further increase of the voltage. There is no immediate explanation as to why this drop takes place; possibly it is due to an increasing influence of the absorption, since at higher voltages the electrons penetrate the layer of crystals to a greater depth — the fluorescent light accordingly has to travel further before being finally emitted. It is

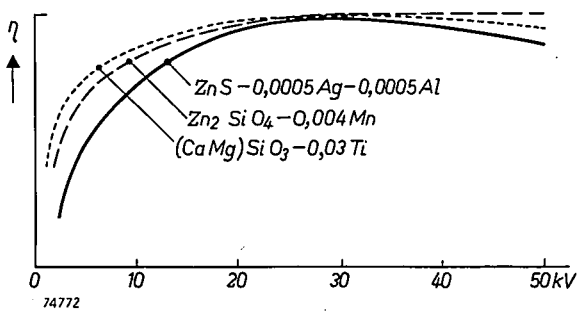


Fig. 6. Effect of the acceleration voltage on the efficiency η of various phosphors.

also possible that, when the primary electrons are accelerated, the chances of elastic collision are increased — at the expense of inelastic collisions — with a corresponding increase in the number of reflected primary electrons.

Comparison of cathodoluminescence with ultra-violet fluorescence

As will be seen from Table I, the efficiency of fluorescence due to electronic bombardment can be quite considerable, viz. up to 25%. This is particularly noteworthy in view of the many different ways in which energy is lost (elastic collisions, transitions without radiation etc.) mentioned in the introductory paragraphs.

Ultra-violet irradiation entails fewer causes of loss, and higher efficiency values are therefore often obtained. The quantum efficiency (chance of emission of a quantum of visible light per ultra-violet quantum absorbed) will quite frequently have values in the neighbourhood of 100%. However, since ultra-violet quanta embody more energy than the emitted light quanta, the energy efficiency, even in the most favourable cases, is naturally lower, being usually not more than 50%.

Table II shows the efficiency of a number of phosphors when bombarded with electrons and also when irradiated with ultra-violet at wavelengths of 2537 Å and 3650 Å, from which it will be seen that, in terms of energy, the fluorescence of sulphides

[6]) Actually, the spectral distribution does depend on the temperature (it usually increases in width when the temperature is increased), but, even allowing for the spectral variation of sensitivity of the selenium cell, no very great error is entailed in the measurement of the energy efficiency as a function of the temperature. At the same time, any such variation in the spectral distribution may have a marked effect on the colour of the fluorescent light.

under electronic bombardment is well able to compete with that produced by U-V irradiation and in some instances is greater.

Table II. Efficiency of the fluorescence of some phosphors under electronic bombardment and ultra-violet irradiation.

Phosphor	Energy efficiency		
	Electrons 20 kV	Radiation 2537 Å	Radiation 3650 Å
ZnS—5×10⁻⁴Ag—5×10⁻⁴Al	23	3	21
(ZnCd)S—Ag—Cl	19	10	35
ZnS—Cu—Cl	18	10	40
Zn_2SiO_4—Mn	8	45	—
$(ZnBe)_2SiO_4$—Mn	7	40	—
$MgWO_4$	2	45	—
$Cd_2B_2O_5$—Mn	5.5	35	—

Light yield from thin phosphor layers

In the foregoing we have shown how the efficiency η_i can be derived from the fluorescence of a very thick layer, with light emitted only from the rear. In practice, however, we are usually concerned with thin layers which are viewed from the front. The total luminous flux $\eta_i I_p = I_0$ produced (measured as energy) then gives rise to a luminous flux T emitted from the front of the layer, and another, R, emitted from the rear. In general:

$$R + T < I_0$$

because some of the radiant energy is lost by absorption in the layer. In *fig. 7* we give the theoretical curves for R, T and $R+T$ as a function of the layer thickness D (see Appendix), assuming that the electron energy I_p is wholly dissipated in the layer. Clearly, with very thin layers, half the light will be emitted from the front of the screen and half from the rear; hence when $D=0$, it may be said that, theoretically, $R=T={}^1/_2 I_0$, or, including a correction for reflection at the glass face, $R={}^1/_2 I_0\,(1+r_G)$, and $T = {}^1/_2 I_0\,(1-r_G)$, where r_G denotes the reflection coefficient of the glass.

It will be seen that the luminous flux T, which is the important element, cannot exceed about 45% of the fluorescent radiation. In actual fact, in the case of very thin layers, the attainable values of T are even lower than those indicated in fig. 7, because thin layers cannot absorb the whole of the electron energy I_p. There are two reasons for this. If the layer thickness is less than the penetration depth of the electrons (about 1μ for 10 kV electrons), some of the primary electrons pass through the crystals without giving up their energy.

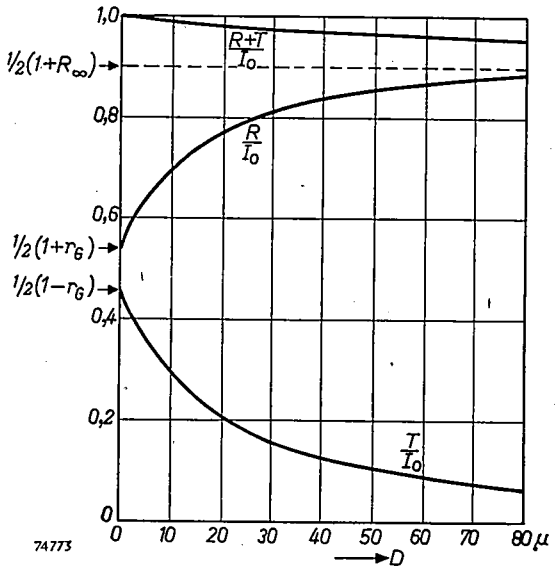


Fig. 7. Theoretical curves showing the ratio of luminous energy to intrinsic energy I_0 as a function of the thickness D. The upper curve refers to the total light output ($R+T$). The other curves are for the light emitted to the rear (R) and to that emitted to the front (T). The curves are computed from formulae (7) and (8) (see appendix; $a = 15$ cm^{-1}, $s = 600$ cm^{-1}, $R = 0.8$, $r_G = 0.08$).

Secondly, owing to the finite size of the crystals, holes occur in the layer if D (or strictly speaking, M, the weight per unit area of fluorescent substance), is very small. The penetration and, more particularly, the incomplete coating due to low values of M result in R and T both approaching zero. Since T, as we have already seen, also approaches zero with very thick layers (scatter and absorption), there must be an optimum layer thickness at which the luminous flux obtained at the front of the screen reaches a maximum. This is clearly demonstrated in *fig. 8*, in which measurements of R and T as

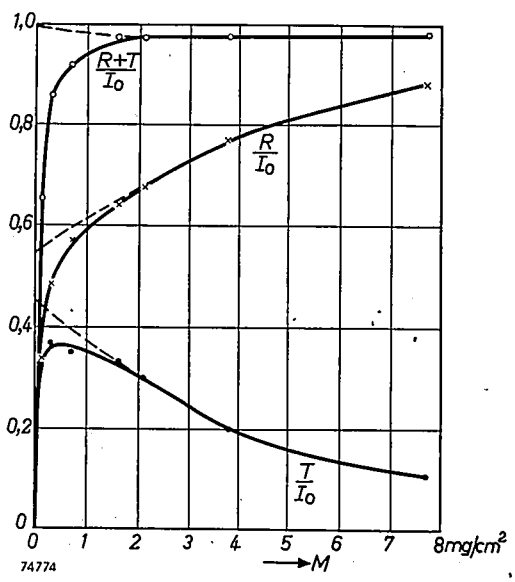


Fig. 8. Radiation from a willemite screen with 10 kV acceleration voltage, from the front (T) and rear (R), and the total ($R+T$), as a function of the coating M. The dotted lines are the theoretical curves shown in fig. 7.

applied to a willemite (Zn_2SiO_4-Mn) screen are shown graphically. With optimum layer thickness, the energy T emitted at the front of the screen is usually 30 to 40% of the theoretical total fluorescent energy I_0. (At the rear we then have 60 to 65%; 5 to 15% is lost through the incomplete coating.)

Very much better results are obtained with a fluorescent screen having a metal backing, i.e. a thin film of metal on top of the phosphor. In some cases the metal backing is used to prevent the build up of charge on the screen (the metallic layer being earthed). It also largely protects the phosphor from bombardment by negative ions (ion burn), and it has yet another great advantage in that the light emitted towards the rear (R) is to a great extent reflected towards the front, thus adding to the effective luminous flux T.

The coating is applied in the following manner. A coating of some organic substance such as nitrocellulose is first produced on the layer of crystals by evaporation from a solution; this serves to fill up the interstices and irregularities, and thus to present a smooth surface. Aluminium is then evaporated on to this prepared surface, the organic filling layer being subsequently removed by heating, leaving the aluminium as a mirror-bright layer resting on the projecting corners of the crystals (*fig. 9*). The

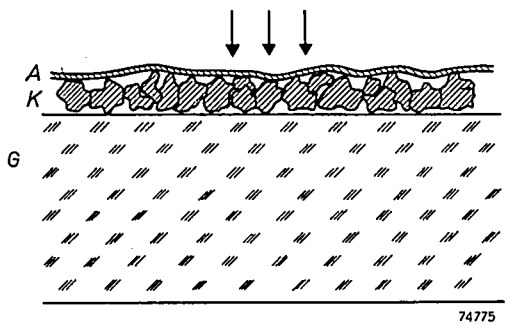


Fig. 9. Glass plate G with phosphor crystals K and aluminium backing A after removal of the nitro-cellulose layer. The arrows represent electrons.

primary electrons now have to penetrate this layer, which is opaque to light, before they can reach the screen proper [7]). It is true that this involves a certain loss of energy (a few kV per micron of aluminium), but the gain in light due to reflection from the aluminium completely outweighs this; the reflection coefficient of the aluminium is very high (in the region of 80%; see Appendix).

A certain minimum thickness of aluminium is necessary before the reflection coefficient of the layer reaches this high value. On the other hand the layer may not be too thick for it would then absorb too much of the electron energy. This layer too, therefore has an optimum thickness. *Fig. 10*

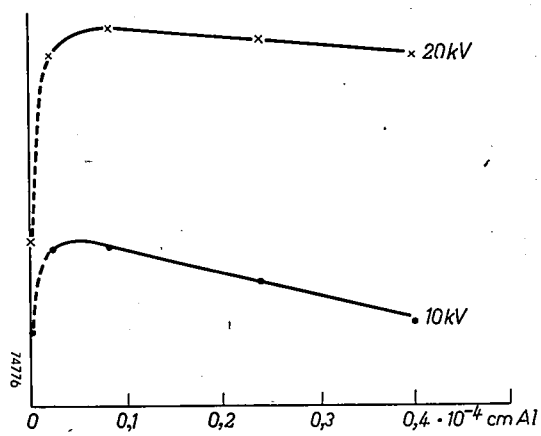


Fig. 10. Radiation (on relative scale) from the front of a willemite screen backed with aluminium, plotted against the thickness of the metal layer for 10 kV and 20 kV acceleration voltages.

shows the quantity of light emitted (on a relative scale) from a thin willemite screen, plotted against the thickness of the aluminium layer; it will be noted that a maximum is reached with quite a thin layer of aluminium (some hundredths of a micron), this being followed by a gradual reduction as the thickness is increased. This drop is a result of the loss of energy of the primary electrons. Extrapolating to zero thickness, it is found that a loss of 10% occurs at 20 kV and 50% at 10 kV, both for 0.4 μ aluminium [8]). The optimum layer thickness is 0.02 — 0.08 μ for 10 kV, and about 0.1 μ for 20 kV.

A maximum in the light output of aluminized screens is also found with increasing thickness of the phosphor layer. This maximum would not occur if the reflection factor of the aluminium layer were 100% and if no light were lost through absorption in the phosphor layer: after being reflected once or several times by the aluminium all the light would finally be emitted from the front of the screen. The light output would thus increase with increasing thickness of the phosphor layer until the screen has become so thick that all the electron energy is being utilized, and would then remain constant. The reflection factor of the aluminium layer is, however, lower than 100%. This means that every time light falls on it some of it is absorbed. The thicker the phosphor layer i.e. the greater the value of R (see fig. 1), the more often the light is thrown back by the phosphor

[7]) See J. de Gier, A projection television receiver, Philips tech. Rev. **10**, 97-104, 1948, in particular pp. 103-104.

[8]) With 50 kV and the same thickness, the loss in energy is only 1 to 2%. Cf. F. Kirchner. Ann. Physik, **30**, 683-688, 1937.

towards the aluminium layer and the higher the absorption losses become. The light output must approach zero at infinite thickness of the phosphor layer even if no losses occur through absorption in the phosphor layer itself. All the light is then lost in the reflection at the aluminium layer.

If we compare the optimum values of T' (with aluminium) with T (without aluminium), we see that the gain amounts to a factor of 2 to 2.2.

In order to determine the light yield to be expected under the most advantageous practical circumstances it is necessary also to take into consideration the fact that, under manufacturing conditions, the efficiency of a phosphor is usually some 10% below its original value. Furthermore, the conditions under which tubes with sulphide screens are normally operated generally entail a certain amount of saturation, which results in a further loss of 10 to 15%.

The intrinsic luminous efficiency of a white-fluorescing sulphide mixture is roughly 60 lm/W, so that, in view of the above-mentioned losses, a mixture of this type can be expected to yield at most 15 lm/W at the front of the screen without metal backing, that is, a specific luminous intensity of about 5 cd/W, or 30 lm/W ($\approx$10 cd/W) with the backing.

The acceleration voltage necessary to produce these results must be at least 10 kV.

Appendix. Theoretical aspects of scatter and absorption

The scattering and absorption of light that take place in the screen can be formulated theoretically following Hamaker, Klasens, Coltman and others [9]), on the basis of a theory given by Schuster (1905).

This theory postulates that the characteristics of the screen are governed by two constants s and a, the first of which refers to the scattering and the second to the spectral absorption of the fluorescent light. It is assumed that the spectral composition of the fluorescent light is not modified by the absorption, so that s and a may be regarded as constant in a given instance. Furthermore, the phosphor layer is taken to be homogeneous, i.e. the grain size may be disregarded in comparison with the thickness of the screen.

According to Schuster, light scattering and absorption in a medium can be represented by the following differential equations:

$$\frac{dJ_1}{dx} = -(a+s)J_1 + sJ_2,$$

$$\frac{dJ_2}{dx} = (a+s)J_2 - sJ_1,$$

where J_1 and J_2 are the luminous flux values at a point x, in opposite directions perpendicular to the screen. At the rear of the screen ($x = 0$), J_2 is equal to R, and at the front ($x = D$), $J_1 = T$.

In the case of the scattering of fluorescent light R and T are evaluated by solving analogous differential equations, with the introduction of suitable boundary conditions. Using the abbreviated form $\sigma = [a(a+2s)]^{1/2}$ the relevant solutions for R and T are:

$$T = \tfrac{1}{2} I_0 \frac{1-r_G}{\frac{a+(1-r_G)s}{\sigma}\sinh \sigma D + \cosh \sigma D} \quad . . \quad (7)$$

$$R = \tfrac{1}{2} I_0 \frac{(1+r_G)\cosh \sigma D + \frac{(1-r_G)(a+2s)}{\sigma}\sinh \sigma D}{\frac{a+(1-r_G)s}{\sigma}\sinh \sigma D + \cosh \sigma D} \quad . \quad (8)$$

Further, it follows from the theory that:

$$\frac{s}{a} = \frac{2R_\infty}{(1-R_\infty)^2}, \quad . . . . . \quad (9)$$

which, in conjunction with (7) and (8) enable us to write:

$$\frac{R}{T} = \frac{1+r_G}{1-r_G}\cosh \sigma D + \frac{1+R_\infty}{1-R_\infty}\sinh \sigma D. \quad . \quad (10)$$

Provided that σ D is sufficiently small (e.g. σ D < 0.3), it may be said as an approximation that:

$$\frac{R}{T} = \frac{1+r_G}{1-r_G} + (a+2s)D. \quad . . . . \quad (11$$

The form of T and R are such that when $D \to 0$, $T = \frac{1}{2}I_0(1-r_G)$, and $R = \frac{1}{2}I_0(1+r_G)$; also $R + T = I_0$, as anticipated. If D be increased, T drops steadily and R rises. When $a = 0$, $R + T$ remains equal to I_0, since no energy is lost through absorption. When $a \neq 0$, while D increases, $R + T$ and also R alone, approach the value $\frac{1}{2}I_0(1+R_\infty)$, which agrees with the efficiency, given above, of $\eta_m = {}^1/_2 (1+R_\infty)\eta_i$.

The functions (7) and (8) correspond closely to the experimental curves for $D \to \infty$, but this is no longer the case as $D \to 0$, because the electron energy is then not wholly absorbed by the phosphor layer.

The quantities s and a, which characterize the optical performance of the screen, are not usually known, so for a number of phosphors measurements have been taken of R, T, R_∞, r_G and D, from which a, s and σ have been computed with the aid of formulae (9) and (10), or (11); the results for the willemite screen (fig. 8) mentioned in the foregoing are now given in *Table III*.

Table III. Measurements on willemite (Zn_2SiO_4-Mn).

M mg/cm^2	σ cm^{-1}	s cm^{-1}	a cm^{-1}
2	45	900	1
4	55	1100	1.5
6	65	1300	1.5
7.5	80	1650	2

It is thus seen that the values of σ and s are not constant, but depend on the layer thickness; the layers having $M = 2$ to 7 mg/cm^2 are in fact too thin for the theoretical formulae to be valid, while for the thicker layers measurements of T cannot be relied upon.

In principle it is also possible to compute the quantity of

[9]) H. C. Hamaker, Philips Res. Rep. **2**, 55-67, 1947.
H. A. Klasens, Philips Res. Rep. **2**, 68-78, 1947.
J. W. Coltman, E. G. Ebbighausen and W. Altar, J. appl. Phys. **18**, 530-544, 1947.
For further bibliography see note [1]).

light using Hamaker's formulae in the case of screens with metal backing (reflection factor r_A); it is found that:

$$T' = \tfrac{1}{2} I_0' \frac{(1-r_G)(1+r_A)}{(1+r_G r_A)a + (1-r_G)(1-r_A)s} \sigma \sinh \sigma D + (1-r_G r_A)\cosh \sigma D \quad \ldots\ldots (12)$$

where $I_0' = \eta_i I_p'$; $I_p'(<I_p)$ represents the energy of the primary electrons after they have passed through the aluminium layer. I_0' is therefore slightly less than I_0. In many cases the quantity r_G (roughly 8%) may be disregarded. Assuming that $\sigma D < 0.3$, our formula now becomes:

$$T' = \tfrac{1}{2} I_0' \frac{1+r_A}{1+[a+(1-r_A)s]D} \quad \ldots\ldots (13)$$

Hence, if $r_A = 1$ and $a = 0$, we should find that $T' = I_0'$, in which case, as anticipated, all the light produced by the electrons would be emitted at the front, no matter how thick the screen were made.

When r_A differs from unity, the expected improvement can be computed with formula (12) or (13), the values of a, s and σ being taken from the actual measurements given above; conversely, r_A can be computed from measurements carried out on screens with and without aluminium backing.

With a given aluminium backing (0.3 μ thick) the relative efficiency of willemite was measured as a function of the layer thickness in mg/cm² of this phosphor; from the curve obtained for a voltage of 20 kV the magnitude of r_A was then determined with the aid of formula (13) in the form:

$$\frac{T_1'}{T_2'} = \frac{1+[a+(1-r_A)s]D_2}{1+[a+(1-r_A)s]D_1}, \quad \ldots\ldots (14)$$

where D_1 and D_2 represent two values of the layer thickness, and T_1'/T_2' the ratio of the luminous emission values in respect of this thickness. For values of D_1 and D_2 corresponding to layers of 1.5 and 7.5 mg/cm² it was found that $T_1'/T_2' = 1.5$, so that $r_A = 0.83$. Direct measurement carried out on the aluminium layer after the tube had been opened, gave $r_A = 0.82$. In other cases the amount of agreement was not so close, but on the basis of a number of measurements taken, it may be said that r_A is roughly equal to 80%.

Summary. Introductory remarks on the mechanism of cathodoluminescence and the scattering and absorption of light in the screens of cathode-ray tubes are followed by a description of methods of measurement, the purpose of which is to determine the intrinsic conversion efficiency of electron energy into luminous energy. The results of measurements taken from a number of different phosphors with varying temperature and primary electron energy are presented. Further measurements give the luminous intensities at the front of conventional types of screen with and without aluminium backing. There are optimum values of the screen and backing thicknesses. Theoretical support is also lent to the arguments.

ABSTRACTS OF RECENT SCIENTIFIC PUBLICATIONS OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN

Reprints of these papers not marked with an asterisk * can be obtained free of charge upon application to the Administration of the Research Laboratory, Kastanjelaan, Eindhoven, Netherlands.

2021: K. F. Niessen: Curie temperature of nickel-zinc ferrites as a function of the nickel-zinc ratio (Physica **17**, 1033-1049, 1951).

In a series of mixed crystals of nickel ferrite (or any other ferromagnetic ferrite $Me^{2+}Fe_2O_4$) and zinc ferrite, the ratio of the Curie temperature to that of the zinc-free ferrite can be calculated by means of a generalisation of Néel's extension of the Weiss theory to ferrites. In this generalisation the interactions of all magnetic ions are taken into account. The dependence of the Curie temperature on composition leads to the prediction that zinc ferrite is antiferromagnetic.

2022*: H. G. Beljers, W. J. van de Lindt and J. J. Went: A new point of view on magnetic losses in anisotropic bars of ferrite at ultra high frequencies (J. appl. Phys. **22**, 1506, 1951).

When ferrite rods are manufactured in a special way the magnetic losses at frequencies between 100 Mc/s and 10^4 Mc/s can be greatly reduced (in some cases down to 10 percent) by a relatively small, axially directed, magnetic field. Ferrite bars not specially manufactured show losses nearly independent of polarization. The effect may be explained by assuming that extrusion causes elongated air inclusions giving rise to an anisotropic structure, and a certain non-random orientation of the magnetic domains.

2023: A. M. Kruithof and W. de Groot: Photometry at very low brilliance levels (C.I.E. Proc., Stockholm 1951, Vol. 2,).

The paper describes preliminary experiments at luminances between 10 and 10^{-5} cd/m² with a photometer with circular field (central area 4°, annular area 8°) by four observers. A source, having a colour temperature of 2042 °K, is compared with sources of colour temperatures 2042 °K and 7700 °K resp. The heterochromous comparisons show a dependence on the fixation point (foveal and 4° parafoveal), which is apparent when the value of $\log (L_{2042}/L_{7700})$, corresponding to a match in apparent brightness, is plotted versus $\log L_{2042}$ (L = luminance).

VOL. 15 No. 3, pp. 73-104 SEPTEMBER 1953

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
THE PHILIPS INDUSTRIES

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, NETHERLANDS

THE FERROXCUBE TRANSFORMER CORE USED IN THE BROOKHAVEN COSMOTRON

by F. G. BROCKMAN *) and M. W. LOUWERSE **). 621.318.134:621.384.61

For the last ten or twenty years fundamental research in nuclear physics has been associated with the building of ever larger machines for particle acceleration. This evolution which has tied up the physicist's work with that of the engineer has frequently been commented on. However, the evolution continues, and in recent years the personal assistance of the engineer has been superseded by the material cooperation of whole industrial organizations. This article will give an account of the cooperative studies that have been made in the development of ferrites for high energy particle accelerators by the Brookhaven National Laboratories, U.S.A., and the Philips Laboratories in the Netherlands and the U.S.A.

Introduction

In the efforts to produce particles of very high energies for nuclear experiments, a substantial advance was made in the past year when the proton accelerator of the Brookhaven National Laboratories, U.S.A. — the Cosmotron — was brought into successful operation. While the highest particle energy obtained with earlier accelerators did not surpass 1000 million electron volts, the Cosmotron has produced protons with 2300 MeV, and it is expected that the energy can be increased to 3000 MeV.

The completion of this machine, which was achieved by a group of workers including M. S. Livingston, J. P. Blewett, G. K. Green and L. J. Haworth, is a remarkable feat of enterprise and engineering. A few data that will be mentioned presently may give some idea of the vastness of the project. An account of the design was given by these Brookhaven workers in 1950 in the Review of Scientific Instruments, and a detailed description of the completed machine has just appeared in the same journal [1]).

The present article will focus attention on the fact that, in one of the chief elements of the Cosmotron, the "transformer" producing the radio-frequency electric field for the actual acceleration of the protons, a large ferrite core plays an important part. The major portion of this core was assembled from Ferroxcube material manufactured by Philips at Eindhoven. The specifications for this core and the problems encountered in selecting the material and in the manufacturing will be described. A few remarks will be made on the development of other ferrite materials for similar purposes.

Basic design of the Cosmotron

A short description of the design principles of the Cosmotron as published in the first article quoted in [1]), will first be given.

The protons travel in an annular evacuated chamber of cross-section $30'' \times 7''$. This chamber is composed of four quadrants of about 30 feet radius, which are connected together by four short straight sections (plan view, *fig. 1*). The protons are made to follow a path roughly on the axis of the chamber by magnetic fields perpendicular to the plane of the path, which are produced by four powerful electromagnets disposed along the path. In the photograph of the Cosmotron (*fig. 2*) the yokes of these magnets, which enclose the annular acceleration chamber, can be seen.

*) Philips Laboratories, Inc., Irvington-on-Hudson, N.Y., U.S.A.
**) Laboratory of the Philips Ceramics Factory, Eindhoven, Netherlands.

[1]) M. S. Livingston, J. P. Blewett, G. K. Green and L. J. Haworth, Design study for a three-Bev proton accelerator, Rev. sci. Instr. **21**, 7-22, 1950.
The detailed description is contained in a series of articles in Rev. sci. Instr. **24**, 723-870, 1953 (No. 9).

One of the straight field-free sections is used for the injection of the protons into the chamber. The protons are obtained from a hydrogen ion source and accelerated by a Van de Graaff high voltage generator to an initial energy of 3.5 MeV for injection. In another straight section the acceleration of the injected protons is effected: under the influence of a longitudinal electric field alternating at a suitable frequency, their energy is increased by a certain amount after every revolution. To keep the protons (mass M, charge e) on a circular path of radius R, a magnetic field of flux density

$$B = \frac{Mv}{eR} \quad \ldots\ldots\ldots \quad (1)$$

is required. With the momentum Mv corresponding to the *initial* energy of 3.5 MeV, the required value of B is 0.029 Wb/m². In order to keep the protons on the same orbit within the vacuum chamber

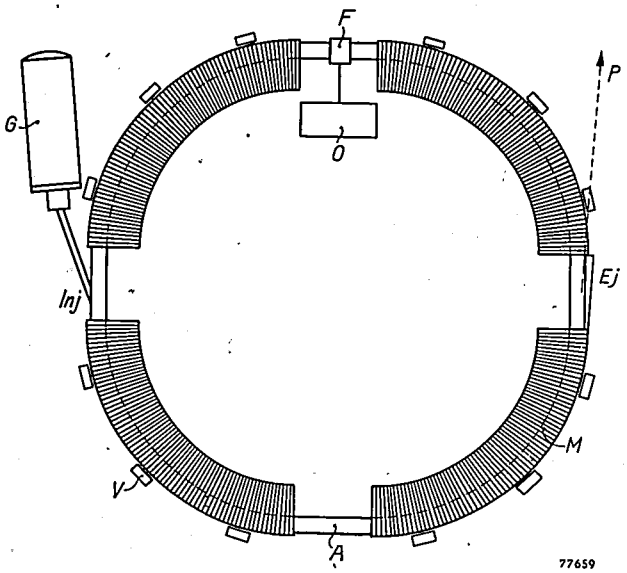


Fig. 1. Plan-view of the Cosmotron. The accelerated protons move on (or near) the axis of the annular evacuated chamber A, which has a diameter of about 60 feet. The protons are injected into the chamber at *Inj*, with an initial energy of 3.5 MeV, from a Van de Graaff high tension generator G. Four electromagnets M keep the protons on their track. The acceleration of the protons occurs at F, where an alternating electric field, generated by the oscillator O, gives the protons a "kick", once per revolution. When the protons have acquired their final energy, they emerge in a beam P from the annular chamber at the ejector Ej. V represent a number of high vacuum pumps for the evacuation of the acceleration chamber.

while their energy and momentum is increasing, the magnetic field is increased continuously, up to a maximum value of 1.4 Wb/m². This value will allow the protons to be kept on their track with a final energy of 3100 MeV. The beam of accelerated protons can be ejected from the annular chamber by a suitable deflector device accommodated in another straight section of the chamber.

At the injection energy of 3.5 MeV, the proton velocity is 2.6×10^7 m/sec. As the path along the axis of the annular chamber has a total length of about 70 meters, the time for one revolution is initially about $^1/_{350\,000}$ second. This means that the accelerator section proper must carry an alternating electric field of frequency about 350 000 cycles/second, in order to give the circulating protons a "kick" each revolution in the correct phase. At a final energy of 3000 MeV, the proton velocity (for the calculation of which the relativistic formula must now be applied [2]) is approximately 10 times the initial value, so that during the course of the acceleration of a burst of protons the frequency of the accelerating electric field must be gradually increased from 0.35 to about 4 Mc/s.

Although this brief description would suffice for our further discussion of the accelerator, a few further details of the design study published by the Brookhaven group [1]) may be mentioned in this section for the sake of completeness. One of the major problems of the project was the supply of the large magnetic energy required for keeping the protons on their track. The field of 1.4 Wb/m² in a space 70 metres long and roughly 36″×9.35″ in cross-section (gross dimensions, as against the above-mentioned net dimensions inside the chamber) involves a magnetic energy of 12 megajoules. This formidable magnetic energy [3]) must be built up during the acceleration time of one burst of protons and taken away relatively quickly after the ejection in order to make the apparatus ready for the next burst. As it would be very uneconomical to dissipate the magnetic energy after each burst, it is stored mechanically in the form of the kinetic energy of a large flywheel and recycled. In order that the build-up of the magnetic energy may be accomplished with a reasonable peak power, and that the acceleration may be achieved with not-too-large "kicks", the acceleration time for one bunch of protons must be relatively long. An acceleration time of 1.0 sec was chosen. (This means that the protons will make roughly three million revolutions during their acceleration.) The cooling of the magnet and the available power of the motor-

[2]) A convenient graph for the general velocity-energy relationship may be found in Philips tech. Rev. **11**, p. 68, 1949.

[3]) For a comparison, the reader is referred to the description of the Amsterdam synchrocyclotron given in this Review **12**, 354, 1950: it was there mentioned that the large electromagnet contains a magnetic energy of 0.65 megajoules, whereas the magnetic energy in the gap of a high power loudspeaker amounts to about 0.25 joules.

generator coupled to the flywheel are such that the acceleration of a bunch of protons may be repeated every 5 seconds.

Another important problem was the simultaneous variation of the magnetic field and of the frequency of the accelerating electric field. Because of the essentially non-linear behaviour of an electromagnet as a circuit element, it would be difficult to realize an exactly predetermined increase of the magnetic field with time. The designers of the Cosmotron avoided this difficulty by regarding the magnetic field as the independent variable. The expected time-variation of the magnetic field (*fig. 3*) was derived from the circuit constants and model measurements. From the rate of change dB/dt of the magnetic field it can be deduced by how much the momentum (and the energy) of the protons must increase after each revolution in order to keep them on the same orbit. This required energy gain per revolution is plotted against time in *fig. 4*; it is seen to decrease during the acceleration period from an initial value somewhat greater than 1000 eV to a final value of rather more than 600 eV per revolution. The velocity of the protons resulting at each instant of the acceleration time (and which depends only on the instantaneous value of B) determines the frequency that is required at that instant for the accelerating electric field; cf. curve f in fig. 3. The problem of controlling the frequency to obtain this time variation with the required precision, was solved at Brookhaven in an elegant manner: the derivative of the magnetic field is measured continuously by means of a pick-up coil, and the integrated signal of this

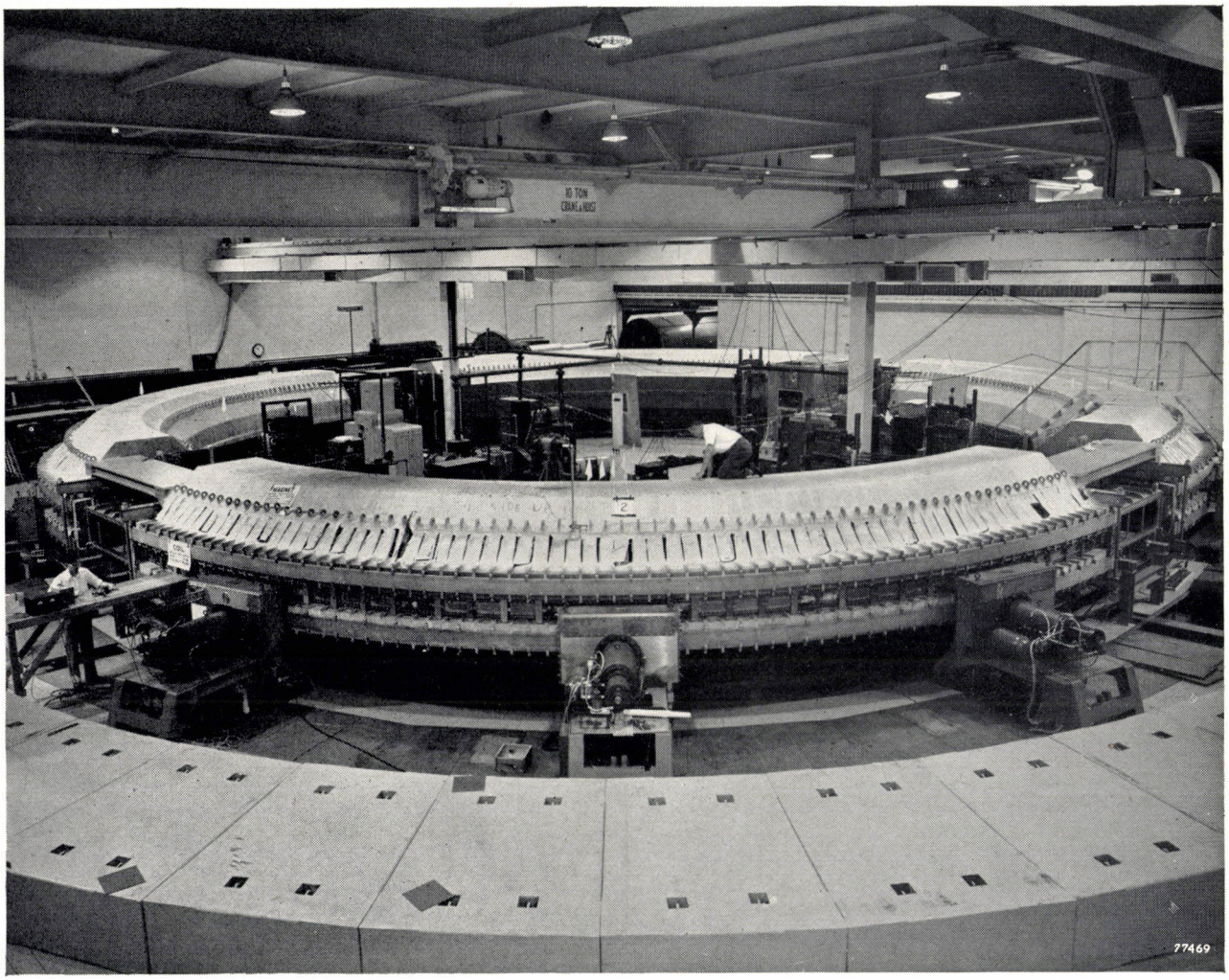


Fig. 2. General view of the Cosmotron, in the Brookhaven National Laboratory, U.S.A. The four large electromagnet quadrants, surrounding the four sections of the annular acceleration chamber, have a total weight of 2200 tons. In the background is the Van de Graaff generator placed in a horizontal tank. Farther to the right is the accelerator unit. The concrete wall seen in the foreground protects personnel from the radiation of the machine.

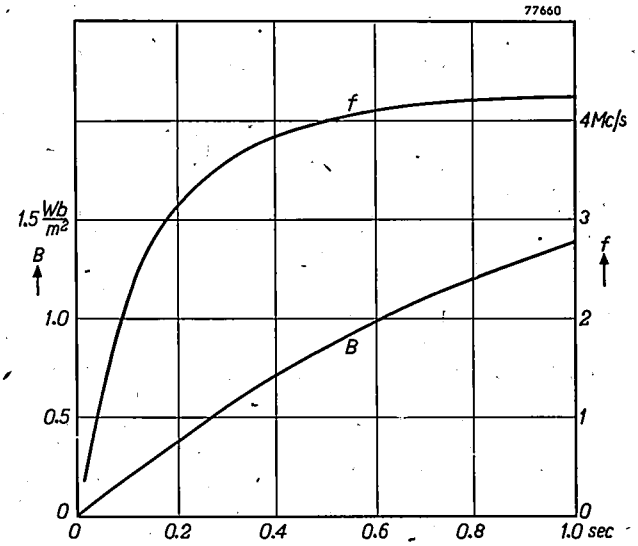


Fig. 3. Variation with time of the flux density B of the magnetic field during one accelerating period. This field, which forces the protons to move in their circular orbit, is regarded as the independent variable: the frequency of the accelerating electric field must then vary according to the curve f, in order to ensure that the protons circulating at an increasing rate will remain on their track inside the annular chamber.

coil, is used for controlling the frequency of a "permeability-tuned" LC-oscillator producing the accelerating radio frequency field [4]). The tuned circuit of this oscillator employs an inductance containing a Ferroxcube III core with an initial permeability of about 1000. The control signal provides an increasing bias magnetization of the core, causing its effective permeability to drop sharply, thus decreasing the inductance. In this way the required wide frequency sweep (fig. 3) is obtained, together with the required frequency stability. A frequency error as small as 0.3% would cause the proton orbit radius to exceed the latitude allowed by the annular chamber. With the oscillator described, when the temperature of the Ferroxcube core is held constant to 0.1 °C, the oscillator frequency is stable and reproducible at any control signal value to better than 0.1%. The oscillation amplitude is held constant at a value such that the peak accelerating field is about twice the highest instantaneous value needed (see above).

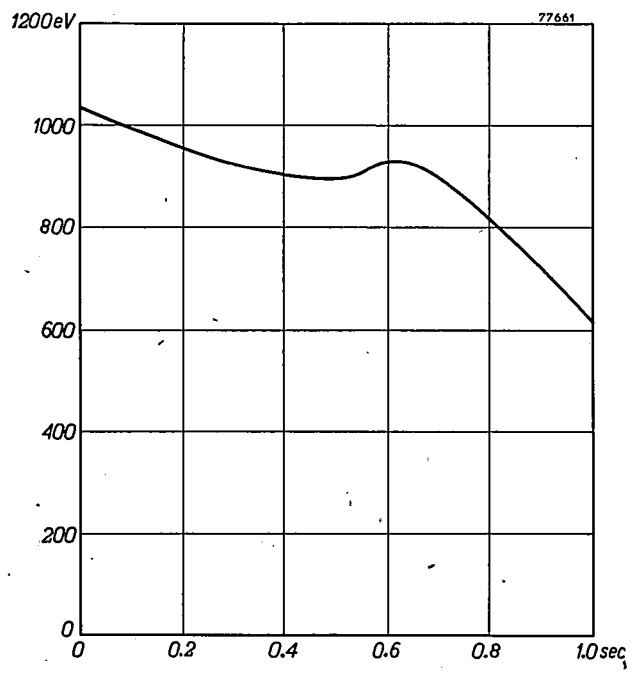


Fig. 4. Required energy gain per revolution of the protons as a function of time.

[4]) A. I. Pressman and J. P. Blewett, A 300 to 4000 kilocycle electrically tuned oscillator, Proc. Inst. Rad. Eng. **39**, 74-77, 1951.

With a view to the phase stability of the revolving protons, these must pass through the alternating electric field at a certain phase angle Θ after its peak. If the protons are focussed to a bunch at a phase of about $\Theta=60°$, the peak value of the accelerating radio-frequency voltage must be about 2200 V. Matters of orbit stability, which of course were of prime importance in the design, need not be considered here, though in the last section they will have to be mentioned.

A Ferroxcube core is thus encountered in the tuning circuit of the oscillator, at the heart of the Cosmotron. This small core, however, did not offer any specific manufacturing problems and need not be discussed in this article. The Ferroxcube core forming the subject of this article is also part of the radio-frequency accelerating device, but serves the purpose of *transferring* the radio-frequency energy to the proton beam.

The accelerating unit

The manner in which the longitudinal radio-frequency electric field is established in the acceleration path of the protons is one of the unique features of the Cosmotron as compared with other types of particle accelerators. The method [5]) may be roughly described by regarding the large circular path of the protons as the single turn secondary winding of a transformer, whose primary winding consists of a coil (having one or more turns) in the anode circuit of a power amplifier (*fig. 5*) driven by the radio-frequency oscillator. Both windings are effectively coupled by a ferromagnetic core (*fig. 6*).

The Brookhaven workers were able to consider the adoption of this novel system because, at the time that the design work on the Cosmotron had begun, the development at the Philips Laboratories

[5]) A proposal to accelerate particles by a method along these lines was already made by A. Bouwers, in his book: Elektrische Höchstspannungen, J. Springer, Berlin 1939, p. 83.

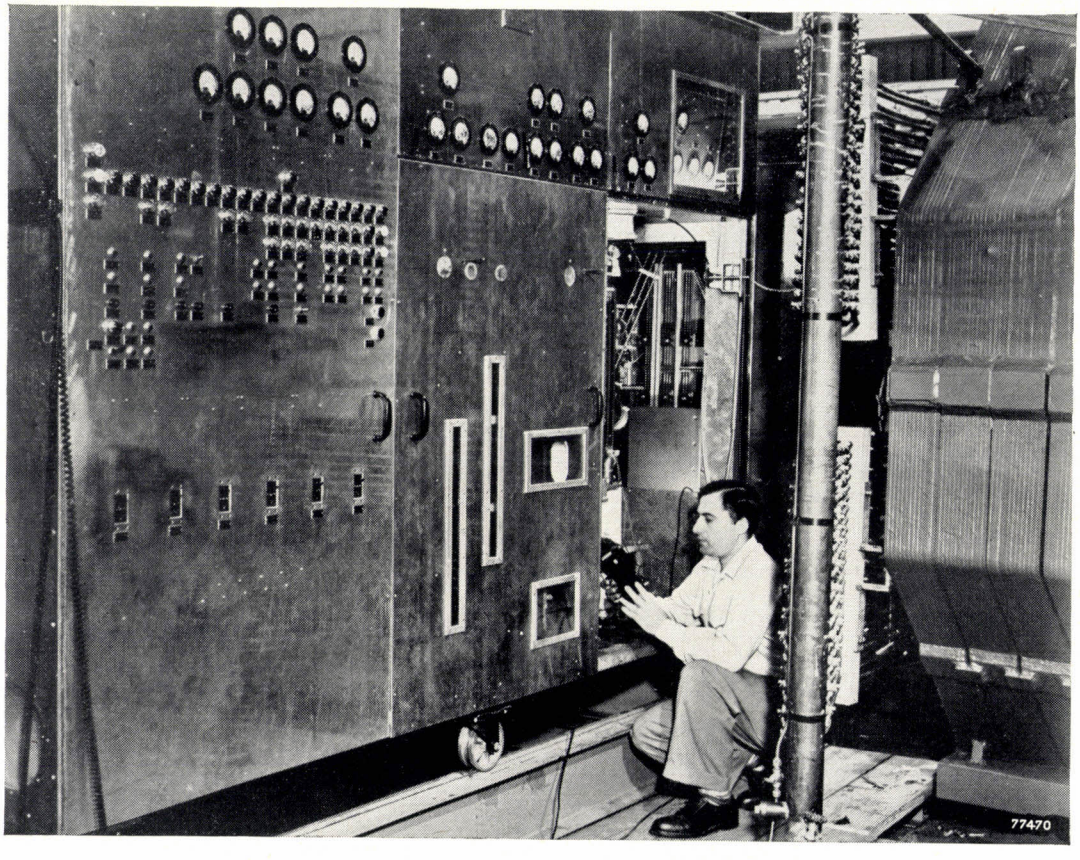


Fig. 5. Power amplifier supplying the radio-frequency energy to the primary of the transformer.

at Eindhoven of ferromagnetic materials suitable for very high frequencies had become known [6]). The usual transformer steel, even with extremely fine lamination, would not do in this case because of excessive eddy current losses. Powdered cores of iron or of "Permalloy" also offered no satisfactory prospects (see below). The new materials, however, owing to their very high resistivity, could be expected to be adequate even with sections several centimeters in thickness.

Quantitatively, the conditions imposed on the core material were derived from the formula for the input impedance Z of the above-mentioned transformer (cf. the first article quoted in [1]), eq. 17). This formula in simplified form says that Z is proportional to

$$f\mu_r\left(\frac{1}{Q}+\mathrm{j}\right),$$

f being the frequency, μ_r the relative permeability of the core material and Q the usual figure of merit of the primary winding including the effects of core losses. In order to obtain the desired accelerating voltage (2200 V peak) with a reasonable anode current of the oscillator, a high impedance Z is desired, while for reducing power loss and heat dissipation in the transformer, the real component of Z should, of

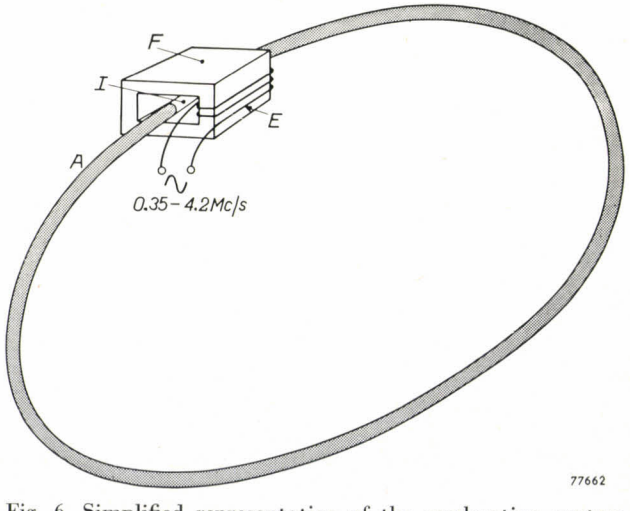


Fig. 6. Simplified representation of the acceleration system of the Cosmotron. The metallic annular acceleration chamber A constitutes the secondary winding of a transformer with primary winding E and ferromagnetic core F. I is an insulating section of the annular chamber, across which the accelerating voltage is produced. — In reality the core F is enclosed in a copper shield having the gap I in the centre. The shield forms an autotransformer with the windings. Thus the electric field is restricted to the gap and radio interference prevented.

[6]) J. L. Snoek, Non-metallic magnetic materials for high frequencies, Philips tech. Rev. **8**, 353-360, 1946; J. L. Snoek, New developments in ferromagnetic materials, Elsevier Publishing Co., Amsterdam 1947.

course, not be too large. A core material with a high permeability μ_r, preferably not less than 500, and low losses resulting in a reasonably high value of Q, say, not less than 0.5 even at the highest frequency used (4.2 Mc/s), was therefore desirable. Comparative measurements of μ_r and Q for ring core samples of several materials were performed by the Brookhaven group. Their results [1]), shown in *table I*, clearly pointed to the superiority of the ferrite material for this application, though they did not yet conclusively demonstrate that ferrites could do the job.

It should be emphasised that the system outlined in fig. 6 was not the only practicable solution for the problem of transferring energy to the proton beam from a radio frequency source. An alternative solution considered by the Brookhaven group involved the use of a *tuned* system, for example a resonant cavity, driven at a varying frequency and continuously tuned over the desired frequency sweep. Although the first system was chosen for the Cosmotron, the alternative system has gained new interest, as it will probably be used in a new type of accelerator which can yield even higher particle energies and which is now the subject of study by a number of groups including that at Brookhaven. We shall return to this further development at the end of this article.

Table I. Properties of magnetic core materials at 1 megacycle, as measured by the Brookhaven group [1]).

Material	μ_r	Q
4-79 Mo Permalloy (0.001″ laminations)	600	0.6
Fe powder (80 mesh grade B)	35	2.0
81 Permalloy powder (120 mesh)	75	1.6
2-81 Mo Permalloy powder (120 mesh)	125	3.0
Typical ferrite	1000	6.0

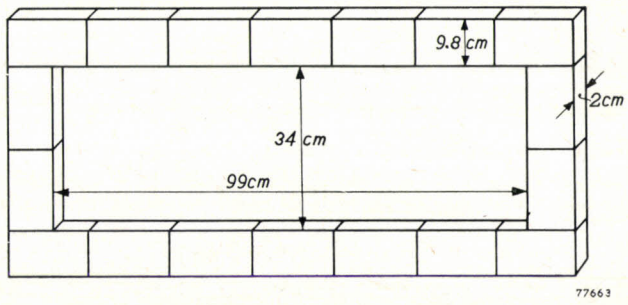


Fig. 7. Dimensions of one of the 24 "picture-frames", made of bricks of Ferroxcube IV, manufactured by Philips for the core of the Cosmotron.

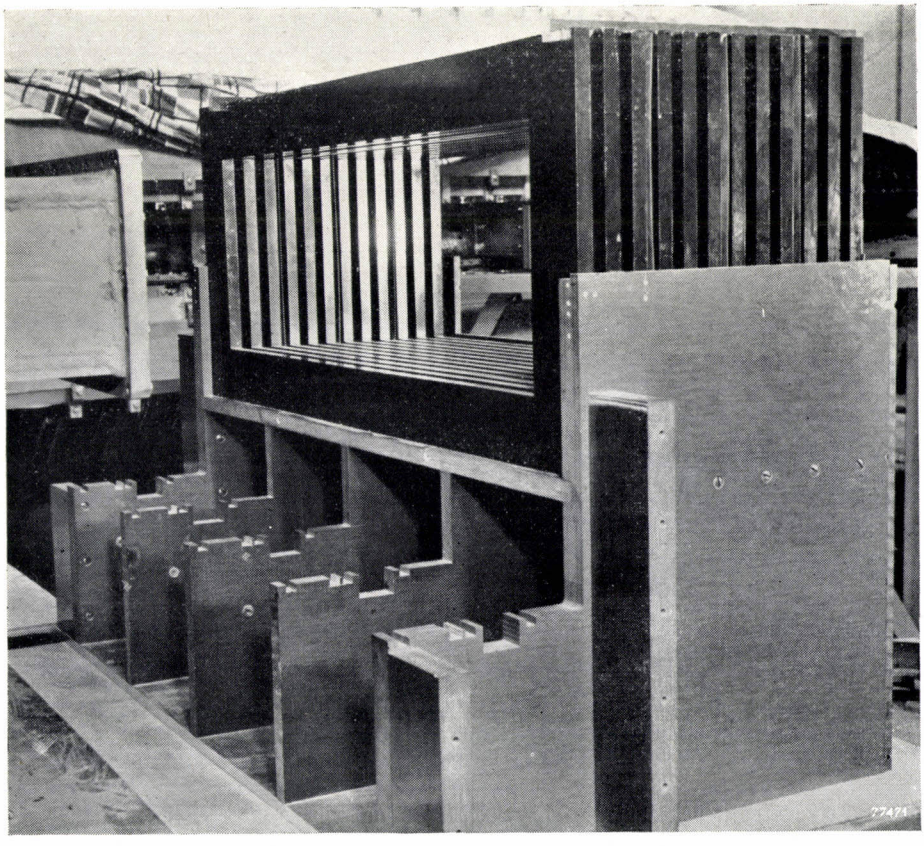


Fig. 8. Bank of 12 Ferroxcube "picture-frames", during the assembly of the core of the Cosmotron. On the left, part of the annular acceleration chamber, which runs through the "window" of the core, can be seen.

The adoption of the transformer system presented the makers of ferrites with two practical problems: the choice of the ferrite material best suited for the purpose, and the actual manufacturing of the large core structure. The "window" of the core must be about 100 cm wide and 34 cm high, to allow space for the annular vacuum chamber and the primary winding. In order to keep the magnetic flux density sufficiently low, a core cross-section of 10×50 cm was required. This would mean a total weight of the ferrite core of about 750 kg.

In 1948 the Brookhaven group consulted the Philips Laboratories, Irvington, regarding a suitable fer-

rite for the core structure. After the characteristics of different materials for the purpose were investigated, part of the final core was assigned to Philips in November 1949. This part consisted of 24 "picture frames" of Ferroxcube IV, of the size shown in *fig. 7* and composed of bricks with dimensions 170.0 × 98.0 × 20.0 mm. The delivery of the whole batch was completed in May 1950. The frames were installed in the Cosmotron in two banks of 12 each. The photograph in *fig. 8*, made during the course of the assembly of the Cosmotron, shows one of these banks.

A few details of the preliminary investigations and of the manufacturing method will now be given.

Choice of material and manufacturing of the Ferroxcube frames

The ferrite chiefly manufactured at the time was Ferroxcube III, a solid solution of manganese-, zinc- and ferrous ferrites. This type of ferrite has by far the lowest hysteresis losses, a property of great importance for telephone applications, and in fact, the use of ferrite in carrier telephony was then commercially the most important [7]). The different varieties of Ferroxcube III have initial permeabilities [8]) varying from 800 to 1700, resistivities of the order of 100 ohm cm (a factor of 10 million times greater than that of iron) and they yield Q-values of 10 at frequencies up to 0.3 to 0.8 Mc/s. From the latter property, it was estimated that core sections with a thickness of a few centimeters would not introduce excessive eddy-current losses. This conclusion appeared to be verified by tests performed on samples of the material of the small size usual in carrier telephony and other applications. When, however, bricks of this material having dimensions of several cm and suitable for the assembling of a test core for the Cosmotron were made at Eindhoven and tested at Brookhaven, it was discovered that their initial permeability μ_i decreased abnormally rapidly with increasing frequency resulting in a very low value at about 2 Mc/s (*fig. 9*). Closer investigation of this unexpected phenomenon at Irvington revealed the existence of a *dimensional effect*, which could be shown to be due to the material possessing a very high dielectric constant [9]).

The rapid decrease of μ_i, which could be interpreted, in general, as a resonance effect, was shown to shift to higher frequencies when air gaps were introduced into the core, an increasing air gap with a given frequency thus giving rise to a larger effective permeability (*fig. 10*). This result made it very probable that the permeability observed with no air gap was not a true constant of the material. Further evidence for this was obtained from experiments with laminated cores. Again, the apparent resonance effect was shifted to higher frequencies as the core was subdivided more and more in a direction perpendicular to the magnetic flux (*fig. 11*). Since the effect of ohmic eddy currents certainly could not

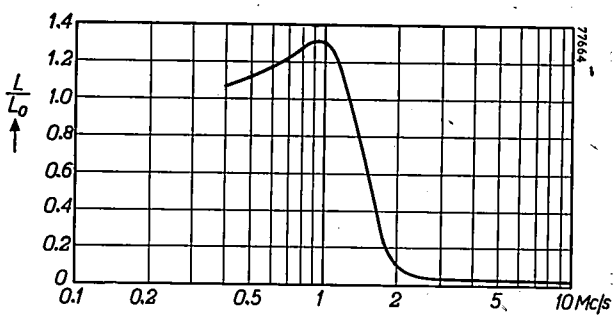


Fig. 9. The permeability of Ferroxcube III, measured on the sample bricks, for the Cosmotron core, plotted against frequency. The ratio of the self-inductance L of a ring-coil with a core of these bricks to that of the same coil at low frequencies L_0, is used as a measure of the permeability.

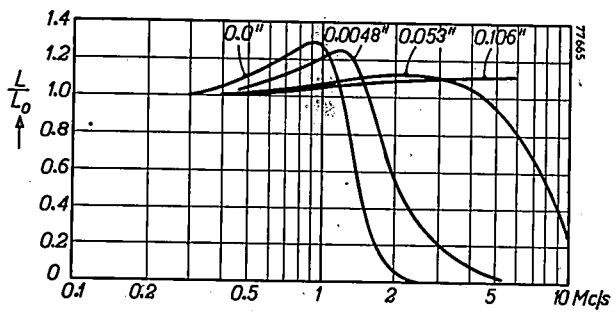


Fig. 10. Same as in fig. 9, with air gaps in the core of 0, 0.12, 1.35 and 2.7 mm respectively. The sharp drop in permeability occurs at higher frequencies as the air gaps get larger.

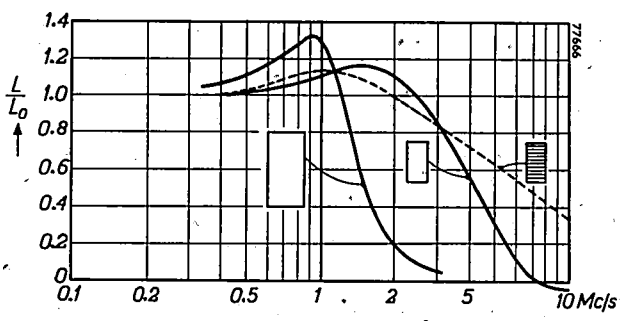


Fig. 11. Same as fig. 9, for cores of different cross-sections and subdivided into laminations as shown. The smaller the dimensions, the higher the frequencies at which the decrease occurs. (In one of the curves, negative values of L are seen to occur; this means that a dimensional resonance frequency has been passed and that the induction in the core then is oscillating in antiphase with the exciting alternating field.)

---

[7]) G. H. Bast, D. Goedhart and J. F. Schouten, A 48-channel carrier telephone system, Philips tech. Rev. **9**, 161-170, 1947 and **10**, 353-362, 1948/49.

[8]) Cf., for example, J. J. Went and E. W. Gorter, The magnetic and electrical properties of Ferroxcube materials, Philips tech. Rev. **13**, 181-193, January 1952.

[9]) F. G. Brockman, P. H. Dowling and W. G. Steneck, Dimensional effects resulting from a high dielectric constant found in a ferromagnetic ferrite, Phys. Rev. **77**, 85-93, 1950.

account for the results, one was led to consider the dielectric properties, and a relative dielectric constant ε_r of quite a surprising value was found, being as high as 50 000. The phenomena observed could then be explained by the superposition upon the usual eddy currents due to ohmic conduction of another type of eddy current which flows by dielectric displacement. Such currents will be out of phase with the ohmic eddy currents, and they can lead to a dimensional resonance in an inductor wound on a core of large enough cross-section. The wavelength of an electromagnetic wave in such a material is given approximately by:

$$\lambda \approx \frac{c}{f} \cdot \frac{1}{\sqrt{\mu_r \varepsilon_r}} \qquad (c = \text{velocity of light in vacuo}).$$

Standing wave effects can be noted when the smaller dimension of a core, perpendicular to the magnetic flux, is $^1/_2$ wavelength.

Interesting though this discovery was from a scientific point of view, it showed that Ferroxcube III would not be a suitable material for the Cosmotron core. However, another type of ferrite, Ferroxcube IV (nickel zinc ferrite), was available at Eindhoven, and a variety of this type met the specifications in a satisfactory manner. This material has a sufficiently small dielectric constant and a much higher resistivity than Ferroxcube III (of the order of 10^6-10^7 ohm cm), so that no dimensional resonance effects could occur at the frequencies to be used, even with bricks of large dimensions. Consequently, the frames for the Cosmotron were made from Ferroxcube IV. The permeability versus frequency curve of one large lamination is shown in *fig. 12*.

The large size of the core structure, which in this way influenced the choice of the material, also created some difficult manufacturing problems. Ferrite materials are prepared by *ceramic* methods, i.e. the powdered materials, after the addition

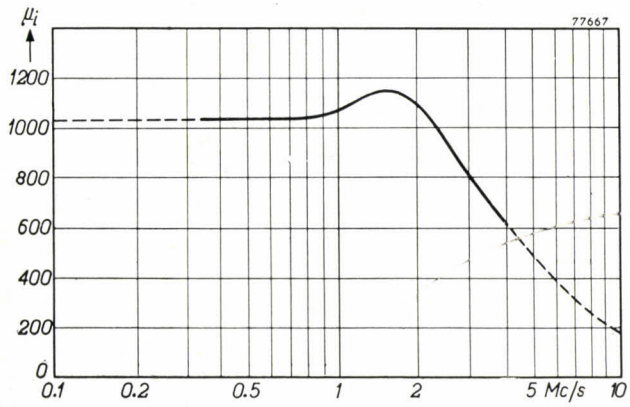


Fig. 12. Relative initial permeability μ_i of Ferroxcube IV, measured on one of the picture frames for the Cosmotron, plotted against frequency. A decrease is seen to occur at much higher frequencies than in the case of Ferroxcube III.

of suitable binders, are given the shape of the objects to be manufactured by pressing, extruding or the like, the shaped parts then being dried and fired at a high temperature. During the firing a shrinkage occurs, which in the case of ferrites may amount to 20-24% (on the linear dimensions). This large shrinkage, which is illustrated by *fig. 13*, imposes very strict requirements on the uniformity of temperature during drying and firing of large bricks such as were required for the Cosmotron core, in order to avoid cracks. The bricks were moulded in an 80 ton press and then dried very slowly, remaining for several days in a large drying chamber at about 40 °C. The firing was then performed in special ovens in an oxygen atmosphere, with a temperature cycle covering a period of several days. Such an oven is shown in *fig. 14*. The firing process, of course, must be

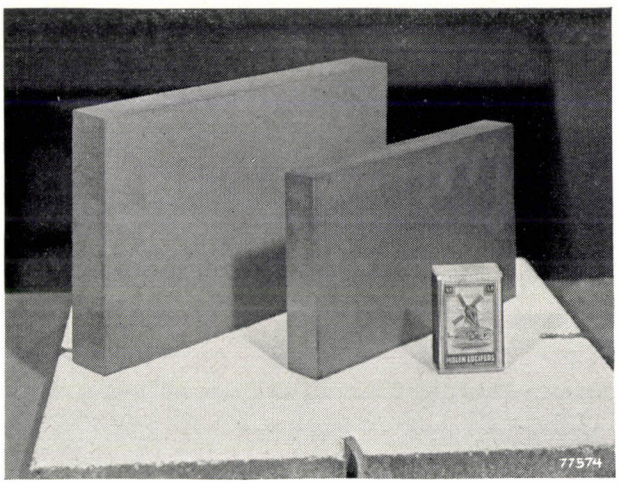


Fig. 13. A brick of Ferroxcube, before and after firing. During the firing at high temperature a shrinkage of 20-24% on linear dimensions occurs.

conducted very carefully, as the reactions occurring during firing are essential for the development of the special magnetic properties of a ferrite.

The bricks then had to be cemented together to form the "picture frames" as outlined in fig. 7. The unavoidable air gaps between adjacent bricks were reduced to a few microns by grinding the faces of the bricks with this precision; it will be appreciated that severe requirements were imposed on the rectangularity of the bricks, the permissible deviation of the 90° angle amounting to 1′ of arc in the largest surface of the brick and 1′30″ in the second largest.

It was established at Eindhoven that a plastic of the aethoxyline group, hardening at a temperature of about 180 °C, would be a suitable cement for ferrite materials. Experience with a sample picture

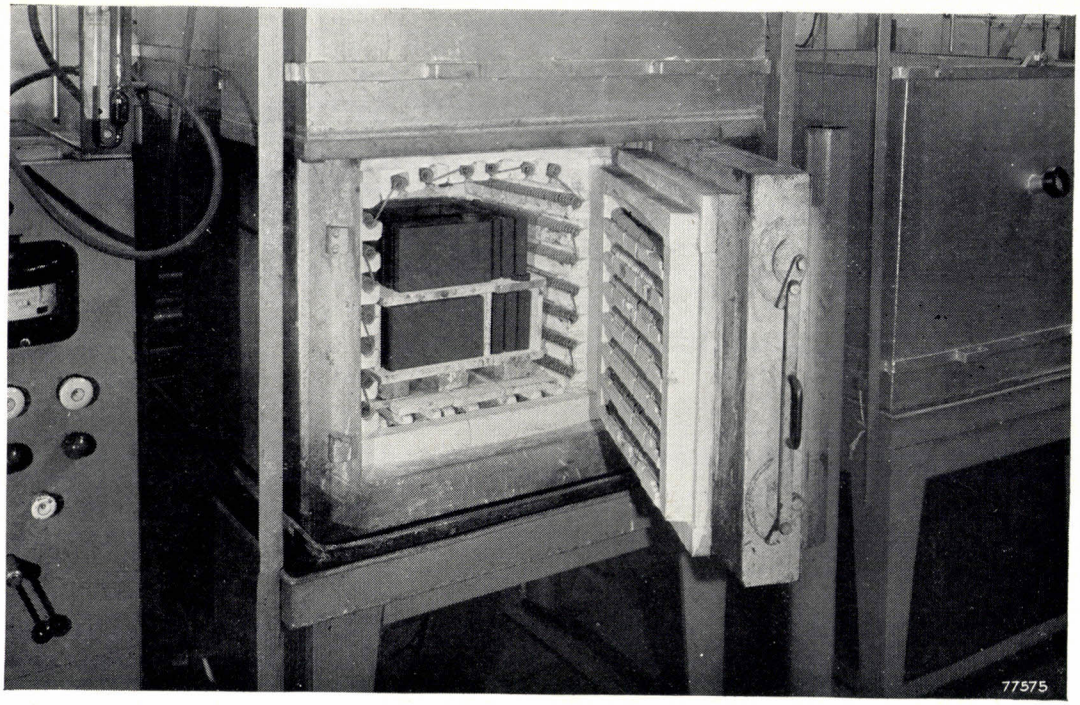


Fig. 14. One of the electric ovens in which the Ferroxcube IV bricks for the Cosmotron core were fired.

frame, used for preliminary measurements, had shown however, that the simple cementing and baking procedures applied to similar joints of other materials and also of Ferroxcube III were not satisfactory in the case of Ferroxcube IV. This could be ascribed to the *porosity* of the sintered Ferroxcube IV material, a property essential for giving the oxygen gas access to the innermost parts of the sample during firing, but unfortunately also giving water vapour, entering from the sides, access to the thin plastic films in the joint, causing deterioration in a few months.

A special procedure was therefore developed for the cementing of the Ferroxcube IV bricks. The main steps of this procedure involved a thorough cleansing of the surfaces to be cemented, a repeated impregnation in vacuo with the plastic cement and baking of the bricks. In the cementing

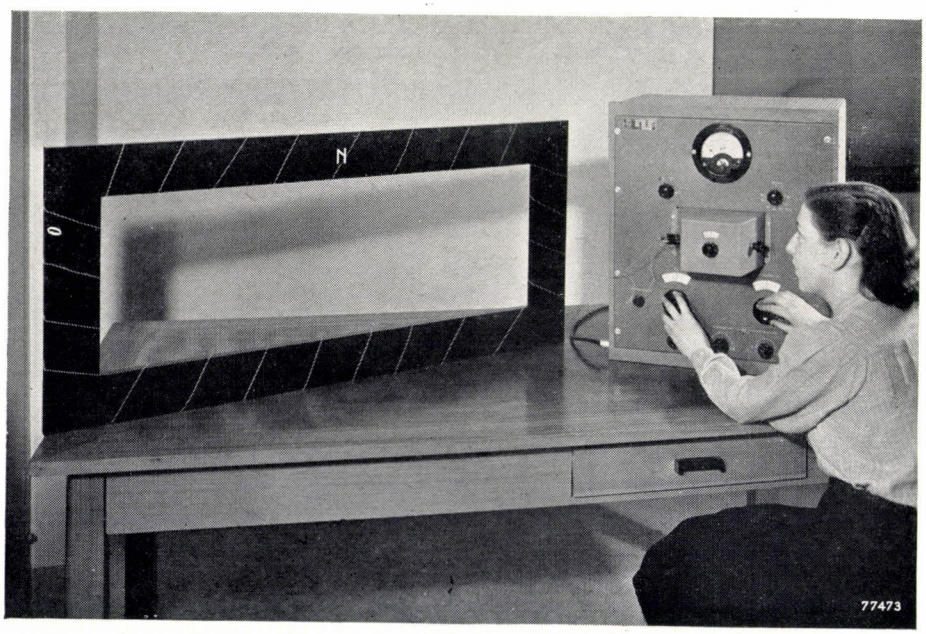


Fig. 15. One of the "picture-frames" for the Cosmotron core, during control measurements in the Ceramic Factory at Eindhoven.

process, a special tool enabled the bricks of a frame to be pressed together by a constant force during baking, so that most of the plastic layer applied between adjacent bricks is driven out of the joints during that part of the polymerization period when the viscosity is lowest. In this way the gap lengths

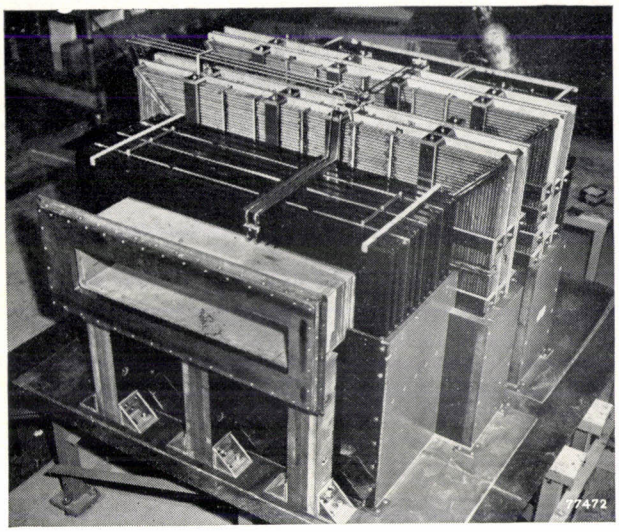


Fig. 16. The completed ferrite core of the Cosmotron, assembled in position. The Ferroxcube "picture frames" are seen in the centre and in the background of the photograph.

of non-ferromagnetic material were reduced to the small dimensions mentioned above. As a final precaution, the cemented frames were coated with a water repellent polyvinyl lacquer.

The breaking strength of the joints obtained is considerably larger than that of the Ferroxcube IV slabs themselves. This had been established by experiments, and it was proved again in shipping the frames to Brookhaven: when a box containing one of them was dropped from the hoisting crane, the frame was broken at two places, but not at any of the joints.

The photograph *fig. 15* shows one of the completed frames in the course of testing at Eindhoven. *Fig. 16* is a photograph of the complete core, assembled at Brookhaven and containing the two banks of picture frames.

Operation of the completed core under the load conditions imposed at Brookhaven has been reported to be completely satisfactory. The frames are separated by air spaces, and since the core is surrounded by a copper shield, forced air cooling can be applied. Even uncooled, however, the core functions without appreciable temperature rise. For further information about the performance of the oscillator and accelerator unit as a whole, the reader should refer to the recent publication by the Brookhaven group [1]).

The application of ferrites in "strong-focusing" accelerators

Since the Cosmotron was put into operation great interest has been aroused by the discovery by members and consultants of the Brookhaven group of the so-called "strong-focusing" principle [10]). Without entering into a detailed discussion, it may be said that a new method of reducing the amplitudes of radial and vertical oscillations of the revolving particles will allow the cross-section of the annular vacuum chamber to be reduced considerably, compared with the cross-section of the Cosmotron. Thus, with about the same magnetic energy — and therefore, very roughly, with a total cost of the same order — particles could be kept on a circular track with a diameter, say about 10 times that of the Cosmotron. In this way a particle energy of about 10 times that of the Cosmotron, i.e., of about 30 000 MeV, could be obtained (the momentum according to eq. (1) is proportional to the product BR, and at these high energies the energy is approximately proportional to the momentum; see also the graph of fig. 10 in the article quoted in [2]).

Plans for accelerators using this principle are actually being studied at several places. Such an accelerator would not make use of a single accelerating unit like that of the Cosmotron, but it would contain a large number of accelerating units, based on the continuously tuned resonator principle mentioned above as an alternative possibility. No

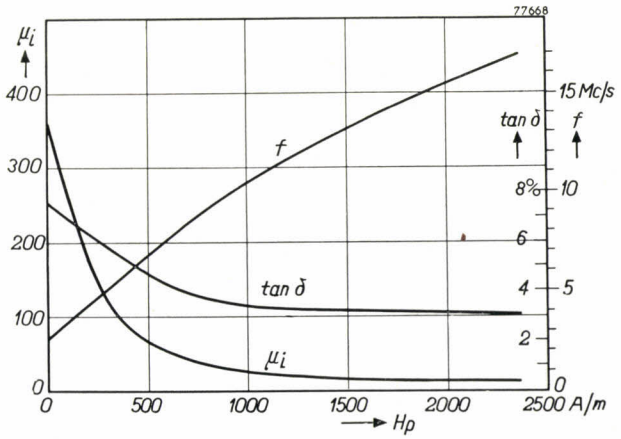


Fig. 17. A Ferroxcube material suitable for accelerators employing tuned resonant cavities has been developed at Eindhoven. The permeability μ_i of this material is varied by means of a polarizing field H_p and the curve f shows the frequency variation obtainable in this way (1 : 6). The figure of merit Q (= 1/tan δ) of a cavity filled with this material is greater than 10 even at the lowest frequencies.

[10]) E. D. Courant, M. S. Livingston and H. S. Snyder, The strong-focusing synchrotron — a new high energy accelerator, Phys. Rev. **88**, 1190-1196, 1 Dec. 1952.

transformer then would be needed, but oddly enough, ferrites again would have to be employed, on an even larger scale than in the Cosmotron: the resonators would be resonant cavities, containing massive cores of ferrite. The resonant frequency of such a cavity can be changed by a saturating magnetic bias field, much in the same way as in the master oscillator of the Cosmotron described above. As several tons of ferrite would go into each cavity, a total quantity of several hundred tons would be needed for a 30 000 MeV accelerator! The ferrite to be used will have to meet very exacting requirements. One of the tentative designs calls for a lowest resonant frequency of 3 Mc/s and an upper limit of 17 Mc/s. This frequency sweep (about 1 : 6) implies a corresponding permeability change of 1 : 36, to be obtained by applying a polarizing magnetic field. Another requirement is a Q at all frequencies of interest of 10 or more, at a maximum flux density at the lower frequency limit of 0.01 Wb/m^2. In co-operation with the Brookhaven group, the Philips Laboratories at Eindhoven have developed a ferrite which satisfies these requirements, as is illustrated by *fig. 17*. Samples of this ferrite are now under test at two laboratories associated with the high energy accelerator design.

Summary. The Cosmotron proton accelerator, designed and built by the Brookhaven National Laboratory, U.S.A., has already delivered a beam of protons with an energy of 2300 million electron volts and is expected to yield protons of 3000 MeV. The accelerator unit of this machine consists of a large "transformer", the ferromagnetic core of which is excited by an alternating current whose frequency during the acceleration period of one second varies from 0.35 Mc/s to 4.2 Mc/s. At these high frequencies only ferrite materials, which have very low eddy-current losses, could be used for the core. The major part of the ferrite core was manufactured by Philips at Eindhoven after a suitable variety of Ferroxcube meeting the permeability and figure of merit requirements had been selected by the Brookhaven Laboratory in co-operation with the Philips Laboratories at Irvington, U.S.A., and at Eindhoven. Some glimpses of the problems encountered in the manufacture of the Ferroxcube core are given in this article. Mention is also made of the development at Eindhoven of another variety of Ferroxcube which will be suitable for the accelerator units of machines based on the principle of "strong focussing", which are expected to yield particles of energies of 30 000 MeV or more.

THE PRINCIPLE OF THE MAGNETIC RECORDING AND REPRODUCTION OF SOUND

by W. K. WESTMIJZE.

621.395.625.3

The hysteresis and non-linear character of ferromagnetism would appear to render the magnetic recording of sound a hopeless task. Despite these difficulties, sufficiently faithful recordings on ferromagnetic tape were accomplished as early as 1907 by pre-magnetizing the tape. Nowadays, the same result is obtained by superimposing an HF alternating current on the signal current. For a correct application of this principle, a careful study of the magnetization process of a magnetic tape is required. Such an investigation is described here, covering the recording and reproduction processes themselves, the frequency characteristic, distortion, noise, and other phenomena which influence the quality of the reproduction. The conclusion from this work is that magnetic recording not only competes successfully with other sound-recording methods, but may even surpass them.

Introduction

The nature of the processes involved in the recording and reproduction of sound on a ferromagnetic tape will first be discussed. The apparatus (*fig. 1*) and some properties of the ferric oxide coated magnetic tape have been discussed in an earlier publication [1]). It was there pointed out that the material of the tape should have a certain remanence and not too low a coercive force, and consequently must show a certain degree of hysteresis. Special measures are then necessary to ensure linearity, i.e. to ensure that the resultant of two or more superimposed signals will be solely the sum of these signals.

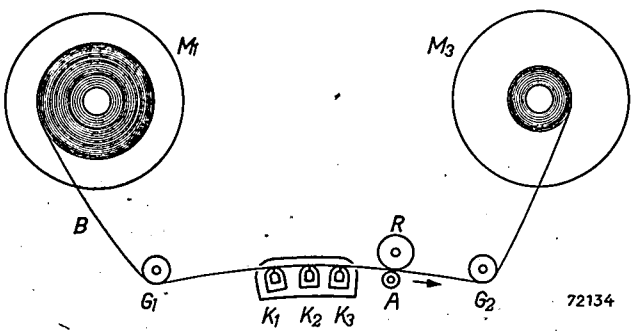


Fig. 1. Schematic representation of a tape-recorder. M_1 supply reel, M_3 take-up reel, B magnetic tape, K_1, K_2, K_3 erasing, recording and reproducing heads. A drive capstan, R rubber pressure idler, G_1 and G_2 idlers.

The process of magnetization as it occurs in the material of the tape will first be examined with the aid of a schematic representation. Next, the recording and erasing of signals will be discussed. Other matters to be dealt with include distortion (non-linearity between magnetization and signal strength), the reproduction process, the frequency characteristic, noise and the "print" effect.

Finally, the merits of the magnetic recording process will be compared with those of other methods of recording sound.

Schematic representation of the magnetizing process

As a rule, when working with ferromagnetic substances, a distinction should be made between reversible processes (in which the magnetization [2]) closely follows the field intensity), and irreversible processes. With small variations in the intensity of the external field, the magnetization is practically reversible. Irreversible changes occur when the variation in the field intensity exceeds a certain limit. Localised sudden changes — Barkhausen

[1]) D. A. Snel, Magnetic sound recording equipment, Philips tech. Rev. **14**, 181-190, 1953 No. 7; hereafter referred to as I.

[2]) In a material uniformly magnetized by a solenoid, the induction B is composed of two parts:

$$B = \mu_0 H + J.$$

In this $\mu_0 H$ represents the contribution of the magnetic field of the current and J (the magnetization, i.e. the magnetic moment per unit of volume; $J = \mu_0 M$) the contribution of the material (B and J measured in Wb/m^2, H and M in A/m). The same formula is valid at every point of a piece of material of any given form, whether or not it has been placed in a magnetic field:

$$\mathbf{B} = \mu_0 \mathbf{H} + \mathbf{J},$$

where **B**, **H** and **J** represent, respectively, the induction vector, the field vector and the magnetization vector, at any point in the field. In general the magnetic *field* is built up from various contributing elements:

$$\mathbf{H} = \mathbf{H}_u + \mathbf{H}_m,$$

$\mathbf{H}_u$ representing the external field, originating in electric currents or other magnets, and $\mathbf{H}_m$ the field of the "magnet poles", which, due to the magnetization, occur on the surface and at discontinuities in the material.
In the case of a toroid $\mathbf{H}_m = 0$. In other cases, $\mathbf{H}_m$ and **J** are usually in opposite directions and, since a decrease of $\mu_0\mathbf{H}$ gives a greater decrease of **B**, the magnetization **J** will be smaller than it would be if $\mathbf{H}_u$ only were active. This is known as *demagnetization*. It should be borne in mind that the induction lines are always closed lines, as the induction **B** is subject to the law.

$$\text{div}\,\mathbf{B} = 0.$$

jumps — then occur in the various domains. When the field intensity changes in the opposite direction the magnetic processes will again be at first reversible and then irreversible. In general, the critical field intensity will have different values for various elementary domains of the material. The resultant hysteresis loop (figs 2*b* and *c*) is obtained by superposition of the magnetization curves of the various domains (*fig. 2a*). This complicated behaviour may be represented [3]) by the idealised magnetization curves of figure 3, which are built up of two fixed irreversible branches, connected by two (less steep) reversible branches. The slope of the latter corresponds to the initial permeability (*fig. 3*).

Comparison with *fig. 4* shows that the real magnetization curve possesses the same character, apart from the rounding off of the sharp edges.

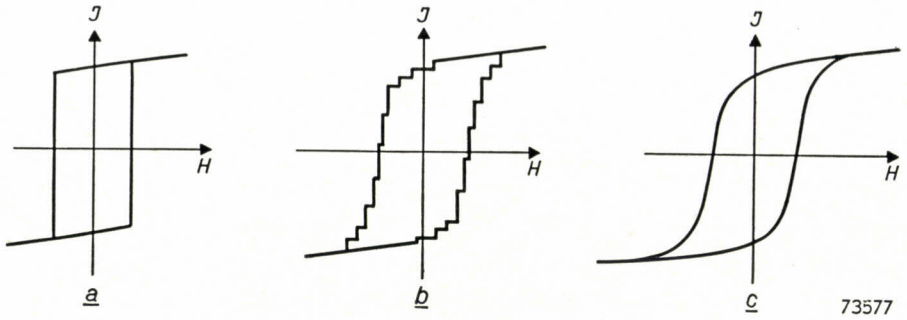


Fig. 2. Idealized and actual magnetization curves: (*a*) shows the curve for a particle magnetizable only in a positive or negative direction, (*b*) the superimposing of a small number of these curves, and (*c*) the same for an infinite number of particles.

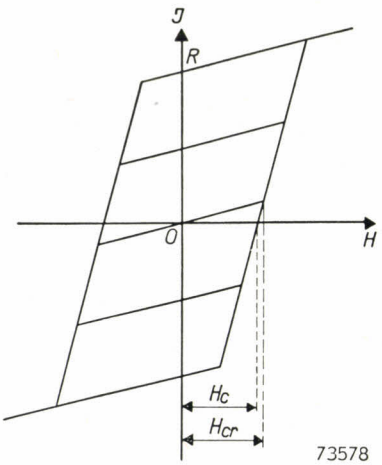


Fig. 3. Schematic magnetization curve, built up of two steep irreversible branches, connected by less steep reversible branches. The origin *O* represents the non-magnetic state, *R* represents the maximum remanence at zero field strength. H_c represents the coercive force, H_{cr} the critical field strength at which the magnetization becomes irreversible.

Recording method

In the magnetic recording of sound, the tape — consisting of a homogeneous magnetic material or a carrier coated with magnetic material — is passed across the gap of the recording head (*fig. 5*). This consists essentially of an electro-magnet, the magnetic field of which, within certain limits, may be assumed to be proportional to the energizing current. The latter is derived from the original signal via an amplifier.

The intensity of the magnetic field of a "ring-shaped" recording head is at its peak value in the gap, where, however, it cannot be used. The active portion of the field is the *stray* field, in the vicinity of the gap. It decreases rapidly in intensity in the lengthwise direction of the tape and in a direction perpendicular to the tape (i.e. the direction of the thickness). If, for example, the recording gap has a length [4]) of 10 μ, it may be assumed that the

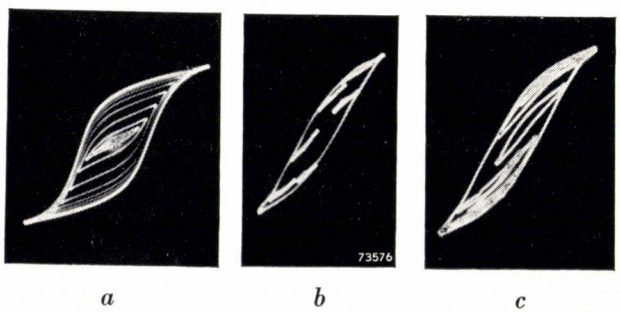


Fig. 4. Actual magnetization curve, which, by means of an integrating circuit is displayed on a cathode-ray tube and photographed. Curve *a* is obtained by using an alternating current with a gradually decreasing field intensity, curves *b* and *c* by superimposing an alternating current with a higher frequency and a small amplitude on a given alternating current. It is clear, particularly from *c*, that irreversible processes also occur on the reversible portions, which for reasons of simplicity were neglected in the idealised magnetization curve. The scale of representation differs for the three figures.

[3]) A similar procedure was followed by H. Toomin and D. Wildfeuer (Proc. Inst. Rad. Engrs. **32**, 664, 1944).

[4]) The "length" of the gap in this type of recording head is defined from the lengthwise direction of the tape. The "width" of the gap corresponds to the width of the tape.

effective field extends only some tens of microns from the gap. Consequently, when using a homogeneous tape of thickness say, 60 μ, only a layer of about 15 μ facing the head will be magnetized.

Starting with a demagnetized tape, no permanent magnetization will be imparted to it unless the field strength of the applied signals is greater than a certain critical value H_{cr}. This may be seen

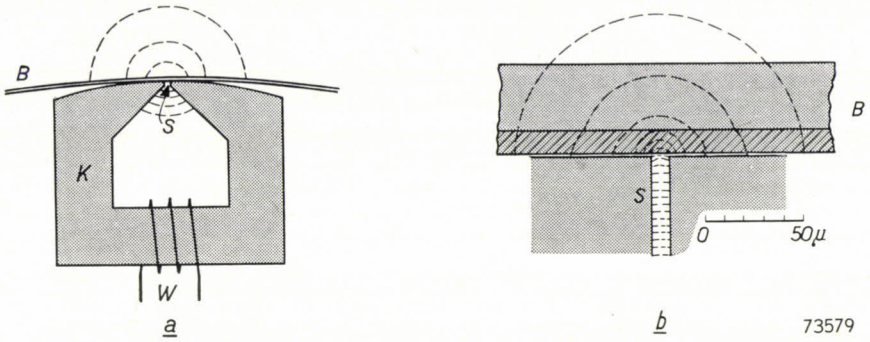


Fig. 5. Recording head. *a*) A core of laminated soft iron (K) with a very short air gap (S) is magnetized by a current flowing through winding (W). This creates a stray field (dotted lines) around the gap. The tape (B) is carried across this gap and every particle traverses this stray field, the intensity of which increases when approaching the gap and decreases after passing the gap.
b) enlarged drawing of gap with tape.

For a coated tape, there is therefore no point in making the coating thicker than 15 μ.

When a particle of the magnetic material moves past the recording head, it is exposed to a magnetic field varying from point to point as regards intensity and direction; in general it will also vary with respect to time. For simplicity, we will first confine ourselves to the case where the frequencies of the signal are so low that during the time that a particle is exposed to the active field, the signal field is independent of time. With a tape speed of 0.7 m/sec and an active area of 50 μ, giving a transport time of about 10^{-4} sec, this condition will be satisfied when the frequencies are lower than 1000 c/s.

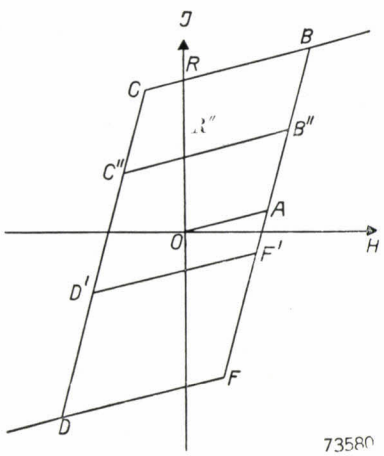


Fig. 6. If the material is exposed to a strong field which first increases and then decreases, the reversible portion OA is first described and then the loop $BCDF$. If now the field starts to decrease when the point B'' is reached, then $B''C''$ is described. A similar effect is obtained for a negative field after it has reached its largest value, say, at D' (line $D'F'$).

from the schematic magnetization curve of fig. 3. (In fact, there will be a very small remanent magnetization.) Only when the peaks of the signal

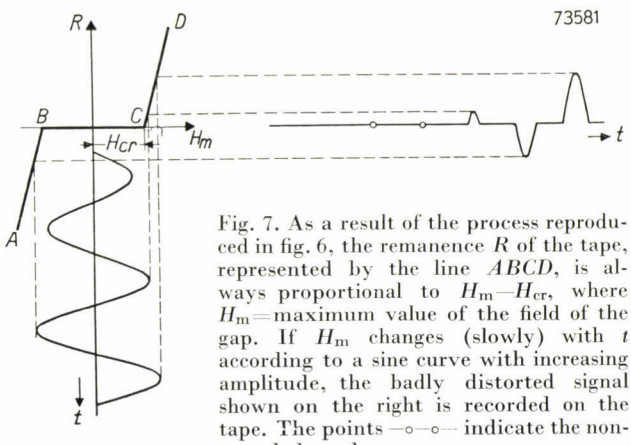


Fig. 7. As a result of the process reproduced in fig. 6, the remanence R of the tape, represented by the line $ABCD$, is always proportional to $H_m - H_{cr}$, where H_m = maximum value of the field of the gap. If H_m changes (slowly) with t according to a sine curve with increasing amplitude, the badly distorted signal shown on the right is recorded on the tape. The points —o—o— indicate the non-recorded peaks.

field exceed the critical value H_{cr} will the magnetization change, say, from O via A to B'' (*fig. 6*), and from there fall back to R'', rendering the tape permanently magnetized. Evidently the remanence OR'' is not now proportional to the value H of the signal field, but to $H - H_{cr}$, with the result that the signal is badly distorted.

The result of the recording of a sinusoidal signal of a low frequency and gradually increasing amplitude would be as represented in *fig. 7*.

Of course, the distortion could be diminished by the use of a material with a lower coercive force. However, as pointed out in article I, a certain coercive force is needed to prevent distortion of the

sound track by stray fields and to counteract the influence of demagnetization in regions of sharp changes in the magnetization (i.e. small wavelength, corresponding to high frequency).

One solution of the linearity problem which has been used from the early days of magnetic recording [5]) is as follows. Before the tape is carried past the recording head, it is exposed to a direct field of such intensity that it becomes saturated (*E, fig. 8*) and, consequently, when passing the head, it is in the state R_V. A direct current is now superposed on the signal current of such strength that the point *A* would be reached if no signal were present. In this case, on leaving the region of the head, the remanent magnetization is proportional to the signal field; if this is zero, the tape again becomes unmagnetized. Linearity can indeed be obtained by this method, and usable records can be made. However, the noise inherent in tape that has gone through this cycle is much stronger than that of a virgin tape. The causes of this phenomenon will be given later when discussing the problem of noise.

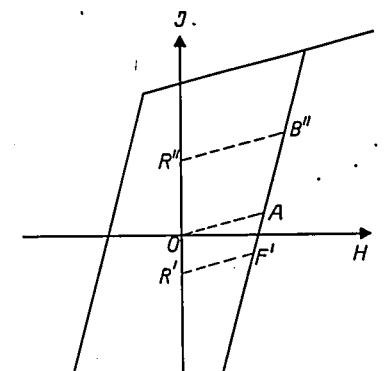


Fig. 8. D.C. biasing. The tape is first saturated (*E*) and then passed across the recording head. The direct current biasing field alone would lead to a state *A*; with modulation by the signal, such states as e.g. *B''* and *F'* would be reached. The remanence left in the tape after leaving the active field in the head is then 0, *R''* or *R'*.

Modern methods do not make use of a direct field; nowadays an alternating field is superimposed on the signal, of such a frequency that several cycles occur during the passage of the tape through the active portion of the stray field round the gap.

Assuming again a transport time for the tape of about 10^{-4} sec, it is clear that the frequency of this biasing field should be higher than 10^4 c/s. In practice a much higher frequency is used, in the ultrasonic region, in the neighbourhood of 10^5 c/s, so as to avoid the disturbing effects both of this frequency itself and of audible beat-tones between signal and biasing currents.

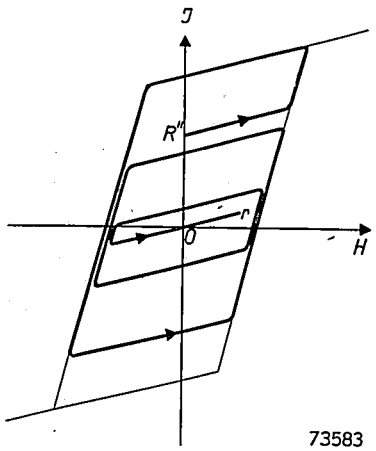


Fig. 9. The process of erasing. Imagine that the remanence *R''* is the starting point. The increasing alternating field will saturate the material after a certain number of periods and make it traverse the large cycle. If now the intensity of the field decreases so slowly that the positive and negative peaks are nearly identical, the magnetization will follow the path shown till finally, the reversible branch *r*, through the origin, will be reached, reducing the remanence to zero.

The amplitude of this field will, of course, vary in the length direction of the tape. If this field alone is applied, a particle arriving at the gap will first be taken through several hysteresis cycles of increasing amplitude until it is saturated, and will then go through a number of diminishing cycles, so that it is demagnetized as it leaves the head. (*fig. 9*). This process occurs whether tape was previously magnetized or not, and is put to use in the erasing head.

Consider the case when the high-frequency alternating field is superposed on the relatively steady signal field of the head. Now the positive and negative peaks of the field are not situated symmetrically with respect to zero field, but are symmetrical about the field corresponding to the signal. If the amplitude of the alternating field is now decreased below the critical value H_{cr}, the magnetization process again becomes reversible, although these reversible states are not represented on a line through the origin but on a line parallel to that line (*fig. 10*). For simplicity, suppose that only the amplitude of the alternating field decreases as the tape moves, the signal field being substantially constant: then the remanent magnetization will be that corresponding to the signal field and this will remain even when, finally, the signal field, too, is removed. The final state is such that when all fields have been removed, a remanence *R''* remains.

[5]) V. Poulsen (inventor of the magnetic recording method, see I) and P. O. Pederson, U.S. Pat. 873083, 1907.

The essential principle of the process is, therefore, the reduction of a hysteresis loop to a single reversible branch inside that loop. This transition takes place when the amplitude of the high frequency field (*i.e.* the difference in the value of H at points B' and D', fig. 10) is exactly equal to $2\,H_{cr}$. The average value of the intensity of the signal field at that moment is the determining factor of the resulting remanent magnetization. This remains true though in reality its intensity changes at the same rate as the intensity of the alternating field, as the tape passes through the recording head. It is irrelevant that the signal field is not constant with time but is really a relatively slowly alternating field.

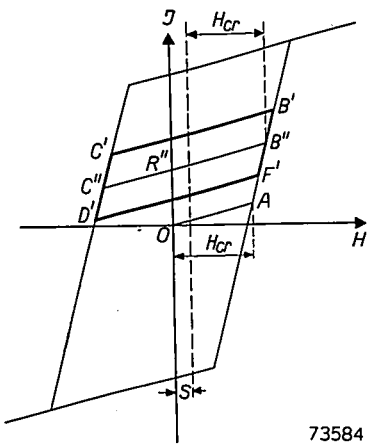


Fig. 10. Recording using a high-frequency biasing field. A rapidly alternating biasing field is superimposed on the constant field S. The positive peaks are always higher than the negative ones. When the biasing field has reached its maximum positive value in B', the cycle $B'\ C'\ D'\ F'$ is traversed. If the amplitude of the biasing field falls below the value H_{cr}, the magnetization remains on the reversible branch through B''; consequently, after taking away the biasing and signal field, the remanence R'' remains.

The remanence is always determined by the mean value of the two extreme field intensities (equal to the signal field) when the difference between them is exactly equal to twice the critical field intensity. This situation arises immediately after the gap has been passed, when the particle has already entered the decreasing part of the strong field: the above value of the remanence is now recorded on the tape.

Distortion during recording

It is clear that owing to the introduction of the HF biasing field, the remanence has become proportional to the intensity of the signal field because the irreversible branch is represented by a straight line and all reversible branches are parallel to each other. Thus the distortion occurring in the absence of the biasing field has been eliminated in principle. A certain amount of distortion still remains, however, due to various causes.

Distortion is inevitable when the signal current is so strong that a higher field exists than that corresponding to the maximum remanence of the material (R, fig. 3). Consequently, care must be taken that the signal strength does not exceed a certain level.

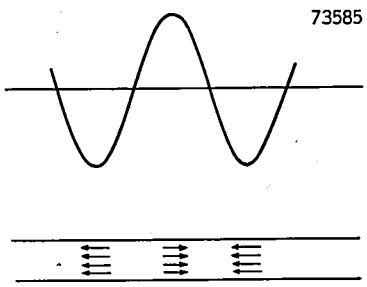


Fig. 11. Intense demagnetization occurs in the sound track of a sinusoidal signal with short wavelength (high frequency).

Distortion may also arise due to the demagnetization effect of adjacent parts of the tape. Consider, for example, a sinusoidal signal; if the positive peaks give rise to a magnetization directed to the right, the magnetization corresponding to negative peaks will be directed to the left (*fig. 11*). At the point under consideration, the oppositely directed magnetization in adjacent domains generate a field that tends to weaken the magnetization, especially when maxima and minima are close together (corresponding to shorter wavelength, i.e. higher frequency) and when the magnetization is strong. The relation between this demagnetizing field and the magnetization causing it, is represented in *fig. 12* for a fixed, low frequency, by the line I.

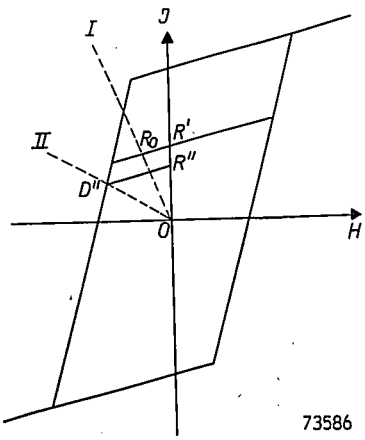


Fig. 12. As the demagnetizing field intensity is proportional to the magnetization, the resulting magnetization can be represented by a line through the origin, the slope of which is determined by the geometry of the magnetization pattern, for example, by the wavelength. With a small demagnetizing factor (I), the remanence R_0 is proportional to the signal, but not if the sloping line (e.g. II) and the reversible branch would intersect outside the hysteresis parallelogram.

When it has left the head, the part of the tape under consideration then attains a state corresponding to point R_0. As long as the reversible branches intersect the line I at points within the hysteresis loop, the proportionality between signal and remanence is not impaired, not even when the demagnetization is diminished due to the magnetic "short-circuiting" effect of the soft iron of the reproducing head (as a consequence of this, the line I in fig. 12 is replaced by a line with a steeper slope).

The proportionality is lost, however, when the intersection point would lie outside the hysteresis loop. In fig. 12 this is the case when, at higher frequencies, the demagnetization is represented by the line II. This causes a shift of the remanence from the line $R'R_0$ to the line $R''D''$.

Optimum value of the biasing current

As is evident from the above, the linearity is influenced by the magnitude of the biasing current as well as by the strength of the signal. If the biasing current is zero, the magnetization is zero until the signal field H exceeds H_{cr}; it then increases in proportion to $H—H_{cr}$, until saturation is reached. This is the large distortion mentioned in the beginning, which occurs when no biasing field is used. Even when a biasing field is present, the same state of affairs is reached if its amplitude is too small, the only difference being that the increase in magnetization starts somewhat earlier. With a biasing-field amplitude equal to H_{cr}, the magnetization is proportional to the signal, except for the limits set by the saturation. In practice (for reasons to be given later) the biasing field in the neighbourhood of the gap is chosen higher than H_{cr}. In that case the signal is recorded on the tape a little further on, at the point where the biasing-field amplitude is exactly equal to H_{cr}. The more intense the biasing field, the further away from the gap this occurs and the weaker, at a given signal current, the operative signal field. The more intense the biasing field H_b, the more it must decrease to reach the value H_{cr}. The signal field must suffer a corresponding decrease; hence, for a given signal current, the remanent magnetization decreases as the biasing current increases.

These conditions are represented in *fig. 13a*. As the curvature of the line representing magnetization as a function of signal field is a measure of the distortion, it is clear that with weak signal currents, the distortion decreases with increasing biasing current (H_b). For strong signal fields, however, (depending on the exact values) the distortion first decreases with biasing field and may then suffer an increase due to saturation, followed by a further decrease. With average signal strengths the behaviour lies between those stated above.

In practice, a value of the biasing current is chosen such that it corresponds nearly to the secondary maximum for strong signals. This is a compromise, adopted because, as will be seen later, a higher value of the biasing current must be avoided, whereas with lower values, distortion may occur owing to the fall off in intensity of the field for particles furthermost from the tape surface.

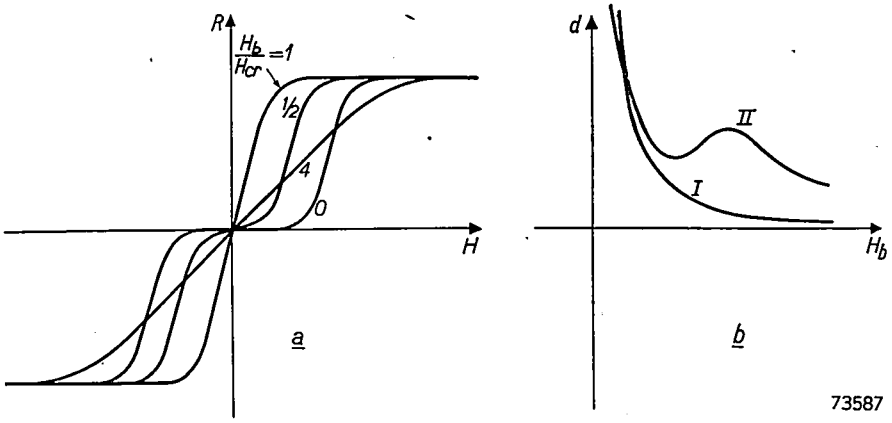


Fig. 13. *a*) Relation between magnetization and signal with various amplitudes of the HF biasing field H_b. *b*) Curve representing the distortion for a weak signal (I) and a strong signal (II), as a function of the HF biasing field.

Effect of the biasing current on the frequency characteristic

The sound signal is recorded on the tape at the moment when the intensity of the biasing field is equal to H_{cr}; in practice, the value of H_{cr} differs

slightly for different particles, so that there is a spread of the critical biasing field, and the signal is not completely recorded until the lowest value of the critical biasing field is reached. The position and the length of the recording zone of course depend on the strength of the current producing the biasing field. If during the recording of a rapidly alternating signal the value of the biasing field does not decrease quickly enough, it is possible that a part of the tape, after having passed the spot in which the intensity of the biasing field is exactly equal to the critical value, is brought again on the irreversible branch by the rapidly increasing signal field. This may cause a false rendering of the signal. The higher the signal frequency, the sooner this result is to be expected, because in that case the signal field, when passing the above-mentioned zone, changes in intensity more rapidly. This false rendering results in a certain attenuation of the recording of the high frequencies.

There are other and more important consequences of the differing coercivities among the particles of the tape. Two particles at the same spot on the tape but with different H_{cr} values will not record the signal at the same moment and therefore will record a different phase. This phase shift of the signals on the sound track results in "interference attenuation" at high frequencies.

At a given biasing current, the points where $H = H_{cr}$ will not lie on a plane perpendicular to the tape, but on a curved surface. Consequently, particles with the same H_{cr} but situated in different depths of the tape will meet the critical field intensity at different distances from the gap. This also makes the phase shifts of the recorded signal dependent on the depth in the tape. During *reproduction*, it is precisely those particles which lie instantaneously on the plane of symmetry of the gap which reproduce the signal, so that the summation of the out-of-phase signals again results in attenuation of the higher frequencies.

If the biasing current is stronger, the region where $H \approx H_{cr}$ lies farther away from the gap and is more extended. In this case the attenuation is noticeable at lower frequencies.

The distortion caused by demagnetization, mentioned earlier, which becomes worse with increasing frequency, can also attenuate the high frequencies, for, owing to this distortion, not only are the higher harmonics superposed on the fundamental, but the fundamental is itself attenuated.

It should be noted that the schematic representation of the magnetization process which has been used in this article, is only an approximation and cannot adequately explain every detail. To go into further detail would be to go beyond the scope of this article [6]).

The reproduction process; influence on the frequency characteristic

In order to convert the magnetic pattern on the tape back into an electrical signal, the tape is carried past the reproducing head, the construction of which is similar to that of the recording head (*fig. 14*).

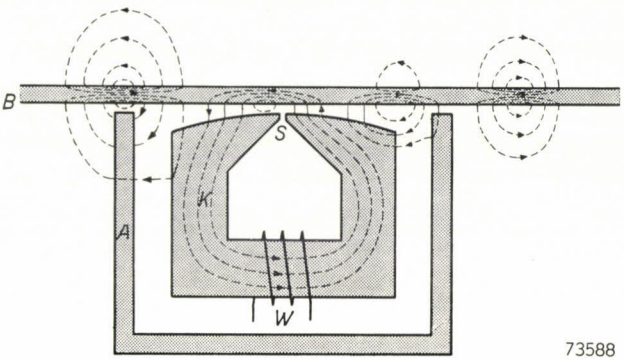


Fig. 14. Reproducing head K with winding W, gap S and screening A. B represents the magnetized tape. The lines of force are indicated schematically.

Considering an arbitrary cross-section of the tape, the recorded magnetization J gives rise to a total magnetic flux Φ. As the lines of magnetic induction are always closed, they must close outside the tape, so that an equal and opposite flux is found there.

When the tape makes contact with the head, the lines of induction will close mainly via the material of the head (owing to its high permeability) and so link the reproduction coil. Thus a flux flows through this coil, which to a first approximation is equal to the flux traversing the cross-section of the tape facing the gap. When the tape moves on, changes of this flux induce the signal voltage in the coil. Due to the good magnetic conduction of the iron circuit, whereby the demagnetizing field is reduced, the flux mentioned is even greater than it would be with a tape moving freely in the air.

The closer the tape comes into contact with the head, the more will the demagnetizing field be reduced. If the surface of the tape is not completely smooth it will be in *true contact* with head only at certain points. In this case, the demagnetizing the field is not completely removed, and at small wavelengths a decrease in the magnetization will occur. This decrease will be greater as the slope of reversible branches becomes steeper (and thus the reversible permeability greater).

[6]) See W. K. Westmijze, Philips Res. Rep. **8**, 148-157, 161-183, 1953.

It should be noted, however, that not all the induction lines will enter the pole pieces and hence do not all link with the coil. One reason for this is that the magnetic resistance of the iron circuit can never be quite zero, so that small magnetic potential difference will always remain across the gap S, and as a result, a flux proportional to this potential difference will pass across the gap. Therefore it is desirable to have a gap with a magnetic reluctance that is high compared with that of the iron circuit, i.e. a "long" gap with the small "height" [4]).

Taking into consideration constructional demands, namely the mechanical strength of the poles, the machining precision and the duration of life of the head, the height of the gap has been reduced as far as possible. It should be borne in mind that during use, the tape is moving continuously across the head at a high speed and with a fairly high pressure and that the magnetic particles of the tape are in fact a form of the well-known polishing agent jeweller's rouge. The resulting wear reduces the height of the gap to such an extent that after say, 500 hours, the head must be exchanged. The life time of the head of course also depends on other details of its construction and on the quality of the tape used.

A large gap length cannot be used, however, for more fundamental reasons: while it is true that for long wavelengths practically all lines of force emanating from the tape enter the poles and link with the coil, this is no longer true when the wavelength is comparable to the gap length. This occurs at relatively low frequencies if the gap length is

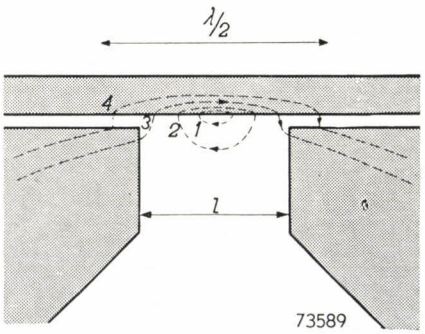


Fig. 15. The form of the induction lines, when the wavelength λ is comparable with the length l of the gap (the cause of the fall of the reproduction characteristics at high frequencies).

large. In this case a large number of the lines of force do not enter into the poles at all, but close through the air of the gap (*fig. 15*). To a first approximation this can be expressed as follows: the flux through the coil (gap length $= l$, wavelength $= \lambda$) is approximately proportional to

$$\frac{\sin(\pi l/\lambda)}{\pi l/\lambda}.$$

For small values of l/λ, i.e. if $\lambda \gg l$, this expression is practically equal to unity and independent of λ. With higher values of l/λ, i.e. at high frequencies, the value of the expression decreases and with it the frequency characteristic. If $\lambda = l$ and consequently $l/\lambda = 1$, the expression $= 0$ and no flux at all flows through the coil. Therefore the gap must be shorter than the smallest wavelength essential for reproduction.

A second cause for the lines of force not linking with the head lies in the fact that at very long wavelengths (*fig. 16*) the lines of force take a path

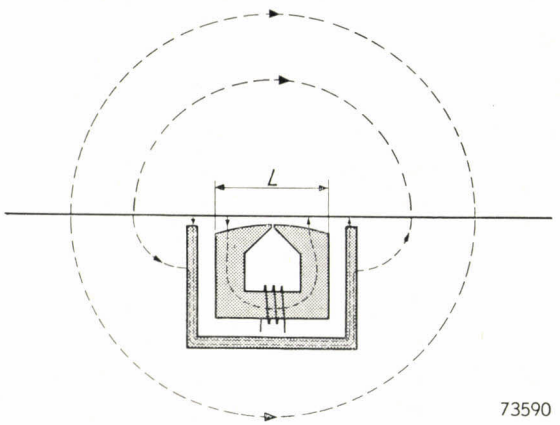


Fig. 16. The form of the lines of induction, when the wavelength is comparable to the dimension L of the head (a source of irregularities in the frequency characteristic at low frequencies).

around the head, especially so when it is surrounded by the inevitable shielding (A). Supposing now that only the lines of force over a certain length $L \gg l$ of the tape contribute to the flux through the coil, the flux becomes proportional to:

$$\int_0^L \sin(2\pi x/\lambda)\mathrm{d}x = \frac{\lambda}{2\pi}\,[1 - \cos(2\pi L/\lambda)].$$

This means that for wavelengths $\lambda \gg L$ only a negligible part of the flux traversing the tape at the point where it faces the gap, enters the coil. As the wavelength decreases, a maximum is reached for $\lambda = 2L$, and with still shorter waves, maxima and minima (= zero) would alternate for $L = (n+{}^1/_2)\lambda$, and $L = n\lambda$, respectively, in which n represents a whole number. This is not so serious as it sounds, because the transition between the flux making its way through the coil and the flux taking other paths is not sharp but gradual, which

results in the maxima and minima being smoothed to a considerable degree, making them imperceptible for $\lambda \ll L$. Consequently, the result of this effect is mainly a decrease in the useful flux for wavelengths greater than L (*fig. 17*).

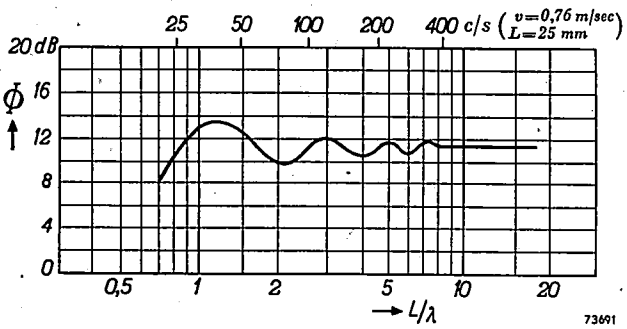


Fig. 17. The form of the useful flux Φ (in dB) as a function of the wavelength at long wavelengths. Φ is plotted against L/λ (thus the frequency increases from left to right). If $L =$ 25 cm and the speed of the tape = 0.76 m/sec, the maximum at the left corresponds to a frequency of 37 c/s.

In general, the decrease of the useful flux at *short* wavelengths is more serious than at long wavelengths. The frequency at which this decrease occurs can in principle always be raised by increasing the speed of the tape. In practice the tape speed is increased to such a value that the remaining frequency correction can be successfully applied in the amplifier, without introducing too much noise and other disturbances. On the other hand, the amount of tape used becomes larger and the playing time decreases when the speed of the tape is increased. Moreover, the decrease of the useful flux at longer wavelengths, which corresponds at small speeds of the tape with frequencies below the audible range, is noticed as an irregularity in the characteristic and as an attenuation at lower, audible frequencies. It will also be seen below, that at higher tape speeds the "print-effect" becomes stronger at higher frequencies (and consequently becomes more disturbing). On account of this it is necessary to choose a tape speed which is a compromise between the reproduction of the high frequencies and the objections just mentioned.

Overall frequency-characteristic: "transimpedance"

In the preceding section the influence of the recording and the reproducing process on the frequency characteristic have been discussed separately. In practice it is difficult to separate the one influence from the other, as the magnetization brought about by the recording process can be measured only with a reproducing head. The voltage generated at the reproducing head can then be measured as a function of the current sent through the coil of the recording head. If we assume that this current is a pure sinusoidal current the relation between the voltage and the current mentioned can be determined as a function of the frequency $f = \omega/2\pi$. This relation has the dimensions of an impedance and may therefore be called the "transimpedance".

If it is desirable to study the influence of one of the components of the system — recording head, tape or reproducing head — separately, this can only be carried out by relative measurements, which means that one of the components should be replaced by a corresponding part and a comparison made between the transimpedance measured before and after the exchange.

Ideally, i.e. excluding very low and very high frequencies, the transimpedance increases linearly with the frequency. It is therefore comparable with the impedance of an inductance.

The transimpedance depends not only on the frequency but also on the number of turns of the coils of both heads. Within certain limits we have a free choice in these, and thus the value of the transimpedance can still be varied at will.

Sometimes it is useful to consider, apart from the transimpedance Z_{12}, another quantity which it is convenient to call the "transfactor" (T), defined by the relation:

$$T = \frac{Z_{12}}{(Z_1 Z_2)^{1/2}}.$$

(Z_1 is the impedance of the recording head, Z_2 that of the reproducing head.)

This factor is dimensionless, and is independent of the number of turns of the coils. In the ideal case it is a constant, independent of the frequency (*fig. 18*), of magnitude about 0.01.

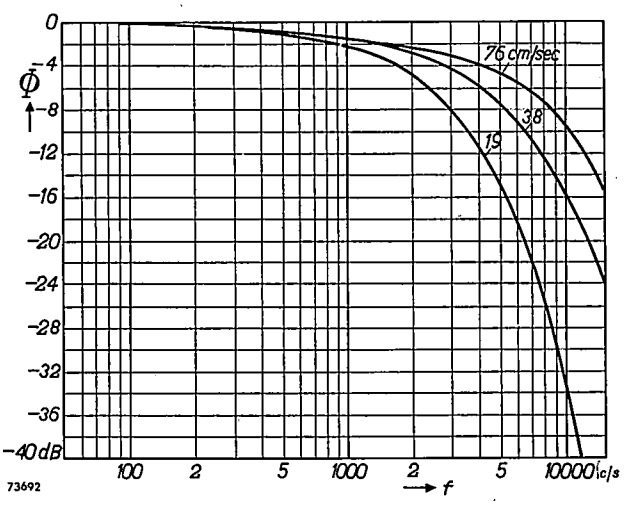


Fig. 18. Corrected frequency-characteristic (flux Φ through the reproducing head) for tape velocities of 19, 38 and 76 cm/sec resp. At the same time the curves approximately represent relative values of the transfactor as a function of the frequency f.

Noise

Ground noise

The magnetic layer of a recording tape is built up from very small particles. If the layer is examined under an electron microscope, the size of the particles (*fig. 19*) is found to be about 0.1 μ. Now we can assume that every particle still contains some Weiss domains, separated by walls. If the tape is completely demagnetized, the fields of the Weiss domains will largely compensate each other, but still a weak, multi-polar field will emanate from every small particle. As a result of this, some lines of force will leave the material and create a micro-field, which irregularly changes its sign along the tape. This micro-field is easily detected by running the tape across the reproducing head, with its narrow gap of 10 μ. If then, a demagnetized or virgin tape is carried across the reproducing gap, it will continuously induce low irregular voltages into the coil of the head, which, after amplification, are audible as noise, the ground noise. This noise limits the weakest signal still reproducible and determines in that way the "dynamic range". By this we understand the difference in level, measured in dB, between the largest signal which can be recorded without appreciable distortion and the smallest signal still audible above the ground noise.

In general, the more particles available to record the signal, the larger the dynamic range. If the average size of the particles is 0.1 μ, and the tape velocity is 0.76 m/sec, with a tape of width 6.3 mm and a magnetic layer thickness of 15 μ, the number of particles passing the gap per second will amount to about 10^{13}. A simple calculation of probabilities shows that in the frequency range $0<f<10^4$ c/s the dynamic range will be at least 90 dB (in reality, due to partly unknown causes, about 70 dB is usually measured).

Modulation noise

A magnetized tape possesses more noise than a clean tape. It is true that this higher noise level is partly masked by the signal, but it is still audible, especially when the signal consists of a single note (pure sine function), which now sounds somewhat harsh. In order to understand this extra noise, it is necessary to consider the structure of the tape. The particles of the magnetic material are not evenly distributed through the carrier, but are grouped in clusters (fig. 19). In the demagnetized state, in every cluster the magnetic moments of the separate particles are irregularly distributed. In the magnetized state, in every cluster a more or less evenly directed magnetization of the particles is present. Due to the distance between the clusters and due to the fact that each cluster now represents a dipole, there is more chance that lines of force from the microfield will emerge outwards. This may also be expressed in the following way. Due to the grouping of the particles in clusters, the sensitivity of the tape shows fluctuations over long distances as well as short ones [7]). This causes the signal to be multiplied by a factor which is a function of the point x on the tape. If we call this factor $1+\delta(x)$, it is found that $\delta(x)$ irregularly changes its polarity, even in an interval comparable with the smallest wavelength which can be registered. Still another way of expressing this would be to say that an interference signal $S(x)\cdot\delta(x)$, proportional to the signal, is superimposed on the signal $S(x)$. The noise caused by this is called the modulation noise. The ratio of signal/modulation noise is about 40 dB.

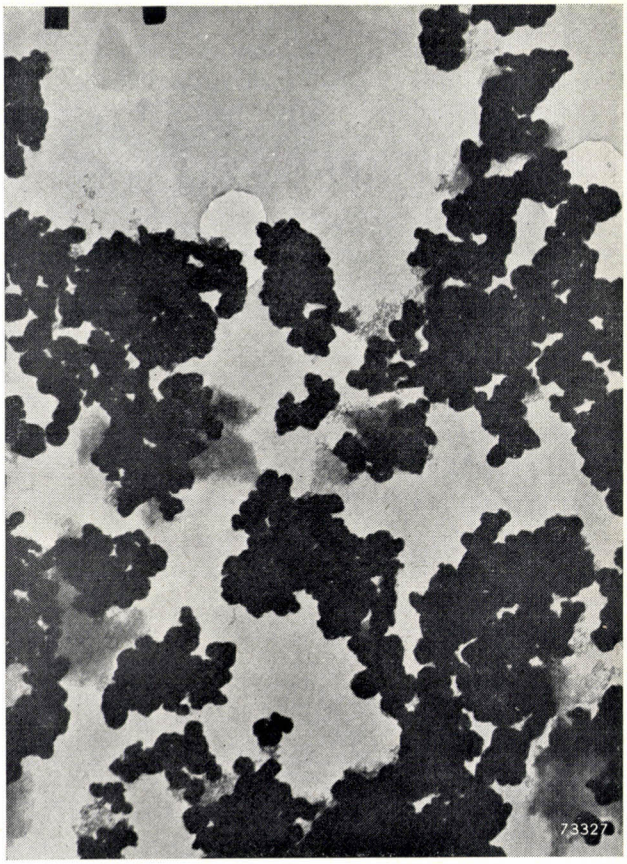


Fig. 19. Electron microscope picture of a thin layer of the coatings covering the magnetic tape. Enlargement 10,000 ×. The separate particles (size ≈ 0.1 μ) group together in clusters.

The measuring of the modulation noise is carried out by passing a direct current (in addition to the biasing current) through the recording head, equal

[7]) Fluctuations in the thickness of the magnetic layer, which may occur up to 1%, have the same effect.

in amplitude to the maximum admissible sinusoidal signal, and to measure the voltage induced in the reproducing head. Clearly a D.C. magnetization of the tape should always be carefully avoided, as the modulation noise of this D.C. magnetization would always be present. This magnetization may be caused, for example, by a permanently magnetized recording head, or a D.C. component of the biasing current. Even if the biasing current does not contain a D.C. component, it can cause D.C. magnetization, viz. when the positive and negative peaks differ in height. This is a result of the specific property of the tape to react only to the peaks of the magnetic field. It can now also be understood why the Poulsen method of using a D.C. biasing field may cause a higher noise level in the tape, especially in the weaker passages. This noise would not arise if the opposing field could exactly eliminate the D.C. magnetization; incomplete compensation, however, leaves a remanent magnetization in the tape. The advantage of the application of an HF biasing field is that, when no signal is present, the tape is automatically left in the demagnetized state.

Print-effect

If the magnetic material really behaved as suggested by the scheme given in fig. 3, only exterior fields of higher intensity than the critical value would be able to affect the recording. Such a strong field will not readily occur accidentally under ordinary circumstances. If appears, however, that weak fields, e.g. those originating from an adjoining layer in a rolled tape, may also influence the magnetization. Consequently, if parts of a tape carrying a very weak part of the recording happen to lie near parts where a sudden strong signal occurs, a part of the magnetization due to the latter signal may be transferred to the adjacent layer. In this way the strong signal may be weakly audible during reproduction once or even a number of times, before and after the real signal. This is the so-called "print-effect". The transfer of magnetic energy is here effected by the thermal agitation. For the reversal of the magnetization of a Weiss-area, a certain amount of energy E is required. Till now we have considered only the case when this energy is completely supplied by an exterior magnetic field. It can, however, also be supplied partly by the thermal agitation. The probability of a transition from one orientation to another as a result of the thermal agitation is given by the well-known expression $\exp(-\Delta E/kT)$, where ΔE represents the energy required for the transition, k the Boltzman constant and T the absolute temperature. When an equilibrium has set in, a weak field, in itself not strong enough to supply the energy for reversing the magnetization (i.e. roughly speaking, smaller than the local coercive force) may still cause a shifting of the equilibrium with the aid of the thermal agitation. After the interfering field has been removed the area does not resume its original state, at all events, not at once. It appears that the print-effect, as would be expected on this theory, strongly increases with increase of temperature, and is also dependent on the duration of the influence (it increases almost linearly with the logarithm of time, (*fig. 20*). When the field is removed it decreases

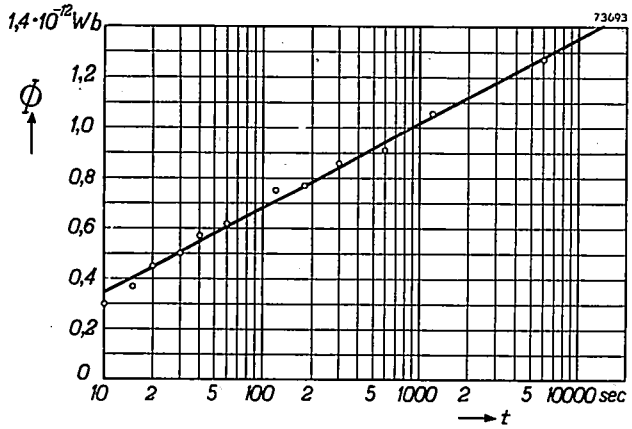


Fig. 20. Measured "print-effect" plotted as a function of the duration of the influence (log. time-scale). The graph shows the flux Φ, caused in a clean tape by contact with a tape at 30 °C having a saturated varying signal of wavelength 0.5 mm.

again, but it does not disappear completely, as there will be some particles in such a stable state that it is highly improbable that they will return to their original state by the thermal agitation.

It will be understood that stray AC fields can also take over the part played by the thermal agitation and may cause an increase of the print-effect. With the aid of a high-frequency A.C. field it is even possible to transfer the recording from one tape to another in contact with it. In this way contact copies can be made [8]) in a manner similar to photographic processes.

The print-effect is not equally strong for all wavelengths. The magnetizing field is in reality the field of the neighbouring layer, the distance between this and the magnetized layer being equal to the thickness of the tape. The intensity of this field decreases exponentially with the ratio of the tape thickness (Δ) to the wavelength (λ), but it is also proportional to the ratio d/λ, in which d represents the thickness of the magnetic layer, which is nor-

[8]) M. Camras, Electronics **22**, Dec. 1949, p. 78.

mally smaller than Δ (*fig. 21*). This is expressed in the formula

$$\mu_0 H = \frac{\pi d}{\lambda} \cdot B_0 \cdot \exp(-2\pi\Delta/\lambda),$$

where B_0 represents the average induction in the magnetized layer. The formula shows that the strength of the field has a maximum value when

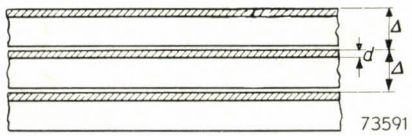


Fig. 21. Some layers of rolled tape. The figure shows the relative dimensions of the tape-thickness (Δ) and the thickness of the magnetic layer (d).

$\lambda = 2\pi\Delta$. When the thickness of the tape $\Delta = 55\ \mu$, this maximum is reached for $\lambda = 350\ \mu$, which with a tape speed of 0.76 m/sec. corresponds to a frequency $f = 2200$ c/s. This frequency is lower at smaller tape speeds, so that the audibility of the print-effect is then less.

Comparison of the magnetic process with other recording methods

In comparison with other sound recording and reproducing methods, the magnetic process has both its advantages and its disadvantages.

At the speeds used for professional purposes, i.e. at a tape speed of 30″/sec = 0.76 m/sec, the quality of the sound reproduced is superior to that of all other systems, thanks particularly to the favourable signal/noise ratio, which is perhaps equalled by that of a newly cut lacquer disc. The dynamic range of the gramophone records commercially available and of the optical systems (photographic and Philips-Miller processes) is certainly smaller.

Since, in optical processes, the number of electrons emitted by a photo-electric cathode is of the same order of magnitude as the number of magnetic particles passing the reproduction head per second in the magnetic process, it might be expected that the two systems would exhibit roughly the same dynamic range. However, the number of electrons emitted is not the only factor which influences the dynamic range in optical processes. The noise in optical systems is also dependent on a variety of other causes, e.g. the grain of the film and small irregularities in the film.

Another advantage of the magnetic process lies in the fact that both the negative and the positive magnetization contribute to the recording, which makes the "zero track", necessary in optical recording, superfluous. This eliminates the noise generated by a zero track, and at the same time, the necessity to suppress the zero track during weak passages ("noiseless reproduction").

Magnetic recordings are free from noise caused by dust particles, an additional source of noise found in all other systems. Dust particles are normally non-magnetic, so they do not contribute to noise in the magnetic process.

On the other hand, disturbing magnetic fields may unfavourably influence the magnetic recording, especially as regards the print-effect, which is not present in other recording processes. For some purposes the fact that the modulation is not visible may be a disadvantage. Nevertheless, the magnetic process is being used more and more in studios for broadcasting, and as an intermediate link in the gramophone and film industries.

The simple handling, during both recording and reproduction, is an important feature. This is especially important for the recording process, where, in all other systems, an expert technician is needed. The door is thus opened to a large number of amateur or semi-professional applications, at home, in the office, and for studying music or languages. The simplicity of the apparatus makes a compact construction possible, a very important feature for designing portable equipments.

A unique advantage of the magnetic process lies in the fact that a recording can be erased and a fresh one recorded at once. This has considerable economic advantages in some applications, and also

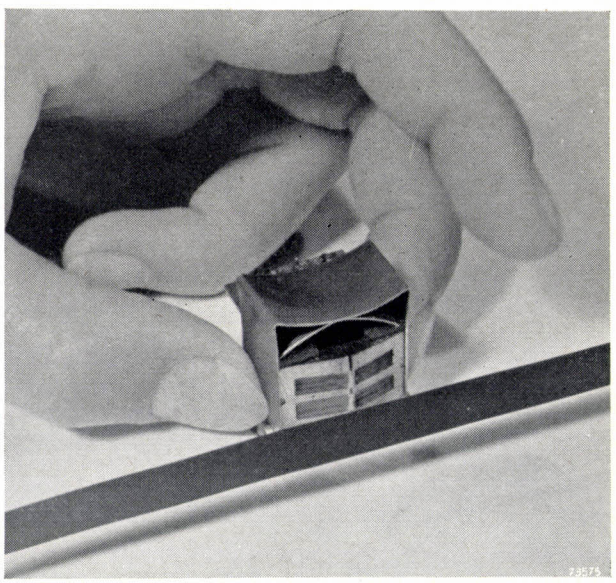


Fig. 22. Recording head for the stereophonic recording of sound on a magnetic tape.

opens the door to a number of quite new possibilities, e.g. the magnetic recording of data in the "slow-memory" of modern computing machines. In this way intermediate results may be stored until needed for further computations.

As a tape is used in the magnetic process, it is very suitable for stereophonic recordings. This is also the case with the optical and with the Philips-Miller systems, but it is more difficult to achieve with the gramophone record. The two sound tracks necessary for stereophonic reproduction can be recorded side by side by using a double recording head (*fig. 22*) and similarly reproduced later on.

Summary. The processes occurring during the magnetic recording and reproduction of sound on a ferromagnetic tape are discussed in detail with the help of a simplified representation of the hysteresis loop and of the magnetization process in the iron-oxide particles of the tape. Linearity between signal and magnetization is attained by the application of a high-frequency biasing field. The strength of the biasing field influences both the linearity of the signal and the frequency characteristic. Distortion and attenuation of the high frequencies during recording may also be caused by demagnetization processes. Deviation from the flat characteristic during reproduction occurs with long waves (low frequencies) as well as with short waves (high frequencies). In the latter case, the tape speed together with the length of the gap of the reproducing head determine the highest frequency which can be reproduced. When further discussing the total frequency characteristic, it is useful to introduce the concepts "transimpedance" and "transfactor". The noise of the tape is determined largely by the grain structure. An additional noise (modulation noise) sets in during magnetization, as a result of fluctuations in the thickness of the magnetic layer and the grouping of the magnetic particles into clusters. The transfer of magnetization to another tape is sometimes useful (for copying purposes), and at other times deleterious (print effect).

A comparison of the magnetic recording process with other recording processes shows that the quality of sound reproduction of the magnetic process is superior, in many respects, to others. Moreover, the specific qualities of this system (easy handling, erasure) favour its use in many applications.

METAL-DETECTORS

by E. BLASBERG and A. de GROOT. 621.318.4:621.317.18:669-493

In many manufacturing processes, the detection of metallic inclusions is of great importance. Examples are in the textile, paper and plastics industries, where small metal particles may damage the machinery or spoil the finished product. Undetected pieces of metal such as nails or parts of tools in timber or coal may have serious consequences in sawing or crushing operations. Many devices have been designed for the detection of metal objects in non-metallic materials. In the present equipment, intended specially for tracing extremely small particles of metal, the sensitivity has been increased to a point where it is possible to detect iron particles as small as 0.1 mg and non-ferrous particles down to about 0.4 mg.

During the manufacture or processing of non-metallic materials the intrusion of metal particles into the raw material or finished article is very difficult to avoid. In many cases this may be disastrous to the quality and good name of the article, apart from the damage that such metal particles can cause to the processing machines. It is often difficult or even impossible to see the foreign body and remove it in time, for example when it is buried beneath the surface of the material; the inspection, moreover, demands constant attention on the part of the personnel.

Efforts to eliminate the human element and to provide a more reliable check than is possible by mere *visual* inspection, have led to the development of special equipment — metal detectors — sensitive to metal particles, which gives a warning signal when such particles are encountered [1]). If desired, the particular machines can be stopped automatically, or that part of the material which contains the metal can be removed.

Many metal-detectors are intended only to indicate the presence of fairly large metal objects, such as nails or bullets in trees which are to be cut into planks, or small tools amongst coal being fed to a crushing machine. In other instances it may be desirable to detect extremely small particles of metal — for example, during the manufacture of gramophone records — even a very small metal particle in the plastic moulding materials can put a whole matrix out of action. In textile factories, too, a small metal object such as a pin or wire staple will often cause much damage to both machine and product.

A description is given in this article of two units which were specially designed for the inspection of non-metallic materials for the presence of small metal particles. The first of these is intended for the inspection of bulky materials in bales or on conveyor belts, packed foods and so on, whereas the second is more for the detection of iron particles in materials in the form of a long strip, such as paper or fabrics.

Metal-detector for bulk materials

Description of the detector

The principle on which this detector works may be seen in *fig. 1*. The material to be inspected moves on a conveyor belt A through a set of four coils L_1, L_2, L_3 and L_4. Of these, the so-called generator coils L_1 and L_2 are wound in the same direction and

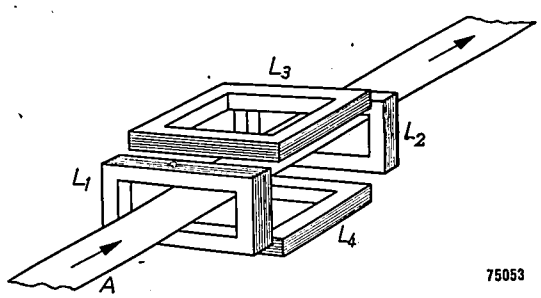


Fig. 1. Diagram showing arrangement of the coils in a metal-detector for bulk materials. The material to be inspected lies on the conveyor belt A. L_1 and L_2 generator coils. L_3 and L_4 pickup coils.

are connected in series; they carry an alternating current of 1400 c/s. *Fig. 2*, which depicts a cross-section of this coil system, also shows diagrammatically some of the lines of force of the magnetic field that occurs in and around the coils. As will

[1]) See G. S. Elphick, A. R. Woods and S. Y. Logan, Metal detection in industry, Symposium on electronics, Chapman and Hall, 124-139, 1949.
C. R. Schafer, Industrial Metal Detector Design, Electronics, **24**, 86-91, Nov. 1951.
M. Grobtuch and D. J. Williams, Pulp-Log Metal Detector, Electronics, **25**, 124-126, July 1952.
W. Zandra, Ein neues Metallmeldegerät, Elektrotechn. Z. **73**, 487-489, 1952.

be seen, coils L_3 and L_4, the "pickup" coils, include lines of force emanating from both L_1 and L_2, these lines cutting the coils in opposite directions. When the whole system is in equilibrium, no voltages are

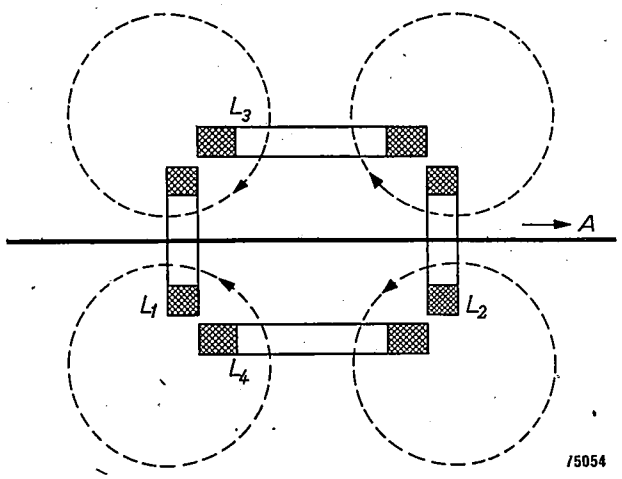


Fig. 2. Cross-section of coil system shown in fig. 1. The broken lines indicate the mean magnetic lines of force.

induced in L_3 or L_4, but, should there be a metal particle in the material under inspection passing through on the belt A, this will, if it enters coil L_1, modify the lines of force of this coil. The balance between the voltages induced in L_3 and L_4 by L_1 and L_2 is thus disturbed, and an A.C. voltage occurs in L_3 and L_4. Now, as these coils are wound in opposite directions and are connected in series, they will supply an A.C. voltage to the amplifier to which they are connected. The amplified voltage can be applied to a relay to operate a warning system or to switch off a motor.

An RC oscillator is employed as generator, this being a simpler way of producing a stable frequency and constant output voltage than by means of an oscillator with LC circuit. The oscillator is coupled to a push-pull amplifier with phase-inverter valve in accordance with the circuit dagram shown in *fig. 3*. The oscillator proper comprises valves T_1 and T_2, a network R_1-C_1-R_2-C_2 producing positive feed-back; negative feed-back is provided through R_3 and R_4. For this circuit to oscillate, the alternating anode voltage from T_2 must be in phase with the alternating grid voltage of T_1 and, as will be seen from a simple calculation, this takes place with an angular frequency given by:

$$\omega^2 = \frac{1}{R_1 R_2 C_1 C_2} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

When R_1 is of the same value as R_2, and C_1 the same as C_2, as is usually the case in RC oscillators:

$$\omega = \frac{1}{R_1 C_1} = \frac{1}{R_2 C_2} \quad . \quad . \quad . \quad . \quad (2)$$

The ratio of the A.C. anode voltage from T_2 to the A.C. grid voltage of T_1 is then 3 : 1, from which it follows that the amplification of the two valves need be only 3 for oscillation to occur. Without the negative feed-back already mentioned (R_3,R_4), the gain would be much greater, but it automatically adjusts itself to 3 in view of the fact that R_3 has a high negative temperature coefficient (NTC resistor [2]). When a current flows in this kind of resistor, the temperature rises and the resistance

[2]) See E. J. W. Verwey. P. W. Haayman and F. C. Romeyn, Semi-conductors with large negative temperature coefficient of resistance, Philips tech. Rev. 9, 239-248, 1947/1948.

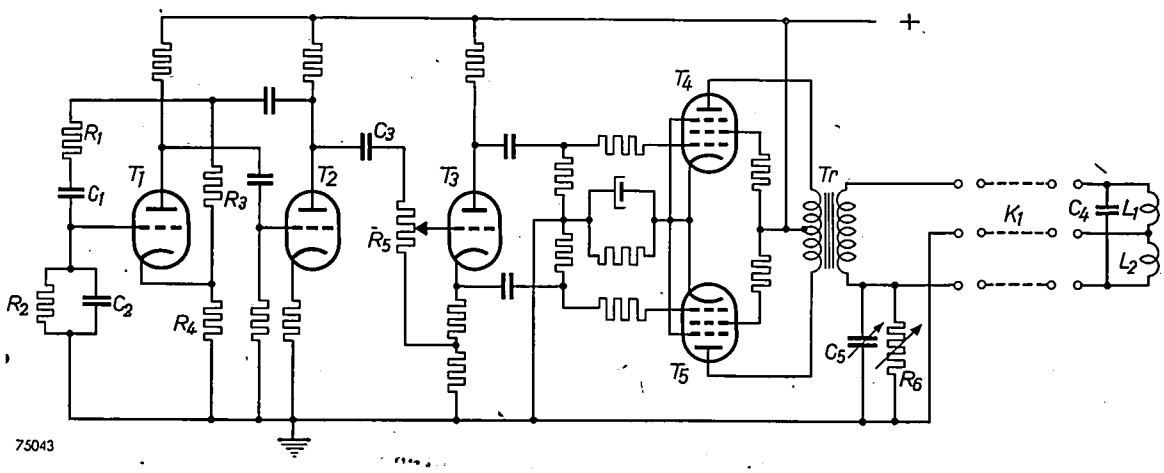


Fig. 3. Circuit diagram of the generator section of the metal-detector. Valves T_1 and T_2 constitute an RC oscillator. T_3 is an amplifier and phase-inverter valve; T_4 and T_5 are output valves. K_1 cable connecting the unit to the generator coils L_1 and L_2.

drops. The negative feed-back factor in the circuit therefore increases when the voltage across R_3 rises, and this will occur as long as the gain from T_1 and T_2 (allowing for the negative feedback) exceeds a factor of 3.

If the amplification is very high, so that a considerable negative feed-back is necessary, a voltage division will take place across R_3 and R_4 in a ratio that is practically the same as that of the voltage dividing network R_1-C_1-R_2-C_2 which, as already mentioned, is 3 : 1. The output voltage of the RC oscillator accordingly rises [3]) to a level where R_3 is almost equal to 2 R_4. For further details of this type of oscillator reference may be made to other publications [4]).

The voltage delivered by the oscillator is applied through a capacitor C_3 and volume control R_5, to the grid of T_3, from the anode and cathode of which two voltages are obtained, of opposite phase, which are in turn fed to the two push-pull valves T_4 and T_5. A cable K_1 is used to connect the output transformer of the oscillator to the two generator coils L_1 and L_2. Together with capacitor C_4, these coils form an oscillatory circuit tuned to a frequency of 1400 c/s. As it is a practical impossibility to make these coils absolutely identical, a variable resistor R_6 and variable capacitor C_5 are connected across one of them, enabling the current flowing in it to be adjusted so that the sum of the induced voltages in L_3 and L_4 (see figs. 1 and 2) will be exactly zero.

Fig. 4 shows the skeleton circuit diagram of the amplifier and detector section, to which the voltage induced in the pickup coils L_3 and L_4 is applied via a cable K_2. This section comprises a two-stage amplifier V_1 for amplification of the A.C. voltage induced in L_3 and L_4. Although only a voltage of fixed frequency is applied to the amplifier, it is essential that the amplifier should have a certain minimum selectivity, as the gain is fairly high, and this helps to reduce the noise originating in the initial stage of the amplifier. Again, it is undesirable to amplify harmonics of the oscillator frequency, for the following reasons. The A.C. voltage applied to the coils L_1 and L_2 by the oscillator is not purely sinusoidal. The "equilibrium", in which the sum of the induced voltages in L_3 and L_4 is zero, cannot be obtained simultaneously for voltages at the fundamental frequency and its harmonics. By designing the amplifier to be fairly selective, these two difficulties of the equilibrium adjustment are avoided. Two filters are used to give the required bandwidth.

The A.C.voltage delivered by the amplifier V_1 is rectified by a diode D and, when a metal object passes through coils L_1 and L_2, the voltage induced in L_3 and L_4 is modified, and the D.C. voltage supplied by the diode varies. The resultant voltage impulse is amplified by a two-stage amplifier V_2, the output voltage of which is applied to the grid of a thyratron Th; the phasing is such that when the A.C. voltage applied to the diode D rises, a positive voltage impulse occurs on the grid of Th, which then strikes.

The voltage from the diode D is simultaneously passed to the grid of an electronic tuning indicator TI, to facilitate adjustment of the equipment. This is done by adjusting resistor R_6 and capacitor C_5 in the oscillator section (fig. 3), so that the total voltage induced in the pickup coils is as small as possible; the fluorescent part of the tuning indicator is then at its minimum.

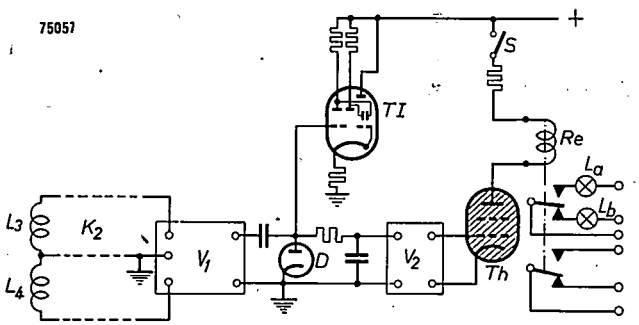


Fig. 4. Circuit diagram of the amplifier and detector section of the metal-detector. V_1 4-stage amplifier for the voltages at the generator frequency of 1400 c/s. The voltage rectified by the diode D, is amplified in a 2-stage amplifier V_2, being simultaneously applied to an electronic tuning indicator TI. A relay Re is included in the anode circuit of the thyratron Th. S switch for quenching the thyratron; L_a and L_b signal lamps.

The amplifier V_2 is not a D.C. amplifier: capacitors are used for coupling purposes. A slow drift in the equilibrium adjustment (minimum induced voltage in the pickup coils) is unavoidable; the causes of this may be temperature variations, or vibrations to which the apparatus may be subjected. If V_2 were a D.C. amplifier, it would respond to these slow changes, even in the absence of metal near the coils. However, as V_2 is an A.C. amplifier, only rapid changes in the voltages induced in L_3 and L_4 operate the warning system.

[3]) The NTC resistor is the limiting element in this circuit. If this method of limitation is not employed, the output voltage is usually limited by the occurrence of grid current or curvature in the valve characteristic. In this case distortion of the output voltage is much more pronounced.

[4]) M. G. Scroggie, Audio signal generator, Wireless World **40**, 294-297 and 331-334, 1949.
J. McG. Sowerby, Selective *RC* circuits, Wireless World **41**, 223-225, 1950.
D. J. H. Admiraal, *RC* oscillators, Electronic Applications Bulletin **12**, 111-131, 1951.

In order to prevent the condition of equilibrium from being disturbed too much by effects such as those described above, the unit should be set at regular intervals to give minimum deflection of the tuning indicator.

The sensitivity

The sensitivity of the equipment drops as the generator and pickup coils are made larger. It has been found in practice that with coils 16 sq. cm in area an iron particle 0.1 mg in weight can be detected. To illustrate the relationship between the voltage delivered by the amplifier and the particle size of the iron, *fig. 5* shows the A.C. output voltage

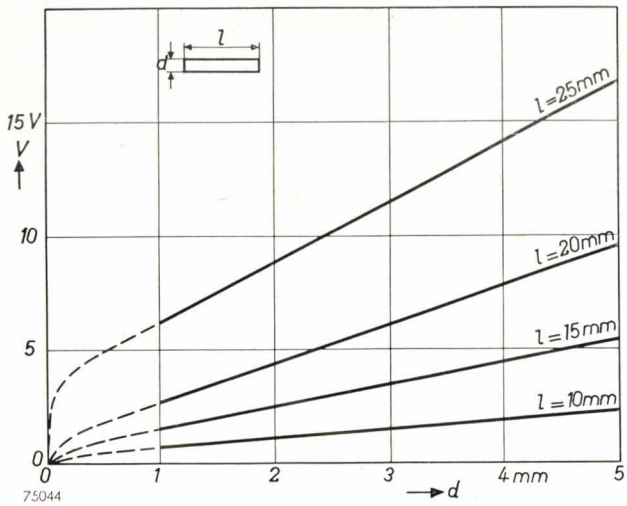


Fig. 5. Voltage V delivered by the amplifier V_1 (see fig. 4) as a function of the diameter d of a cylindrical iron particle of length l placed in one of the generator coils. Aperture of the coils: 10 cm × 10 cm.

plotted as a function of the thickness of a cylindrical iron particle placed coaxially in the coil, for different particle lengths. The coil area in this case was 100 cm². From this figure it will be seen that, provided the particle is not too small, the voltage obtained is a linear function of the particle thickness.

As the voltage induced in the pickup coils, for a given flux, is proportional to the frequency, the sensitivity will in the first instance increase with frequency, but here it is necessary to make a distinction between

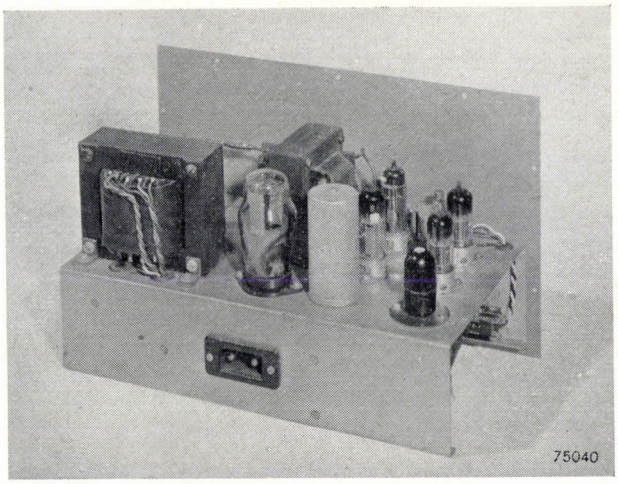


Fig. 6. View of the oscillator section of the metal-detector for bulk materials.

iron and non-ferromagnetic metals. As the frequency is increased, more and more eddy currents are produced in the foreign particles; the detection of non-ferromagnetic objects is based on this fact, so that a high frequency is advantageous for two reasons. On the other hand, with iron particles, which are detected by reason of their high permeability, such eddy currents are undesirable as they reduce the effective permeability, so that on this score, the frequency should not be too high. In a number of units a frequency of 1400 c/s has been employed, this being the frequency at which the results shown above were obtained. For most of the metals in general use, the specific resistivity of which is not so low that the occurrence of eddy currents is greatly hampered, e.g. copper, brass,

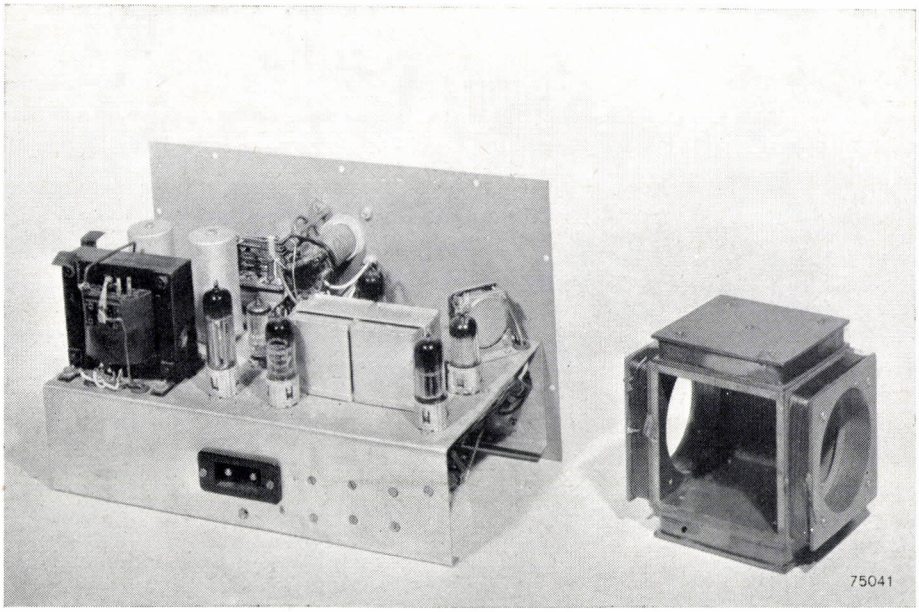


Fig. 7. The amplifier and detector section of the metal-detector for bulk materials. At right: coil system.

aluminium, the voltage induced in the pickup coils at a frequency of 1400 c/s is roughly one quarter of that produced by the presence of iron particles of the same size.

If the equipment is to be highly sensitive, every care must be taken to ensure a constant anode voltage for the valves, since small variations in this voltage will have the same effect as the appearance of a metal particle. On the other hand, if the metal objects to be detected are comparatively large, it will not usually be necessary to stabilize the supply, and the equipment is correspondingly simpler.

Fig. 6 depicts the oscillator section of the metal-detector, and *fig. 7* shows the amplifier together with a coil assembly.

A metal-detector for checking the materials used in the manufacture of gramophone records is seen in *fig. 8*. As a high sensitivity is needed for this purpose, coils having an aperture of only 16 sq. cm are employed.

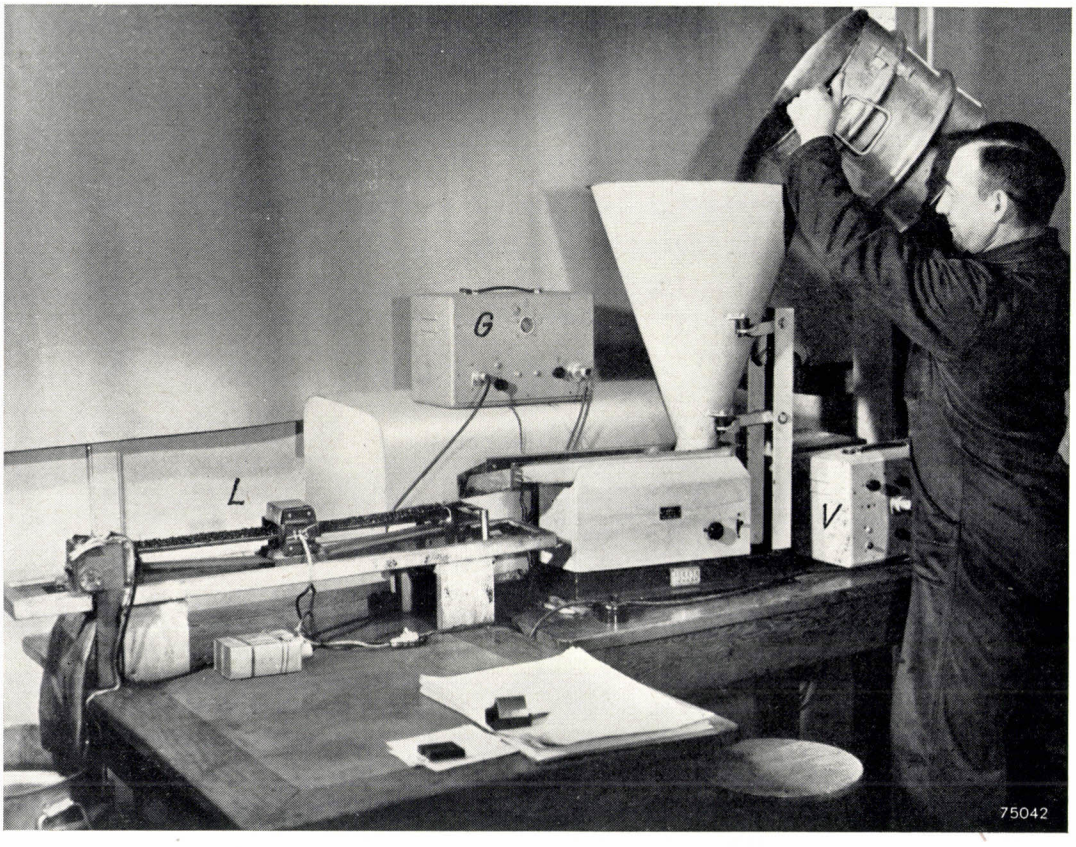


Fig. 8. Metal-detector employed for materials used in the manufacture of gramophone records. *L* coil system. *V* oscillator unit. *G* amplifier and detector unit.

Iron-detector for strip material

Description of the detector

In the textile industry the woven fabric is often subjected to processes intended to increase the gloss or to improve the texture. The cloth is fed through a number of smooth, hollow rolls (calenders), usually of steel, and heated. Calendering processes are also widely used in paper making and in the plastics industry. Should a metal particle become fixed to one of the rolls, it would make an impression in the material once every revolution, thus producing periodic flaws; alternatively, if *loose* particles find their way on to the surface of the rolls, they will become embedded in the material. Often, too, the rolls themselves may become damaged, involving re-surfacing, which is a costly operation.

Another process that many fabrics undergo is one of trimming, which consists of removing the fibres which project above the general surface. The fabric is passed over a roll, above which a fast-moving cylinder with helical cutters is mounted; these cutters travel along a straight blade and so remove all projecting fibres [5]). A single piece of metal between the cutters would cause very serious damage.

A metal-detector of the kind described above

[5]) The action is analogous to that of a lawn-mower.

would not be suitable for application to strip material of width say, one metre, as often encountered in the textile, paper and plastics industries. The use of coils of such dimensions that the material can pass through them, is not feasible. For safeguarding textile machines and the like, a metal-detector has therefore been specially designed, based on a different principle. As will be seen from the description that follows, this equipment reacts only to particles of ferro-magnetic metals, but, in view of the fact that most of the unwanted metal particles encountered in the machinery are of iron, the equipment offers a considerable degree of protection. Against this limitation may be set the advantage that this detector is simpler, cheaper and more robust than that designed for bulk materials.

The principle of this iron-detector may be explained in reference to *fig. 9*. A magnetic circuit is employed, comprising permanent magnets M, soft iron blocks B, and pole pieces P, also of soft iron. Coils L are mounted on the blocks B, and between the pole pieces there is a gap Sp which has a brass strip soldered into it to prevent the entry of dust etc. and to increase the robustness. The fabric is passed over the smooth upper faces of the pole pieces, which can be made in any length to suit the width of the material. The magnets are arranged at regular intervals along the pole pieces, the number depending on the material width.

If the fabric contains an iron particle it will, on passing the gap Sp, cause a variation in the magnetic flux, which in turn produces a voltage impulse in the coils[6]). These coils, which are in series with each other, are connected to an amplifier the circuit diagram of which is depicted in *fig. 10*. Two stages of amplification (valves T_1 and T_2) are followed by a thyratron with relay in the anode circuit for operating a warning system or safety device. When the thyratron has functioned it can be quenched by breaking the anode circuit. To simplify operation, the circuit may be so arranged that, when an iron particle appears, the detector is switched off simultaneously with the driving motor on the machine. Once the particle has been removed and the motor is re-started, the detector then automatically comes into action again.

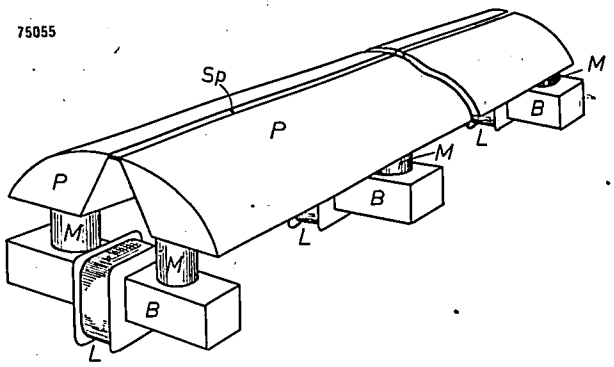


Fig. 9. Principle of the iron-detector for strip material. P pole pieces; M magnets; B soft iron blocks; L coils; Sp gap between pole pieces.

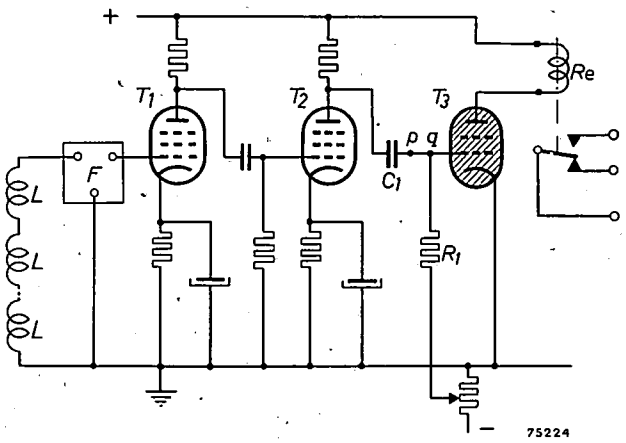


Fig. 10. Circuit of the amplifier section of the iron-detector for strip material. T_1 and T_2 amplifier valves. T_3 thyratron. F filter for suppression of interfering voltages. Re relay.

Since this unit is more sensitive to interference from external magnetic fields than the first detector described, a filter F is included between the coils and the first amplifier valve. The frequency of interfering fields is nearly always 50 c/s, whereas the voltage impulses produced by iron particles, at the speed at which fabrics usually move forward (10 to 20 metres/min), contain chiefly components of much lower frequency (5 to 25 c/s), so that a filter to suppress 50-cycle frequencies eliminates for all practical purposes the effects of interfering fields. Tests in a textile mill have shown that an amplifier without filter delivered an interference voltage of about 16 V, originating from neighbouring motors. The voltage produced by an iron particle was found to be 15 V. With filter fitted, the unwanted voltage proved to be only 1.2 V.

The suppression of interfering voltages can be improved considerably by shielding the filter with magnetic screens, thus preventing the stray fields from inducing voltages in the filter itself. In practice, however, it is found that, even without the screening, the ratio of the required voltage to the unwanted voltage is sufficiently high in almost every instance. In this connection, mention may be made of the interfering voltage encountered when the iron-detector was set up at various distances from a fully loaded $2^1/_2$ h.p. electric motor. At distances of 30, 50 and 100 cm the voltages delivered by the amplifier with filter were 8, 6 and 1 V respectively.

[6]) The action can be compared with that of the playback head of a tape-recorder on a very large scale, the remanent magnetism being in this case not in the tape, but in the head (and constant).

It was found necessary, however, when using the iron-detector on a machine, to shunt the motor-starter with a capacitor. If on this is not done, a voltage impulse may be produced each time the motor is switched on, such that the detector immediately switches it off again. This can be avoided in almost every case by using a capacitor of about 0.1 μF.

As mechanical vibrations in the coils also tend to initiate voltage impulses, it is advisable to mount the whole unit, or at any rate the pole pieces with the permanent magnets, on flexible supports.

The sensitivity

The magnitude of the voltage induced in the coils increases rapidly with the speed at which the iron particle passes the gap between the pole pieces, and also with the size of the particle. To illustrate the sensitivity, *fig. 11* shows the peak value of the voltage impulse as a function of the particle speed, when an iron particle 7 mg in weight moves at a distance of 3 mm across the gap. In *fig. 12* the weight of the iron particle is shown plotted against the distance at which it must travel across the gap to produce an impulse of 1.5 V at the output of T_2; in this case the speed of the particles was constant at 15 m/min.

In the arrangement of the complete appliance the valves T_1 and T_2, together with the filter and associated switch gear, the magnets, coil and pole pieces, are combined to form a single unit. The thyratron with relay and necessary switches are in a separate cabinet. The Philips electronic relay type GM 4801 serves the purpose well; the supply voltages for the amplifier can also be obtained from this relay. *Fig. 13* shows these two units, and *fig. 14* illustrates the manner in which the iron-detector can be mounted on a textile machine.

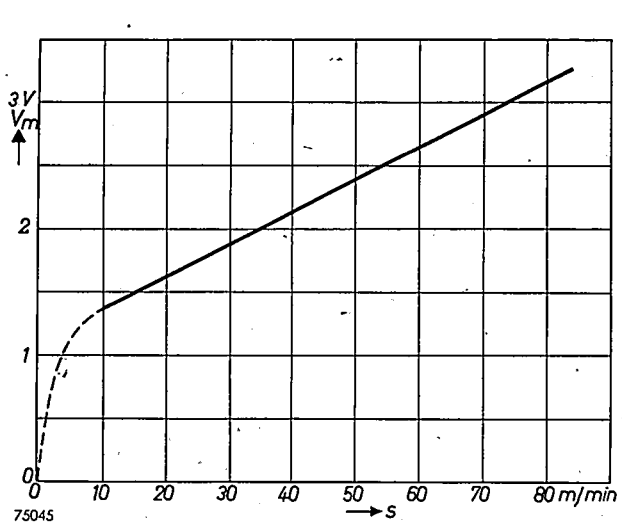


Fig. 11. Peak value V_m of the voltage impulse delivered by T_2 with an iron particle of 7 mg passing the gap between the pole pieces at a distance of 3 mm, plotted against the speed s of the particle. Width of gap: 5 mm.

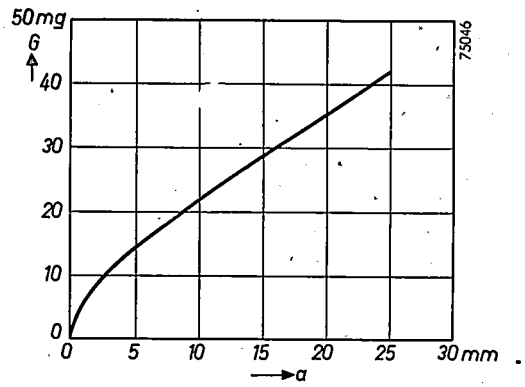


Fig. 12. Weight G of iron particles travelling at 15 m per minute which must move past the magnet gap to produce a voltage impulse of 1.5 V peak value at the output of valve T_2 (see fig. 10), as a function of the distance a from the gap. Width of gap 5 mm.

Other uses of the iron-detector

Apart from its use in protecting machines from the effects of iron particles, the detector described can also be employed to operate some form of warning, or to stop the machine, when any irregularities in the fabric or variation in the thickness occur. A soft iron wheel is made to run on the fabric adjacent to the gap in the pole pieces. As long as the thickness of the material is constant, the distance between the wheel and the gap will remain constant too, and no voltage will be induced in the coils. Any irregularity in the surface will vary this distance and thus initiate a voltage impulse.

The detector has many applications apart from the detection of iron as an impurity. When large iron objects are involved, the equipment described will operate at distances of one metre or more from the object; hence it can be employed for timing purposes in motor racing, or for safety devices or signalling in traffic, lifts or other transport systems. Another special application consists in the counting of a succession of objects, particularly when "optical" counting with a photo-cell is rendered difficult because of vapours or steam, as for example in canning factories. A counting system is then incorporated in the anode circuit of the thyratron, but in this case it is necessary to make a slight modification in the circuit (fig. 10).

When a counting system is employed, the thyratron must of course cut out immediately after it has once fired, and this can be effected by means of a relay which breaks the anode circuit after a slight delay [7]) when energized, and re-makes it again in short interval after it is released. This

[7]) This delay is essential because, when the iron object is not small, more than one voltage impulse will often be produced. It is then necessary to prevent the counting system from operating more than once.

involves a difficulty however, in view of the fact that a voltage impulse caused by an iron object passing over the pole pieces comprises components whose frequencies are very low. Hence the coupling capacitors in the amplification stages have to be fairly large. When the thyratron strikes, the ion current charges the capacitor C_1 up to a certain proportion of the running voltage and, when the thyratron is quenched by the interruption in the anode circuits, this capacitor must discharge across R_1. Owing to the high time constant of R_1C_1 (C_1 = 0.22 μF, R_1 = 2.2 MΩ), roughly 1 second must elapse after the thyratron has cut out before C_1 is more or less fully discharged. Now, if the anode circuit is closed again before this has taken place, the positive voltage still available on the grid would cause the thyratron to operate again, whether or not a positive impulse is applied to the grid. This would make reliable counting impossible,

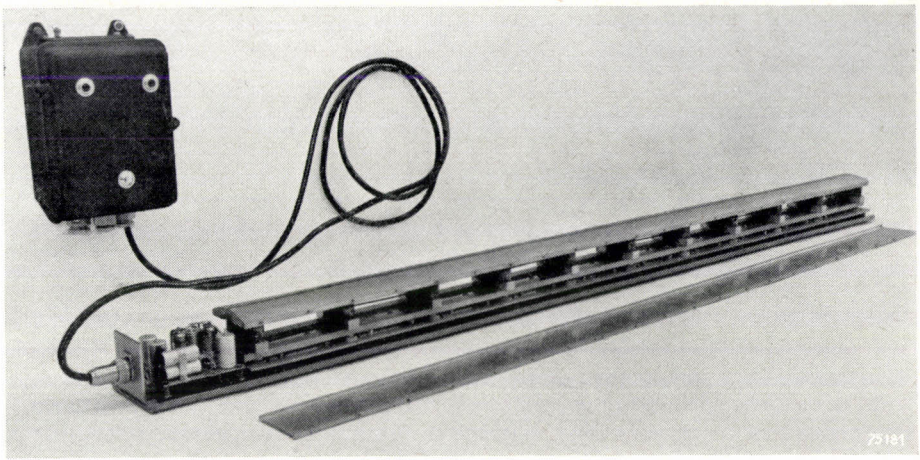


Fig. 13. An iron-detector for material of maximum width 2 metres. with amplifier cover and cover plates of the pole pieces removed. In the left hand top corner will be seen the Philips electronic relay, type GM 4801.

and, indeed, preclude all applications in which rapid operation of the thyratron is required. A simple solution lies in placing a crystal diode between the points p and q in fig. 10, its conduction direction being from p to q. Positive impulses thus have a clear path to the grid of the thyratron, but, in order to charge the capacitor C_1 after the thyratron has struck, a current must flow through the crystal diode in the direction in which it is least conductive, resulting in a fairly long charging time for C_1. When the thyratron is to operate rapidly, then, C_1 will be charged only to a fraction of the thyratron running voltage, so that, after quenching, this small charge has time to leak away before the thyratron anode circuit is re-closed.

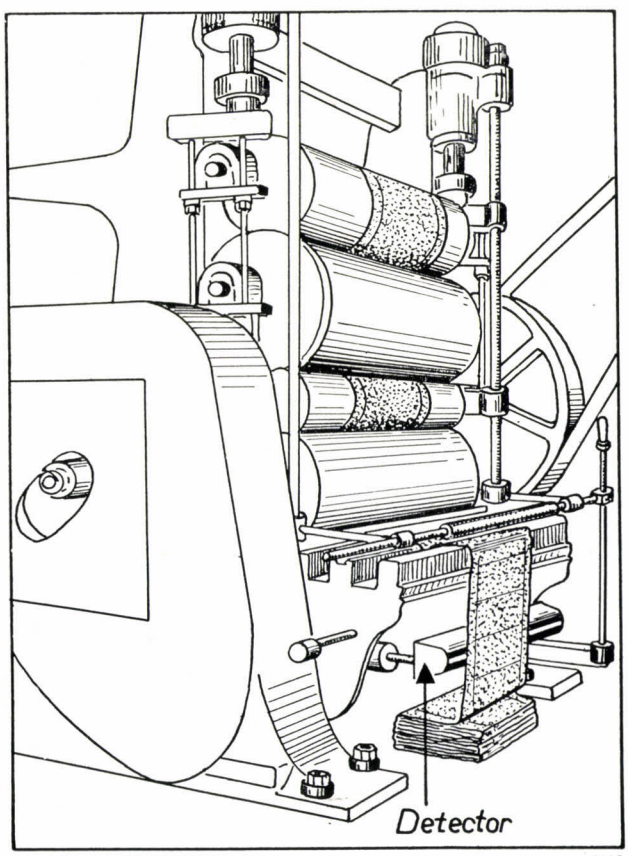


Fig. 14. Iron-detector fitted to a textile machine.

Summary. Two metal-detectors are described, which give warning of the presence of unwanted metal particles in non-metallic materials. The first of these was designed for detecting metal particles in bulk materials, the sensitivity of the unit being such that it will detect iron particles as small as 0.1 mg in weight; for copper, brass or aluminium the sensitivity is roughly one quarter of that for iron.

The second unit was developed for the detection of iron particles in strip materials such as textiles. Compared with the first-mentioned unit, this is of simpler and more robust construction, but it reacts only to ferro-magnetic metals. In addition to the protection of textile fabrics and machinery this detector can be employed for counting iron objects on a conveyor belt, where it sometimes has advantages over photo-electric counting. It can also be applied to other cases where iron objects are involved. e.g. for traffic signalling and for the timing of motor races.

VOL. 15 No. 4, pp. 105-132 OCTOBER 1953

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
THE PHILIPS INDUSTRIES

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, NETHERLANDS

IMPERFECTIONS IN MATTER

by G. W. RATHENAU *).

548.7

*An important stage is reached in a field of research when we are able to distinguish between phenomena inherent in an idealised system and others which are due to disturbances or imperfections in this system. The concept of an "ideal" or "perfect" state is prompted by a desire for simplicity which, it would appear, is not shared by nature. Evidently the dynamic of the human spirit is not based on the same "minimum principles" which govern the processes of nature. Prof. G. W. Rathenau, in his inaugural lecture delivered on 16th March 1953 at Amsterdam, elaborates on this theme and describes its application to the special field of the structure of solids. The main points of this inaugural address, supplemented with illustrations and a bibliography, are reproduced here with the kind consent and co-operation of Prof. Rathenau **).*

Introduction

Matter in the solid state plays an exceptionnally important part in our lives. The tools which man has evolved are witness to the unique qualities of solid matter, and the evolutionary ages of man bear the names of materials. It is an open question what name will be given by posterity to the present age, for which the press has already introduced the unfortunate title of atomic age. One thing is certain, however, that the technical and social significance of the discovery of nuclear energy largely derive from the fact that new fuels are now available for driving our machines.

As far as we ourselves are concerned, the human frame has all the characteristics of solid substance, even though opinions of the importance of this vary between the extremes of Epictetus, who regards the human being as a soul bearing the burden of a body, and of Nietzsche's view "Leib bin ich ganz und gar, und nichts ausserdem; und Seele ist nur ein Wort für ein Etwas am Leibe". (I am wholly body and nothing else; the soul is only another name for a part of the body).

Bravais, a hundred years ago, was the first to regard crystalline solids as consisting of atomic particles arranged at uniform intervals, to form periodic lattices in space. Not until 1912 did the work of von Laue and co-workers brilliantly confirm the main parts of this idea.

The last 40 years have shown that the application of theory, and particularly the quantum theory, to the model of the space lattice paves the way to an explanation of many of the properties of solids.

On the other hand, owing to mathematical difficulties and the lack of precise physical representations, the model of the space lattice has so far not been developed to the point where important properties such as supraconductivity and the magnetic saturation of metals can be satisfactorily explained.

It may, at first sight, seem surprising that so much of the theoretical and experimental work on the solid state in recent years has been concerned with the *imperfections* in the space lattice, while the model of the *perfect* space lattice is still in need of further development. The explanation of this apparently illogical development must be sought in the fact that small divergences in the periodicity have a pronounced effect on the characteristics of the substance. The situation is very analogous to that

*) Professor at Amsterdam University; previously with Philips Laboratories, Eindhoven.
**) The inaugural address has been published (in Dutch) by the North Holland Publishing Company, Amsterdam.

of a cobbled road with a number of stones missing after every thousand. From an aircraft, the road would appear to be perfect, but a motorist would consider it a very poor approximation to a road: from the point of view of driving its usefulness depends on the presence or absence of the defects.

The terms "periodicity deviation" and "lattice imperfection" would seem to imply an unfavourable appraisal not only of the matter itself, but also of this whole field of research. Caution is necessary in any such qualitative appraisal, however. In the history of scientific research it is often the case that divergences found experimentally from an idealised model have proved to be pointers to theoretically and technologically important developments. Furthermore, when imperfections appear to be *inevitable* and should therefore be referred to as the "rule" rather than as imperfections or defects, their investigation is clearly justified.

In fact, particularly in recent years, it has been shown that lattice imperfections are indeed inevitable.

The rôle of entropy in the occurrence of lattice imperfections

It is a well-known fact that the production of lattice defects in a crystal system involves an increase of internal energy. At the same time, it is possible to indicate two general causes of the inevitability of such defects. The first of these is related to a very universal premise which states that with rising temperature the macroscopic state of an atomic system is determined less and less by the requirement of lowest internal energy and more by the requirement of a large number of ways of obtaining this state. Stated in more general terms, this means that at elevated temperatures, nature does not tend towards order, but towards disorder. Hence well-ordered crystals undergo a change at the higher temperatures to the liquid or gaseous state, in which the atoms are less well arranged. The lattice imperfections, whose equilibrium concentration increases with temperature, tend to make the solid state resemble the molten or gaseous phase. From the statistical point of view the state implied by the word "disorder" could, in fact, be termed "order", just as an untidy writing desk, viewed from a higher level may, in fact, be considered as being more ordered than one on which all the papers are concentrated in one heap.

Let us take the case of a crystal consisting of atoms of one kind only. Atoms of one kind may be regarded as indistinguishable from each other, so that the crystal can be constructed in only one way. If one atom of another kind be introduced, however, the conditions are very different. A given macroscopic condition, e.g. a crystal with a foreign atom somewhere within it, can be constructed in just as many ways as there are possible positions of the foreign atom in the crystal. The same will apply mutatis mutandis to atomic vacancies in a crystal, and for all other local departures from uniformity.

Many solids are formed at such high temperatures that the number of realisation possibilities — expressed in terms of the entropy — will significantly influence the concentration of imperfections. It is not improbable that, for special applications, advances in technique may lead to lower temperature methods of preparation of solids, which can now be formed only at elevated temperatures. Often, however, it is found that the limit of chemical purity indicated by entropy has not been reached by a long way. In three dimensions, the analogue of our example of the cobbled road, viz, one impurity at 1000 atomic intervals, may, in fact, be called a model of perfection.

Interaction between physical lattice defects and departures from stoichiometry.

In a study of this kind it is an over-simplification to regard divergences from exact chemical composition, i.e. departures from stoichiometry, as independent of physical departures from uniformity, such as vacancies in the lattice. Often the internal energy of a crystal can be reduced by introducing a chemical defect together with a physical imperfection. A well-known example of this is given by Bradley and co-workers [1]) in the ordered nickel-aluminium alloy NiAl with a surplus of aluminium. From the point of view of minimum energy it is more advantageous to leave two places *vacant* in the nickel sub-lattice than to squeeze a large aluminium atom into one of the places belonging to the small nickel atom. The same argument explains why in ionic lattices such as that of ferrous oxide FeO, a surplus of oxygen is in fact obtained by the formation of unoccupied places in the lattice of the iron.

This last example may be used to show that the introduction of extra oxygen may have further consequences in addition to the occurrence of physical defects in the form of unoccupied places in the ionic lattice. In fact, as far as energy is concerned, it is most advantageous to make the entire crystal electrically neutral. This can be done by the formation of trivalent iron ions as additional lattice faults. One kind of lattice defect is then responsible for a whole chain of defects. Future

work will have to be awaited for an atomic scale-picture of the situation in ionic crystals containing more complex defects such as dislocations. The fact that the valency of the participating ions can be controlled by the introduction of suitable chemical additions in ionic lattices has been shown by Verwey and co-workers [2]). In this way the electrical conductivity and diffusion rate in the solid state, i.e. the reactivity, can be controlled. The problems concerned are particularly important from a technical point of view. Among others, Wagner [3]), Mott [4]) and Shockley [5]) have made contributions to their solution.

Mention has been made of a physical defect arising from a chemical imperfection; this is an extreme instance of interaction. More usually, physical and chemical lattice irregularities already present, attract or repel each other. Examples of this are found in the work arising from Snoek's research on interstitial solutions [6]); small atoms occupy intermediate lattice positions in crystals built up from larger atoms. It appears that these small atoms migrate to those places where there is a physical lattice defect which offers them more space. When the crystal is elastically stretched, they arrange themselves more or less in the direction of elongation (*fig. 1*). If permanent, additional lattice defects are produced as a result of plastic deformation, the interstitially dissolved atoms are captured within them, as shown by Dijkstra [7]). The material can accordingly be purified locally by heaping up the impurities at some other point.

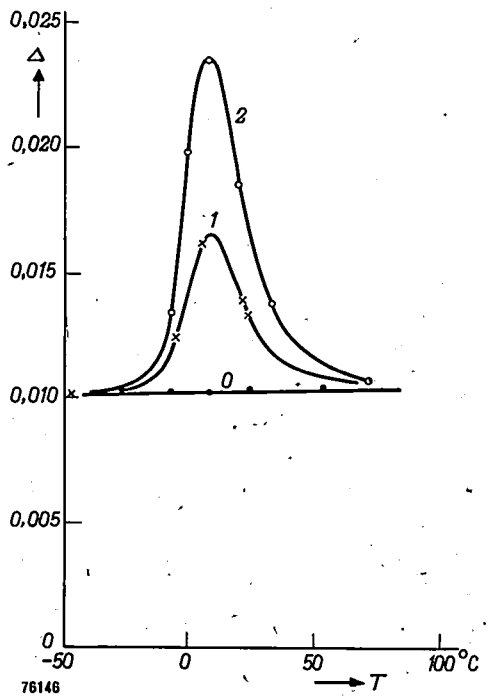


Fig. 1. Damping Δ of torsional vibrations in iron as a function of the temperature T. The damping is caused by the migration of dissolved nitrogen atoms to sites offering them the most space. These atoms thus arrange themselves more or less along the direction of tensile stresses, subject to a certain relaxation time. If the oscillation period is equal to 2π times this relaxation time, a relatively large amount of the vibrational energy is absorbed and dissipated as heat. A damping peak therefore results.
0: wire of pure iron
1: iron wire with dissolved nitrogen
2: as *1*, but with greater quantity of nitrogen
(From J. L. Snoek, Physica **8**, 711-733, 1941).

It is clear that research in this newly opened field, as undertaken for example by Zener [8]), Cottrell [9]) and Fast [10]), is only at a beginning. It is difficult to believe that the accommodation of dissolved atoms is determined only by the amount of space available for them. For that matter it is not at all certain in what state of ionization the interstitially dissolved atoms occur and whether this ionization does not in turn depend on the surrounding structure. It is possible that sensitive methods of measurement may yield further information on these questions.

Acceleration of structural changes by lattice imperfections

As a first reason for the inevitability of lattice irregularities, the entropy gain which they occasion has already been mentioned. When the crystal is at an elevated temperature, the internal energy will rise if irregularities occur, but the free energy — which decides what will really happen — will decrease.

The second reason for the occurrence of lattice defects is that, in general, nature by no means always selects a path for structural changes that leads *direct* to the greatest decrease in free energy of the system. Nature takes the *quickest* route compatible with the requirement of a drop in free energy in the whole system. In systems involving structural changes in the solid state, this speed is often achieved at the expense of the perfection of the lattice. One may be permitted to say that nature behaves as a mass producer, preferring a quick sale at low profit.

Crystal growth from solutions; nucleation as a rate-limiting factor

Frank and co-workers [11]) have concerned themselves with the growth of crystals especially from supersaturated solutions. Their investigations are of great importance as the treatment is straightforward and the confirmatory experiments unambiguous. Consider the case of a perfectly faced crystal in a moderately supersaturated solution,

The first atoms, which by adsorption could form an atomic layer on the perfect crystal face, parallel to it, are very liable to be re-dissolved, for the evaporation pressure of the nucleus of the new layer is very high as long as its volume is small compared with its surface. However, once a nucleus of a certain critical size has been accidentally formed — an improbable event in itself — it can grow rapidly and thus produce another perfect face on which, after an indefinite time, a fresh nucleus of critical size may be adsorbed.

In the example given, the repeated nucleation is the process that determines the speed of the reaction. It can be avoided by building the crystal on the principle of a single continuous surface, much as the surface of a screw thread or that of a spiral staircase can be regarded as constituting a single surface. If a crystal exhibits a *screw dislocation* of this kind, it can grow further in the direction of the axis of the spiral without requiring repeated nucleation from which fresh faces can grow (*fig. 2a* and *b*).

For crystal growth, therefore, what is needed is a dislocation such that the Burgers vector characterising the imperfection (so called after J. M. Burgers who has done much pioneering work on dislocations) includes a component in the direction of the normal to the plane of the lattice. In the absence of a dislocation of this kind, growth is practically at a standstill.

In the case in point, imperfections in periodicity are thus essential to growth. Whoever compares the experimentally found growth spirals — of which fresh examples are given almost daily —

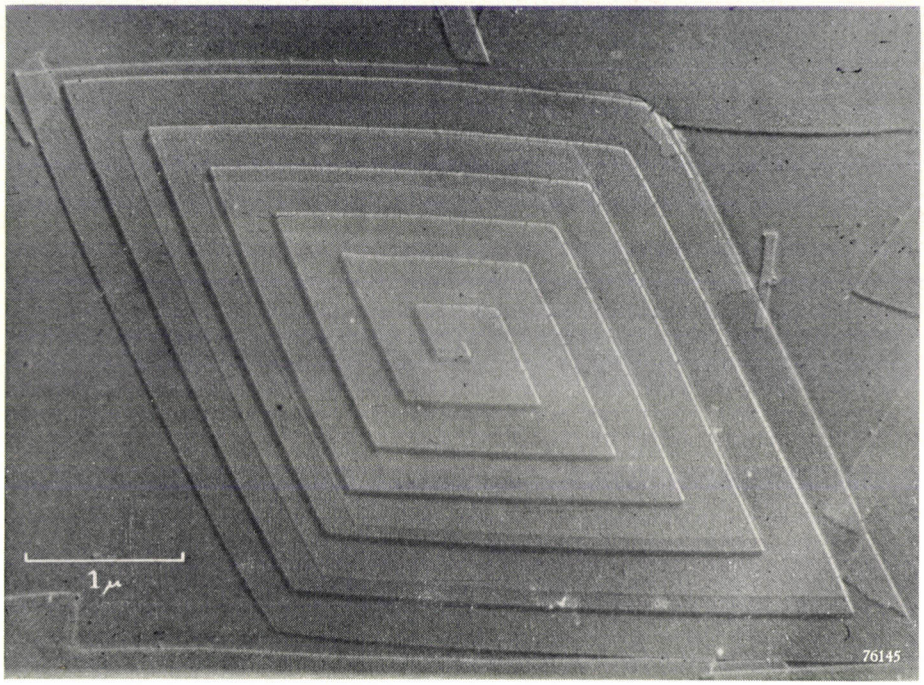


Fig. 2*a*) Growth of a single crystal of n-hectane in the direction of the axis of a screw dislocation. Electron-microscope photograph; specimen shadowed with palladium. (I. M. Dawson, Proc. Roy. Soc. **A 214**, 72-79, 1952, No. 1116).

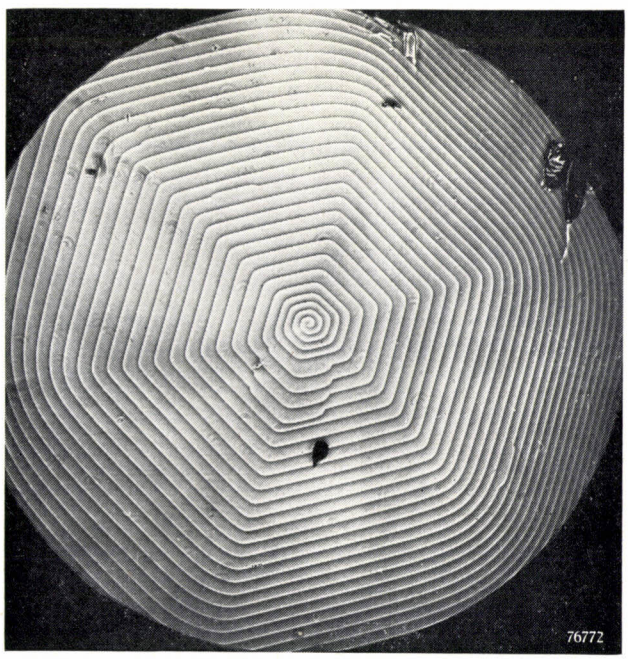


Fig. 2*b*). Growth spiral on a carborundum crystal. (A. K. Verma, Nature **168**, 430, 1951).

with the calculated forms cannot doubt the essential correctness of this statement. If the distance between the successive turns of the spiral along the axis is large, possibly owing to the presence of a large and rather complex unit cell in the structure, it may be expected, moreover, that that space round the axis of the spiral, which is under the most elastic strain, will be "channelled" and not filled with atoms [12]). Until now, examples of this phenomenon have been obtained experimentally only in inorganic substances. It would be surprising, however, if no use were made in living matter of this opportunity for the formation of very fine capillaries.

It might well be asked: to what should we attribute the screw dislocation that promotes the growth of a crystal? It is probable that this physical imperfection will, again, often arise as a result of variations in the chemical composition — corroborating the proverb that accidents never happen singly.

I should not wish to create the impression that the theory propounded by Frank is capable of explaining more than a small part of the phenomena of crystal growth from the liquid phase. Suffice it for me to say that it is at the moment not quite clear how the terms "supersaturation of a solution" and "perfect interfaces" should be interpreted when applied to crystal growth from a melt. It is not improbable that the imperfection in crystals formed from the melt (as revealed by X-ray diffraction) is to a certain extent also unavoidable.

Segregation; diffusion as a rate-limiting factor

In the growth of crystals from supersaturated solutions, of which we have just spoken, the nucleation is the rate-limiting factor. In order to grow *without* repeated nucleation — that is to say, rapidly — the crystal awaits the occurrence of lattice imperfections. In other transformations of solids there are rate-limiting factors of a different kind. A typical example is found in iron-carbon alloys, which form homogeneous crystals at elevated temperatures. At relatively low temperatures these crystals must, in order to attain equilibrium, separate to form two quite different kinds of crystal which also both differ from the original in structure and carbon concentration. Here, it is the diffusion of the carbon that determines the rate of transformation. In fact, the carbon appears almost only in one of the final phases.

In order to increase the speed of the reaction, nature reduces the path along which diffusion has to take place, viz. by forming the two new crystal types as lamellae in alternating layers (*fig. 3*). At the *boundaries* between two such lamellae, imperfections occur, simply owing to the fact that the two lattices do not mate. During the transformation, therefore, a part of the available free energy is transferred to energy at the boundary planes.

According to an estimate of Zener [13]), who has made a theoretical investigation of such transformations, about half the energy released during the transformation should be retained as boundary energy, in order to give the highest reaction rate. The lower the temperature at which the transformation takes place, the lower the rate of diffusion, but the greater the amount of free energy available for the formation of new boundary planes. As a result, a finer spacing of the lamellae occurs, a fact which has been known for some time from experiment. A theoretical minimum for the lamellae spacing, as indicated by the energy at the boundary plane and the heat of transformation would lie in the region of 200 Å.

Even though the principal factors are known, the transformations that take place in steel have yet to be studied very much more closely. It is not yet known, for example, whether and when a relationship exists between the orientation of the lattices of the two lamellate types of crystal.

Segregation; various rate-limiting factors

Whereas in the first-mentioned class of structural changes, nucleation is the rate-limiting factor, and in the second example the diffusion, it is to be expected that frequently one and the same substance will encounter different rate-limiting factors in different temperature ranges. Nature employs various devices to evade these.

Crystals supersaturated with a given element, as investigated by Guinier [14]) and Köster [15]), furnish eloquent examples of this. Below a certain temperature the solid solution is no longer stable. Just below this transition temperature there is so little free energy available that the surface energy essential for the formation of a germ crystal cannot be produced. In this temperature region, germ crystals occur as parasites on any lattice imperfections that may be present, which reduce the surface energy required and thus promote an acceleration of the process. In this way the lattice imperfections become permanently fixed. At temperatures well below the transition temperature, more free energy is available to reduce the supersaturation to zero, but diffusion now takes much longer, and precipitation from the supersaturated solution therefore occurs everywhere in the lattice

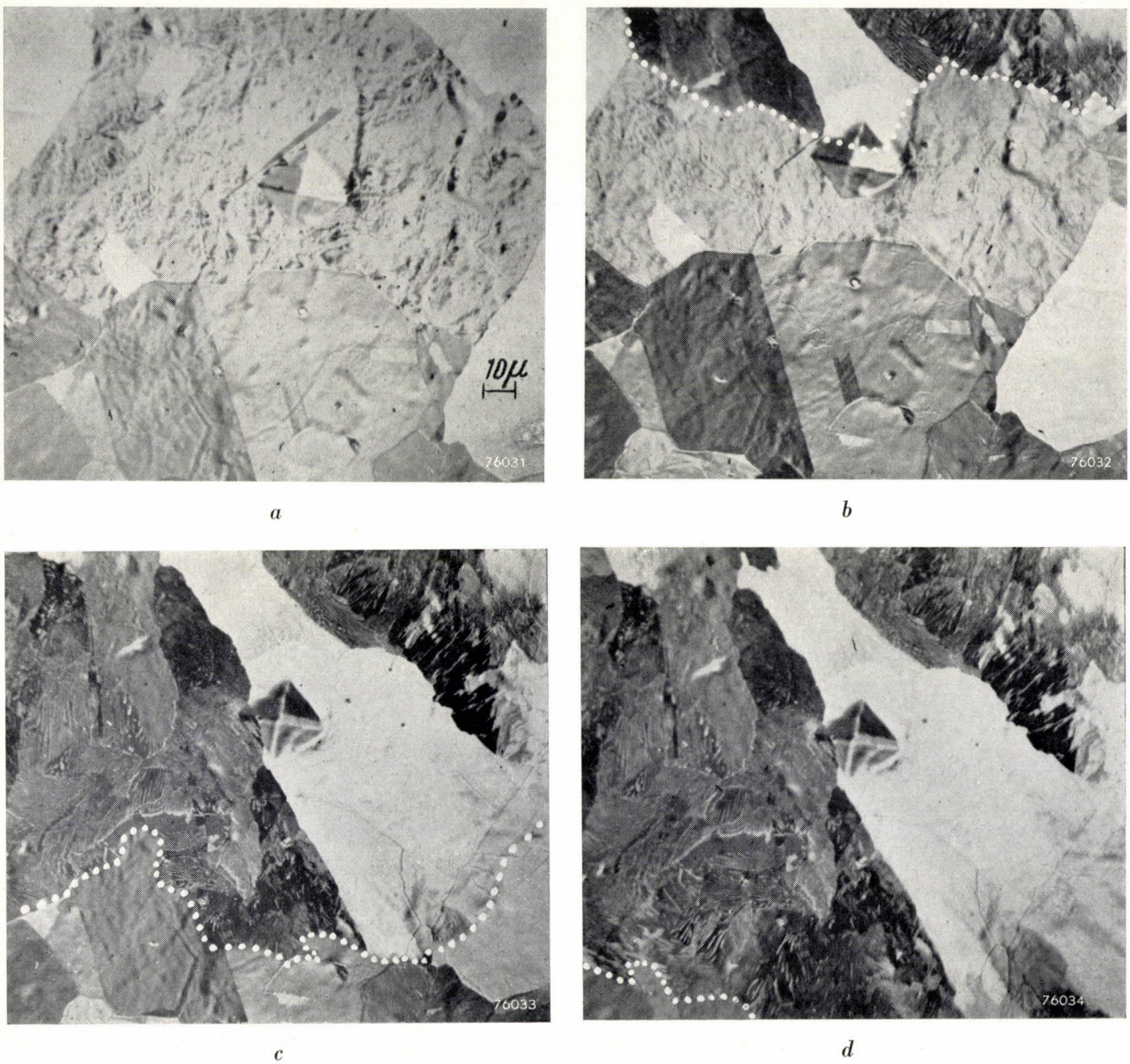
of the matrix. This involves the consumption of considerable energy for the formation of lattice irregularities. The precise physical picture of this process can only be conjectured; one experimental indication is that in this low temperature region the separation appears to be reversible. An increase in temperature results in solution.

Finally, at very low temperatures, the rate of diffusion is so small that concentration differences even over very small distances require a great deal of time to develop. Transformation then takes place as a sudden reaction, the whole structure of the lattice undergoing a change whilst retaining the original composition; this structure is not stable, but it is more stable than the original one in the supersaturated state. The dimensional and angular changes accompanying this instantaneous reaction lead to an upheaval in the lattice and a large accumulation of energy in the resulting lattice defects.

It is possibly relevant to note that the physical picture of the relationships between the orientation of the vanishing lattice and of the lattice in process of formation is still very incomplete.

Recrystallization

During the growth of crystals in the homogeneous



a *b*

c *d*

Fig. 3. Formation of pearlite during cooling of an iron-carbon solution from high temperatures. The pearlite forms in colonies of small lamellae. The progress of the boundary of the pearlite formation is shown by the dotted lines in *b* to *d*. Photograph taken with an emission electron microscope constructed at Eindhoven (See G. W. Rathenau and G. Baas, Physica **17**, 117-128, 1951). Magnification 500 ×.

solid phase it commonly occurs that energy for the formation of lattice imperfections is partly diverted to ensure that the disappearing lattice will vanish as quickly as possible. We know that when metals are plastically deformed, the disturbances in the lattice due to this will, when the temperature is raised, line up in continuous planes, thus forming boundaries between undistorted, submicroscopic zones, with very slight mutual disorientation among the lattices. The phenomena of recovery and recrystallization, on which W.G. Burgers and co-workers [16]) have done such important work, are apparently based on nothing else than the development of these zones into larger areas.

For many years C. S. Smith [17]) has stressed the fact that the motive force in the growth of crystals in the homogeneous solid phase is the interfacial tension at the boundary planes, just as in the growth in a system of soap bubbles. From this it follows that small crystals are no match for larger ones, but that crystals of which the lattice shows only slight disorientation have little tendency to absorb each other. This is because the surface tension is very small when the disorientation angle approaches zero (*fig. 4*), a fact which is quite plausible and

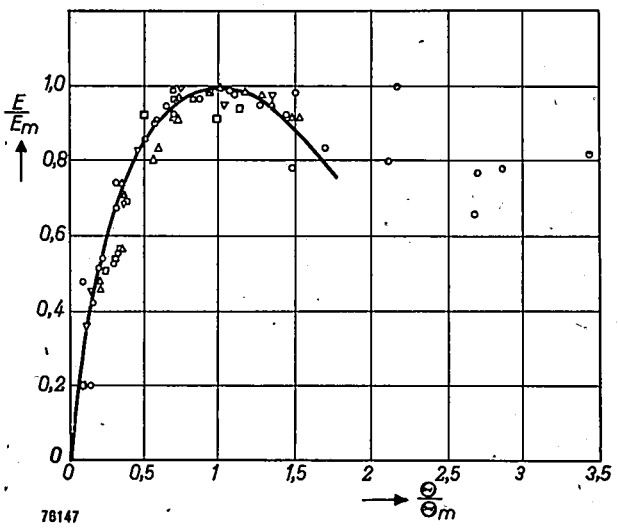


Fig. 4. Interfacial tension E (energy per unit area) of two crystals disoriented by an angle Θ with respect to each other. The various measurement points refer to three different metals. (W. Shockley, L'Etat solide, Inst. Internat. de Phys. Solvay, Brussels 1951, p. 447.) If Θ_m be the angle at which E is at a maximum ($=E_m$), E/E_m appears to be (to a fair approximation) the same function of Θ/Θ_m, independent of the metal.

which is quantitatively demonstrated by the calculations of Read and Shockley [18]) and others, and by the experiments of Chalmers, Dunn and others. The fact that crystals with roughly similar orientation are not in a position to react with each other quickly, means that they can grow independently of each other, and that lattice disturbances become permanent at the common boundaries (even though they may not be very numerous per square cm).

Order-disorder transitions

Lastly, we must not overlook the phenomenon of ordering in systems containing more than one kind of atom, say two kinds. Below the so-called transition temperature, a condition is approached in which each atom surrounds itself uniformly with as many as possible of the other kind of atom. In recent years the theory of order-disorder transitions has made great progress by taking into account the interaction of not only nearest atoms, but also their next-nearest neighbours. It appears, for example, from the work of Linde [19]), Mac Gillavry [20]) and Guinier [21]) that in the transition from the state of disorder to that in which order is maintained even over large distances from an atom, contiguous well-ordered zones are formed, the order being disturbed at the boundaries. In the lattice of one zone the atoms are thus disposed in a different manner from that of the atoms in the lattices of neighbouring zones. It again appears as though imperfections at the boundaries which are difficult to eradicate — the so-called stacking faults — must be accepted in order to obtain a reduction in the free energy within the whole space as quickly as possible. This raises the interesting problem that such faults appear to occur only on certain kinds of lattice plane and that this happens at regular distances.

In this connection it is worth while recalling one of Guinier's observations; in an alloy of one atom of aluminium and one of silver he found an ordering of the atoms within short distances, i.e. in very small zones, even though long-distance ordering and a transition temperature have not been found. It may well be that such short-distance ordering occurs in many systems and has some influence on the physical, electrical and also the very susceptible magnetic properties.

The nature of physical lattice imperfections

In the foregoing we have discussed at some length the fact that, for the sake of an increase in entropy and high reaction rates, lattice imperfections occur as a necessary — but none the less interesting — "evil". We have not yet attempted, however, to give any physical picture of the lattice defects.

It is beyond dispute that screw dislocations are

a physical reality in the growth of crystals from solutions, and there can also be no doubt as to the accuracy of the picture given in 1939 by J. M. Burgers [22]) of the faults that occur at the crystal boundaries. Means were found to "develop" such faults for examination under light and electron microscopes by making them function as points of nucleation during the formation of new phases [23]) (*fig. 5*). The spacing of the faults thus rendered visible correspond to that anticipated from the known disorientation of adjacent crystals.

experience of physical deformation, electrical conductivity and so on, date back a long time. The theory is already in an advanced state of development. Without going into atomic data such as the cohesion between the atoms in a lattice, we have, by geometrical means, constructed models of possible lattice imperfections. In this way the statics of lattice imperfections [25]) and, recently, also their dynamics, have been developed to a high degree and adapted to all the empirically known phenomena. For example, the important new in-

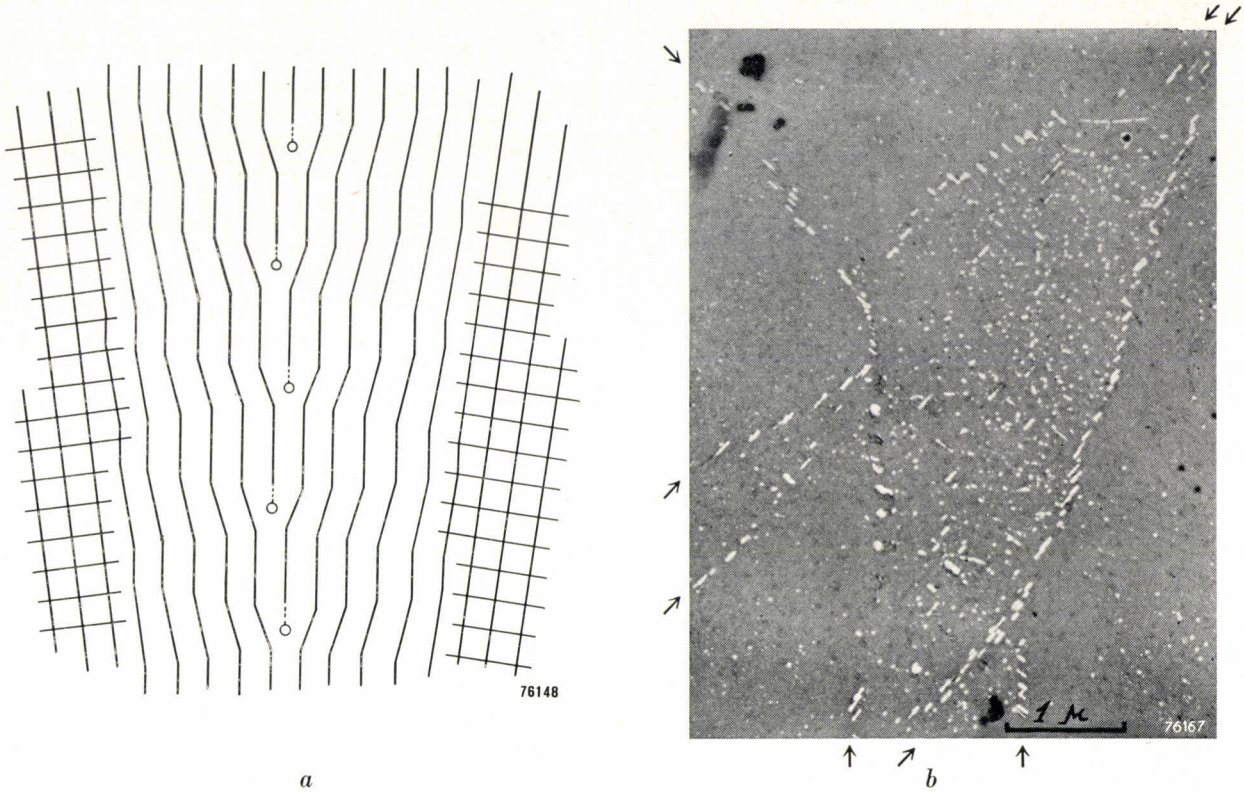
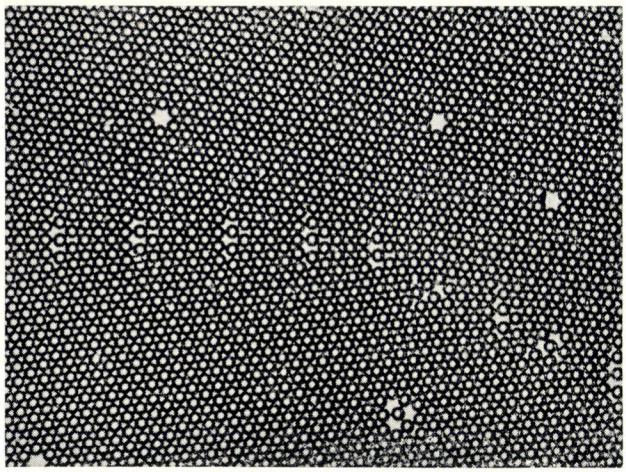


Fig. 5*a*) Geometrical representation of dislocations at the boundaries between two crystals. (Vide J. M. Burgers, Kon. Ned. Akad. Amst. **42**, 293, 1939.)
b) "Developed" dislocations of the kind depicted in a). The distances are as anticipated from the observed disorientation angle of the adjacent crystals. The arrows indicate the crystal boundaries. (R. Castaing and A. Guinier, Comptes rendus **228**, 2033-2035, 1945.)

Druyvesteyn and co-workers [24]) have for some years concerned themselves with the deformation of metals at low temperatures, and their work makes it probable that, in this case, conforming with the ideas of Seitz, the diffusion of atomic vacancies is in part responsible for the recovery of the lattice after deformation.

In all truth it must be admitted that other kinds of lattice imperfection exist which have never been demonstrated in a fundamental manner, but have been deduced indirectly from various experiments. In physics, where the fundamental experiment generally precedes the theory and the latter precedes the practical application, this is a rather unusual situation. Technological applications based on formation regarding diffusion, the so-called "Kirkendall effect", has now been added to the applications of the theory of lattice imperfections. Experimentally, however, the fundamental work has so far consisted only of more or less happily chosen models. There is, for example, the well known two-dimensional model of a disturbed lattice devised by Bragg and co-workers [26]) consisting of a "raft" of uniform soap bubbles (*fig. 6*). However, as recently demonstrated by Shockley, it can be very dangerous to derive quantitative data from a model of this kind.

The state of affairs in the field of physical lattice imperfections may be compared with that in the phenomenon of disorder in aggregates of magnetic dipoles, now familiar in ferromagnetism. The latter



76168

Fig. 6. Two-dimensional model of an imperfect lattice obtained by drifting uniform soap bubbles against each other (Bragg's bubble method). In this example two "crystals" are seen to have formed with a certain disorientation angle. (W. M. Lomer and J. F. Nye, Proc. Roy. Soc. **A 212**, 567-584, 1952.)

presents a much more favourable picture, partly because many of the phenomena are produced on a very much larger scale. The detailed theory of the geometry of Weiss domains, as given by Néel [27]) has been admirably confirmed by experiment. It is to be hoped that in the years to come suitable experiments will enable the present theory of physical lattice imperfections to be verified and applied to the atomic forces at work in a given system.

Regarded superficially, the human body is a beautifully symmetrical structure; yet every surgeon will be able to point out that the internal structure reveals a lack of symmetry that is essential for proper functioning. He would have nothing but admiration for the logic by which these departures from uniformity can be explained. Although, as is usually the case, the analogy is by no means perfect, the physicist, too, studying the imperfect lattice and searching for relationships between the imperfections and many important properties of solids, will be fascinated by the logic of the unfolding pattern and its far-reaching consequences.

Summary. Whereas many of the properties of solids can be explained theoretically on the basis of a model of the periodic space lattice, certain properties and also many phenomena occurring in crystals depend on small departures from the periodicity. The occurrence of such lattice imperfections, in spite of the additional energy required, is shown to be correlated with the increase in the probability of a given state (increase in entropy), or, by the evasion of rate-limiting factors, with an acceleration of changes in structure that are necessary, i.e. which lead to a condition of lower final energy. The first instance is illustrated by a consideration of the interaction between physical and chemical lattice imperfections, for example, in the case of controlled valency substances and in interstitial solutions. The second instance is encountered in phenomena such as the growth of crystals from solutions, segregation, recrystallization and order-disorder transitions, where the processes are in principle determined by the atomic interaction as manifested in nucleation, diffusion, and interfacial tension. Some notes are included on the nature of the lattice imperfections. Although some kinds of imperfection such as screw dislocations, boundary faults, and atomic vacancies are well-known and much theoretical work has already been done on lattice defects, further fundamental work as to the character of all kinds of lattice faults is still required.

BIBLIOGRAPHY

[1]) A. J. Bradley and A. Taylor, Proc. Roy. Soc. **A 159,** 56, 1939.
A. J. Bradley and G. C. Seager, J. Inst. Metals **64,** 1, 1939.

[2]) E. J. W. Verwey, P. W. Haaijman, F. C. Romeijn, G. W. van Oosterhout, Philips Res. Rep. **5,** 173, 1950.
E. J. W. Verwey, P. W. Haaijman and F. C. Romeijn, Philips techn. Rev. **9,** 239, 1947-48.

[3]) C. Wagner, Z. phys. Chem. B **21,** 25, 1933; **32,** 447, 1936; **38,** 325, 1937.

[4]) N. Cabrera and N. F. Mott, Rep. Progr. in Phys. **12,** 163, 1949.

[5]) W. Shockley, Electrons and Holes in Semiconductors, van Nostrand, New York 1950.

[6]) J. L. Snoek, Physica **6,** 591, 1939; **8,** 711, 1941; **9,** 862, 1942.
D. Polder, Philips Res. Rep. **1,** 1, 1945.
L. J. Dijkstra, Philips Res. Rep. **2,** 357, 1947.

[7]) L. J. Dijkstra, Trans. A.I.M.E. **185,** 252, 1949.

[8]) C. Zener, Elasticity and Anelasticity of Metals, Univ. Chicago Press, Chicago 1948.

[9]) A. H. Cottrell, Report Conf. on Strength of Solids, Physical Soc. London 1948, p. 30.
A. H. Cottrell and B. A. Bilby, Proc. Phys. Soc. A **62,** 49, 1949.

[10]) J. D. Fast and L. J. Dijkstra, Philips tech. Rev. **13,** 172, 1951.
J. D. Fast and J. L. Meijering, Philips Res. Rep. **8,** 1, 1953.

[11]) W. K. Burton, N. Cabrera and F. C. Frank, Nature **163,** 398, 1949, and Phil. Trans. Roy. Soc. A **243,** 299, 1951.

[12]) F. C. Frank, Acta Cryst. **4,** 497, 1951.

[13]) C. Zener, Trans. A.I.M.E. **167,** 550, 1946.

[14]) A. Guinier, Acta Cryst. **5,** 121, 1952.

[15]) W. Köster, L'Etat Solide, Brussels 1952, p. 235.
W. Köster and others, Z. Metallk. **43,** 193, 202, 208, 213, 1952.

[16]) W. G. Burgers, Kon. Ned. Akad. Wet. **50,** 452, 595, 719, 858, 1947.
T. J. Tiedema, W. May and W. G. Burgers, Acta Cryst. **2,** 151, 1949.
W. G. Burgers and T. J. Tiedema, Kon. Ned. Akad. Wet. **53,** 1525, 1950.
T. J. Tiedema, Kon. Ned. Akad. Wet. **53,** 1422, 1950.

[17]) C. S. Smith, Trans. A.I.M.E. **175,** 15, 1948; Trans. Amer. Soc. Metals **45,** 533, 1953.

[18]) W. T. Read and W. Shockley, Phys. Rev. **78,** 275, 1950.

[19]) C. H. Johannson and J. O. Linde, Ann. Phys. **52,** 1, 1936.

[20]) C. H. MacGillavry and B. Strijk, Physica **11,** 369, 1946, **12,** 129, 1946.

[21]) A. Guinier and R. Griffoul, Revue Métallurgie **45,** 387, 1948.

[22]) J. M. Burgers, Kon. Ned. Akad. Wet. **42,** 293, 1939; Proc. Phys. Soc. **52,** 23, 1940.

[23]) R. Castaing, C. R. Acad. Sci. Paris **228,** 1341, 1949.

[24]) J. Molenaar and W. H. Aarts, Nature **166,** 690, 1950.
J. A. Mainntveld, Nature **169,** 623, 1952.

[25]) F. R. N. Nabarro, Advances in Physics **1,** 271, 1952.

[26]) W. L. Bragg and J. F. Nye, Proc. Roy. Soc. A **190,** 474, 1947.

[27]) L. Néel, Cahiers phys. **25,** 21, 1944; J. Phys. Radium **5,** 241, 1944.

General: Imperfections in nearly perfect crystals, John Wiley, New York 1952.
L'Etat Solide, Solvay Congress, Brussels 1952.

PREPARATION AND CASTING OF METALS AND ALLOYS UNDER HIGH VACUUM

by J. D. FAST, A. I. LUTEIJN and E. OVERBOSCH. 669.04:621.365.52

This article describes a method whereby metals and alloys of extreme purity and precise composition can be produced in a form that lends itself well to further shaping into test specimens or components. The high degree of purity is demonstrated by chemical analysis and in certain instances also by the striking physical properties exhibited by such materials.

In the fundamental investigation of metals and alloys, as well as in their processing and use, many of the difficulties encountered are attributable to impurities which are always present in commercial grade metals. Particularly the so-called structure-sensitive properties of metals are affected by the nature, quantity and distribution of the impurities. Among these properties are mechanical properties such as tensile strength and ductility, and magnetic properties such as permeability, coercive force and hysteresis losses. In order to reveal the optimum values that these properties *can* attain we must be able to produce the metals and alloys in a very pure state and also to add any desired percentage of impurities [1]). By working with alloys of precisely known composition it is also possible to obtain information regarding the very practical questions as to those impurities which are more particularly to be avoided and the probable effect of any divergence from specified percentages of the main constituents. The effect of the latter, when investigated in commercial grade alloys, is often completely masked by the effect of fluctuations in the percentages of impurities.

A few years ago there appeared in this Review a description of an apparatus by means of which metals and alloys of precisely known composition can be prepared by melting them under vacuum, using induction heating [2]). The present article refers to improved equipment which enables an alloy, when once prepared, to be cast immediately in a mould while still under high vacuum. There were two important reasons for designing equipment along these lines, both of which are related to the further processing of the metal. Whether the final product is to be a component or a test specimen, some further shaping operation is usually necessary, — for example forging, rolling, swaging or drawing into wire. During these processes uncontrolled quantities of impurities may find their way into the metal. If the mechanical properties are such that the alloy can be worked at room temperature, the risk of the intrusion of impurities is only slight; the impurities are then confined to a very thin outer layer which can be removed either mechanically or chemically. On the other hand, if the mechanical properties of the alloy necessitate its working at elevated temperatures, impurities are able to penetrate much more deeply. Especially to be feared is the intrusion of elements having small, rapidly diffusing atoms, such as hydrogen, carbon or nitrogen. Particularly with alloys which do not lend themselves well to shaping operations it is therefore important that they should be prepared by casting under vacuum in a shape that will necessitate a minimum of working to reach the ultimate product. The other advantage of casting is that there is a smaller risk of cavities or blow holes than obtains when the melt is cooled in the crucible. Such cavities can be caused by the considerable reduction in volume that usually accompanies the transition from the liquid to the solid state: thus, in the gradual cooling in a crucible, when the interior of the melt remains liquid for some time after the outside has solidified, cavities form as the interior solidifies. By casting in a suitably designed mould, this can generally be avoided, as will be seen presently.

With the experience gained with both the earlier and the present equipment it is now possible to

[1]) Ordinarily the word "impurity" is held to mean a deleterious constituent, but in this article we shall take it in the wider significance of "a small percentage of alloying constituent", without, however, specifying a sharply defined limit for this percentage. This wider definition has been adopted because it is not a foregone conclusion that all foreign elements are deleterious and because instances occur in which one and the same "impurity" will have an adverse effect on certain properties of an alloy and a beneficial effect on others. — It should also be noted that for the sake of brevity the words metal and alloy are employed in many places as synonyms. This is not unjustified for, strictly speaking, all metals are alloys, as no metals exist in a perfectly pure state.

[2]) J. D. Fast, Apparatus for preparation of metals with an exactly known content of impurities, Philips tech. Rev. 11, 241-244, 1950.

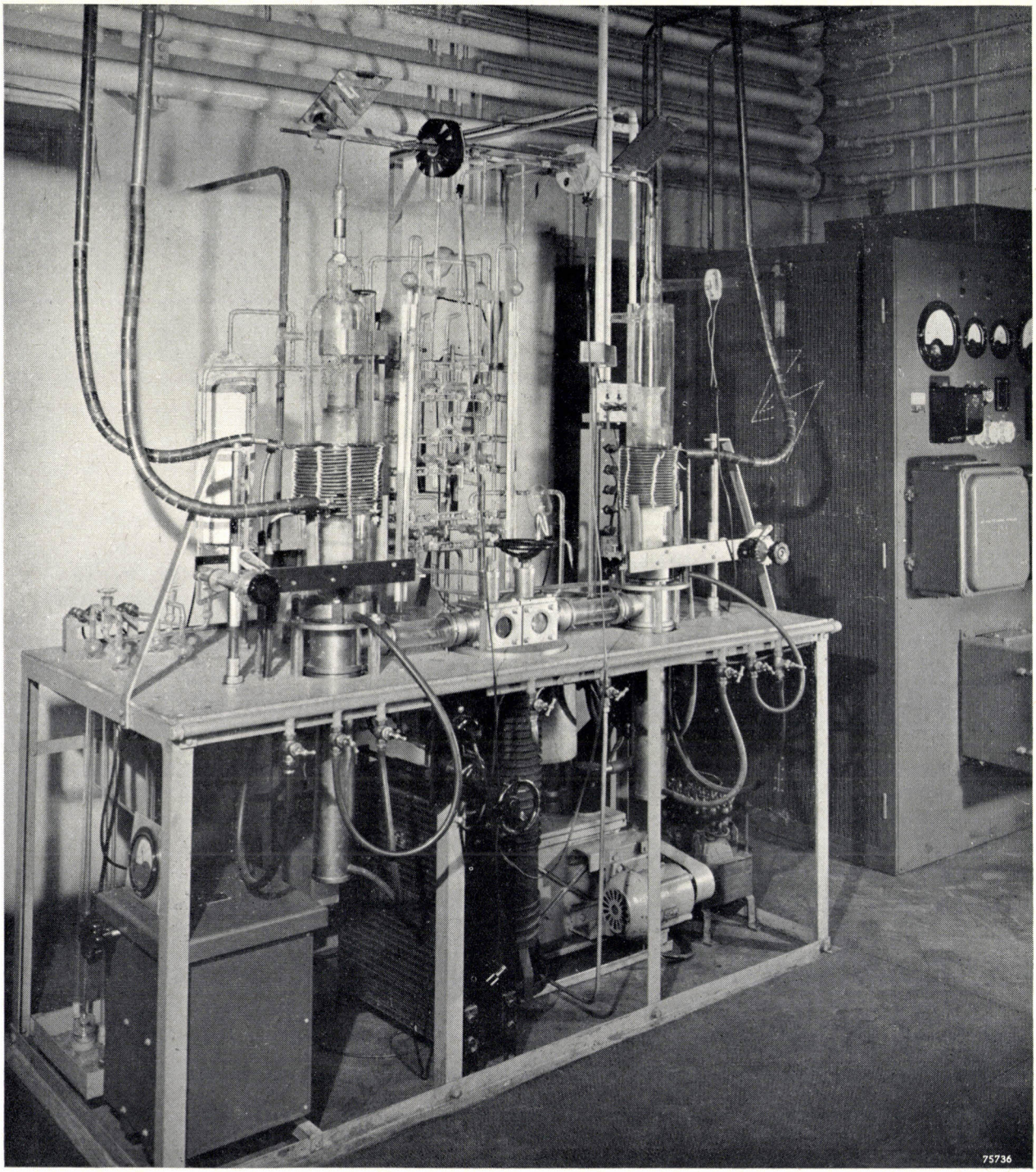


Fig. 1. Melting equipment with two furnaces for melting and casting alloys under high vacuum. The high-frequency induction heating coils are seen in position round the melting chambers. At right (in background) is seen the high-frequency generator (300 kc/s, 25 kW) and, below the bench, the high-vacuum and rotary backing pumps. On the bench, behind the furnaces, will be seen the apparatus for supplying purified gases (hydrogen, argon, oxygen, nitrogen) to the furnaces. Mirrors are mounted above the furnaces to give a clear view of the metal in the crucible.

give a more complete outline of the methods employed in preparing alloys of carefully predetermined composition.

Some examples of the results obtainable by these methods will be given at the end of this article.

Description of equipment

Fig. 1 shows the complete equipment, with the high-frequency generator on the right and the melting apparatus on the left. In the latter, two

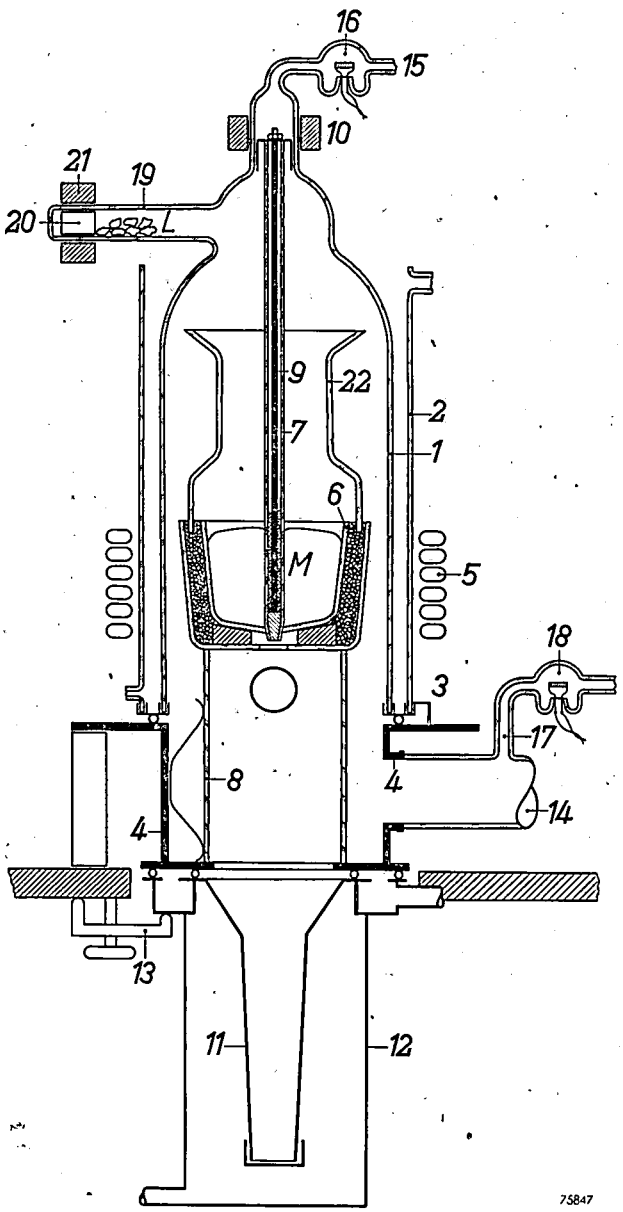


Fig. 2. Diagrammatic cross-section of one of the furnaces (not drawn to scale). The hard glass melting chamber *1* is cooled by the flow of water in the glass jacket *2* and is exhausted through the pumping tube *14*. The metal to be melted (*M*) is placed in a crucible *6* and is heated by high-frequency current in the coil *5*. When the metal has been melted and purified, any required alloying constituents *L* are added from the side tube *19* (pushed out by a ferromagnetic plunger *20*, actuated by a movable solenoid *21*). For casting, the plug *7* is withdrawn by energizing the solenoid *10*. The molten metal then runs out of the hole at the bottom of the crucible into a mould *11*, which is cooled by the flow of water in jacket *12*. Details of construction: the chamber *1* and jacket *2* are mounted on a metal ring *3*, by means of vacuum-tight cement. This ring is held by three clamps on the base *4* of the whole unit; a rubber gasket coated with vacuum grease and lying in a groove in *4* ensures a vacuum-tight joint. The crucible *6* stands on a refractory ring in an outer crucible of quartz, the space between the two being filled with fragments of crucible material. The outer crucible stands on a wide quartz tube *8* which is held in place in the base *4* by means of springs. The plug *7*, which is in effect a ceramic tube, is not attached directly to the ferromagnetic bush within the solenoid *10*, but is carried by a molybdenum rod *9*, this being wedged in the tube by carborundum particles. The metal housing *12* enclosing the mould *11* is held against the base *4* by three clamps *13* with rubber gaskets, to ensure a vacuum-tight seal for the mould and a water-tight seal for the housing. Different gases can be introduced into the chamber through tube *15* and these are removed again through the side tube *17* which is connected to the backing pump. Indicator tubes *16* and *18* serve for checking the oxygen content of the gas; *22* is a quartz funnel.

furnaces will be seen, each of which is fitted with its own induction heating coil.

The pumps used for exhausting the furnaces to a pressure of 10^{-5} mm mercury are housed beneath the bench; they consist of an oil diffusion pump (500 litres per sec.) and a rotary oil backing pump. On the bench may be seen the large metal valves and wide tubes used for connecting the furnaces to the pumps.

Fig. 2 is a diagram of one of the furnaces. The melting chamber proper lies inside a hard glass cylinder (*1*) which is cooled by a continuous stream of water through the jacket (*2*) when melting is in progress. This cooling prevents the liberation of gases from the walls.

At the centre of the chamber is the crucible (*6*) in which the metal is melted. The coil (*5*) for generating the high-frequency electromagnetic field consists of a helix of copper tube, also water-cooled. The coil is mounted outside the melting chamber, so that an alternating current of very high frequency can be used (300 kc/s), ensuring compactness of the equipment [3]). The heat developed in the metal is generated mainly by the eddy currents, but in the case of ferromagnetic metal (iron, nickel, cobalt), hysteresis losses — below the Curie point — contribute towards the heating effect.

The crucible stands on a ring of refractory material in an outer quartz crucible, which is in turn supported by a wide tube of quartz (*8*); this is held in position in a heavy metal base (*4*) by means of three springs. The space between the two crucibles is filled with small pieces of broken crucible material. Both crucibles have a hole at the bottom; the hole in the inner one is closed during melting by a long tube, slightly tapered at the end (*7*) and made of the same ceramic material as the crucible itself. When the metal is ready for casting, this tube is withdrawn; the hole is thus opened and the molten metal flows out into a metal mould (*11*). The tube is lifted by the action of a solenoid (*10*) on a ferromagnetic bush, by means of which the tube is suspended. It is not advisable to attach the ferromagnetic bush directly to the ceramic tube, as the

[3]) If the H.F. coil were placed in the melting chamber.— which is done in other apparatus (employing lower frequencies) in order to permit *tilting* of the whole furnace for casting — gas discharges between the turns of the coil would occur, owing to the low gas pressures now employed.

latter would then tend to break easily, especially when the melt reaches a high temperature; for this reason the force is transferred to the tube by a molybdenum rod (*9*) attached to the bush and wedged within the tube by carborundum granules.

The mould (*11*) consists of a slightly tapered copper tube closed at the bottom end with a nut and a suitable vacuum sealing compound; the upper end is flared and terminates in a flange by means of which the mould is attached to the apparatus (with rubber gaskets to ensure vacuum-tightness). The flange rests on the upper channel-shaped rim of the metal casing (*12*): water flowing through the casing and channel effectively cools the mould and gaskets.

The ample cooling of the mould — at least with the simple form of mould depicted (*11*) — ensures that contraction of the metal on solidifying does not result in blow holes: solidification takes place from the bottom upwards, and the shrinkage results only in a crater at the head of the casting (the last part to solidify). The quantity of metal to be cast can be so chosen that the crater occurs wholly within the tapered head of the casting, thus ensuring that the bar itself is solid throughout its whole length.

Melting in a gaseous atmosphere

The melting of the metal prior to casting must often be done in a gas such as argon or hydrogen. The reason for this may be to minimize the evaporation of relatively volatile constituents, but usually — and this refers especially to hydrogen — it is to reduce the oxygen content of the metal as much as possible. Before the gases are passed through the inlet (*15* in fig. 2) into the melting chamber, all residual oxygen is removed by passing them through a purifier (see article referred to in footnote [2])); this consists principally of a tube containing copper chips heated to 500 °C and two coolers, immersed in liquid oxygen when the gas to be purified is argon, or in liquid nitrogen for hydrogen. The purity of the incoming gas is verified by means of a Cr-Ni steel (18/8) indicator tube [4]) (*16*). The gas leaves the melting chamber through a narrow side tube (*17*), passes through another indicator tube (*18*), and escapes through a mercury trap (not shown in the drawing).

Nickel and iron are among the metals from which the last traces of oxygen can be removed by means of hydrogen. The indicator tube (*18*) shows when practically no more oxygen is being released from the molten metal. At this point, however, the metal will contain a quantity of hydrogen in solution in place of the original oxygen. It is just those metals referred to, i.e. nickel and iron, that are much more capable of dissolving hydrogen when liquid, than in the solid state. When casting is effected in hydrogen, therefore, much hydrogen must necessarily be liberated when the metal solidifies. Solidification takes place so rapidly in the water-cooled mould that many gas bubbles are included in the casting. To illustrate this, *fig. 3* depicts an iron bar sawn through after having been melted and cast in hydrogen at 1 atm. (For the normal examination of castings for the presence of blow holes, the castings are of course not sawn through; instead radiograms are taken, in which cavities are revealed as dark spots.)

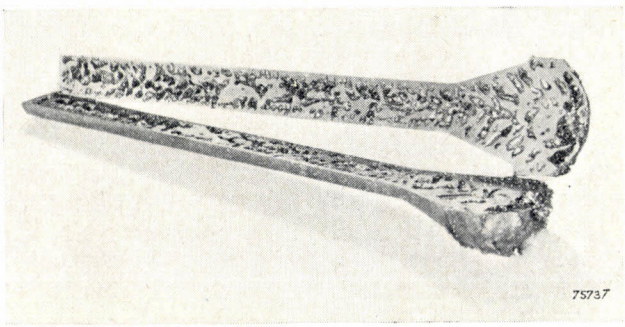


Fig. 3. Iron slug (sawn in half), as produced in the vacuum apparatus after purification and casting in hydrogen at a pressure of 1 atm. Owing to the large volume of hydrogen dissolved, and subsequently liberated by the rapid cooling, the metal is full of blow holes. This can be avoided by pumping away the hydrogen before casting.

The evolution of gas during solidification can be completely prevented by pumping away the hydrogen after purification; in fact, if the gas is pumped away until a pressure of 10^{-3} mm mercury is reached, a cast bar of purified iron shows no trace of blow holes or porosity at all. It is necessary to commence pumping slowly, as the dissolved gas otherwise escapes with such violence that drops of metal are entrained by the gas. To prevent this, the melting chamber is first connected to the backing pump by means of a cock in the narrow tube (*17*); when the greater part of the gas has been removed, the wider tube (*14*) to the main pump is opened and the pressure is reduced to below 10^{-3} mm Hg in a matter of minutes. This pressure is measured at the main inlet tube (*15*), to ensure that the pressure in the melting chamber is not actually in excess of the value measured.

If the required alloy is to contain elements that have a high affinity for oxygen (e.g. aluminium, zirconium, titanium, vanadium) it is better not to

[4]) G. W. Rathenau and H. de Wit, Metallurgia **40**, 114, 1949.

add these to the other metals in the crucible from the start as, during the melting process, they then combine with the oxygen present in the iron, nickel, etc. The oxides thus formed, viz. Al_2O_3, ZrO_2 etc., cannot be reduced with hydrogen. It is true that these oxides coagulate and, owing to their higher or lower specific gravity compared with that of the melt, will either sink to the bottom or rise to the surface. However, in practice, this coagulation and separation take places very slowly. It is therefore better to add the above-mentioned "impurities" to the melt only when those constituents which can be reduced with hydrogen have been fully purified. The materials to be added are placed in a side tube (*19*), and at the appropriate moment they are introduced into the melt by means of a ferromagnetic plunger (*20*) actuated by a movable solenoid (*21*). The special shape of the transparent quartz funnel (*22*) ensures that the metals fall into the crucible. This funnel also fulfils other important functions, such as protecting the glass wall of the melting chamber from spattering metal, and condensing the greater part of the evaporated metal.

Some further general points concerning the preparation of very pure alloys will now be mentioned.

Choice of raw materials

For making alloys of high purity it is obviously necessary to select very pure ingredients. It should be remembered that all metals qualified as "pure" nevertheless contain a certain quantity of unspecified impurities. This applies also to "spectrochemically pure" metals, since spectrochemical analysis, apart from its limited (albeit high) sensitivity, does not usually give any indication of the presence of carbon, nitrogen, hydrogen, phosphorus or sulphur; these elements, even in very small concentrations, generally have unexpectedly pronounced effects on the properties of a metal or alloy. For example, a small percentage of sulphur in nickel is so detrimental that for many applications a nickel containing 0.01% of sulphur, which might indeed be termed "spectrochemically pure", would actually be regarded as highly impure and useless for the purpose (see *fig. 4*). In contrast, nickel containing 0.1% of cobalt may for the same purpose be considered fairly pure and quite usable.

This implies that when a choice is to be made from different qualities of a metal, it is not always the metal with the lowest total content of impurities that will be selected. If the impurities are such that they can be eliminated during the melting process only with difficulty, or not at all (and this includes cobalt and sulphur in nickel), the choice will depend on the harmfulness of the impurities present, from the standpoint of the particular characteristics or application in mind. Also, where the choice lies between alternative impurities *A* and *B*, the first of

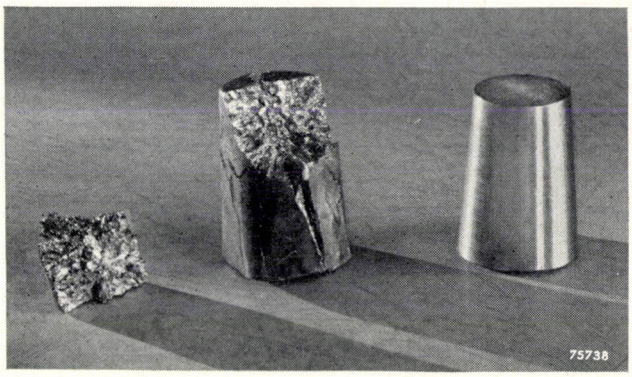


Fig. 4. Test specimens of nickel containing 0.02% sulphur, as obtained by solidification in the crucible (i.e. without casting). A specimen of this kind would have to be processed to give it the desired form by forging, rolling, drawing etc,, but, owing to the sulphur, the metal was too brittle to permit any deformation. The result of an attempt to forge this material is seen in the specimen at the left of the photograph.

which is very difficult and the second relatively easy to eliminate during the melting process, a sample containing a large quantity the impurity *B* might be preferred to one containing only a small percentage of *A*. For instance, the best electrolytic iron is of much higher purity than the best carbonyl iron, but the latter contains almost only carbon, oxygen and nitrogen as impurities, which can be removed during melting (see article referred to in footnote [2])), whereas electrolytic iron is liable to contain metallic impurities that cannot be removed easily, if at all.

Metallic impurities are eliminated without difficulty when the volatility of the basic metal is much lower than that of the other metals. For instance, in the equipment described, which can be exhausted to a pressure of 10^{-5} mm mercury, the manganese content of iron can be reduced to less than 0.001% by volatilization from the molten iron at 1600 °C.

In principle, metallic impurities of which the affinity for oxygen is much greater than that of the basic material can be eliminated by allowing the molten mass to absorb a controlled amount of oxygen; the impurities then undergo preferential oxidation. It has already been pointed out, however, that the oxides formed usually separate from the metal very slowly: for this reason this method of purification is not very practicable.

Effect of the crucible

Although the melting process offers facilities for removing many impurities from an alloy, the possibility of the introduction of impurities from

the crucible itself must always be considered. The choice of crucible material is therefore very important [5]). Generally speaking, it is necessary to employ *ceramic materials* for this purpose.

Metal crucibles are not usually suitable, since nearly all metals form alloys with each other. Iron and lead form an interesting exception. Iron is virtually insoluble in lead, so that lead can be melted in an iron crucible without contamination by the crucible itself.

In exceptional cases graphite crucibles can be employed, viz. for melting copper. In most cases, however, graphite is unsuitable for crucibles, as many metals tend to form carbides with the walls, or at any rate to dissolve appreciable quantities of carbon from them.

For the highly reactive metals titanium and zirconium there is in fact no crucible material that will give entire satisfaction. With graphite crucibles, contamination by the carbon is anything but negligible, and ceramics are not suitable because these metals in the molten state cannot be brought into contact with any oxides (even their own) without dissolving large quantities of oxygen. The graphite crucible in this case appears to involve the lesser of the two evils.

A material much employed for crucibles is pure aluminium oxide and, for a long time, mainly this material has been used in this laboratory. It is important to note, however, that this may not always be the most suitable ceramic material. Thus, if an alloy containing carbon is melted in a crucible of pure aluminium oxide, it will take up metallic aluminium as an impurity, as a result of reaction between the carbon and the crucible material. Even in the case of carbon-free metals, the use of crucibles of Al_2O_3 can lead to the appearance of Al in the melt; the following equilibrium is established:

$$Al_2O_3 \rightleftarrows 2[Al] + 3[O], \quad . \quad . \quad . \quad . \quad (1)$$

in which the square brackets denote the components dissolved in the molten metal. The partial dissociation of the solid Al_2O_3 as indicated in (1) must not be entirely disregarded, for example, when the melt is of iron, because of the high heat of solution of aluminium and oxygen in molten iron. The thermodynamics of this reaction show that the reaction constant at a temperature of 1600 °C is given by:

$$K = [\%\,Al]^2 \cdot [\%\,O]^3 \approx 10^{-13}, \quad . \quad . \quad . \quad (2)$$

where the quantities within the square brackets are the percentages by weight. The purification of carbonyl iron (see above) may be cited as an example: the iron is subjected to prolonged heating at 1600 °C, in an atmosphere of pure hydrogen. If the iron is contained in an Al_2O_3 crucible, the equilibrium illustrated above will be established and, in accordance with (2), the Al content of the iron will rise according as the oxygen content drops. If the oxygen content of the molten metal be reduced to 0.001%, the melt, according to (2), will take up about 0.01% Al from the crucible.

Iron of greater purity can be obtained by using a crucible made of *magnesium oxide*. Since the heat of solution of magnesium in molten iron is opposite in sign to that of aluminium, the reaction constant of the equilibrium corresponding to (1) is much smaller than that given by (2). Moreover, magnesium is much more volatile than aluminium, and the magnesium formed during the process leaves the molten iron by evaporation.

Typical analyses of carbonyl iron with purification in Al_2O_3 and MgO crucibles will be found in *Table I*.

The electrical resistivity at low temperatures is a very reliable indication of the purity of a metal; the purer the metal, the lower this "residual resistance". To obtain some idea of the order of magnitude of this residual resistance, the electrical resistance R at 291 °K (room temperature) may be compared with that at 20 °K (temperature of liquid hydrogen). For iron melted in an Al_2O_3 crucible the ratio R_{291}/R_{20} was found to be 31. The iron melted in an MgO crucible yielded a ratio of 140, and this high value (as also Table I) illustrates the very high purity of MgO iron.

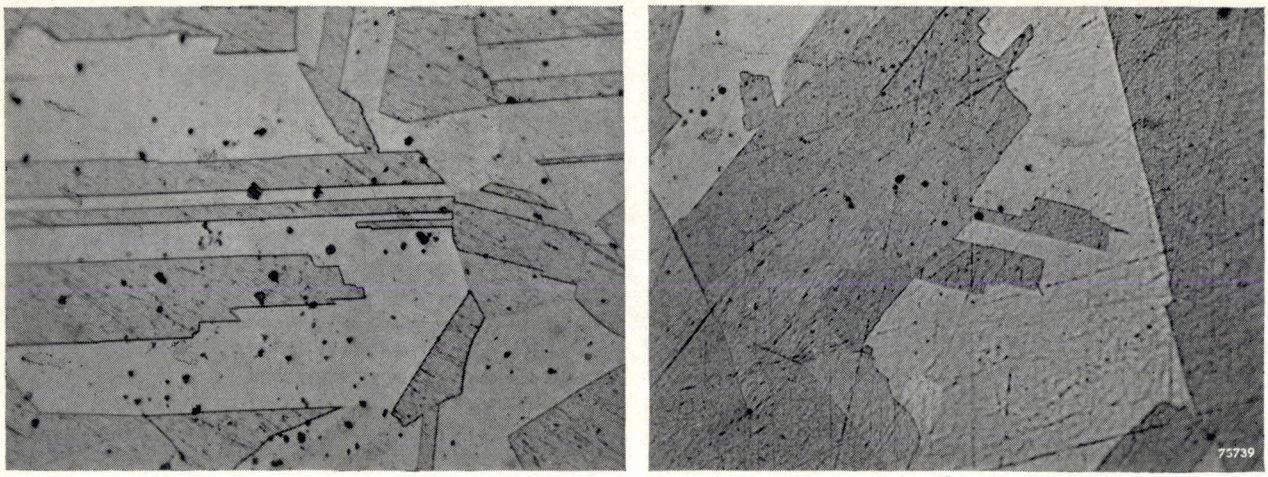
Table I. Percentages of impurities in iron melted in the vacuum apparatus in two different crucibles.

	Al_2O_3-crucible	MgO-crucible
Mg	0.000[7]%	0.001[6] %
Si	0.002	0.002
Al	0.027	0.001
Cu	0.002	0.000[5]
Co	0.001	0.000[5]
Ni	0.006	0.003
C	0.000	0.000
S	0.000	0.000
O	0.000	0.000

Applications of the method

The preparation of pure iron in an MgO crucible, as described above, is representative of the high degree of purity obtainable with this apparatus. Taking this pure iron as a starting point, investigations have been made on the individual influence of various impurities of practical importance upon

[5]) Not long ago a description was given of a method of melting metals in a freely suspended state, i.e. not in contact with a crucible; see E.C. Okress, D. M. Wroughton, G. Comenetz, P. H. Brace, J. C. R. Kelly, Electromagnetic levitation of solid and molten metals. J. appl. Phys. 23, 545-552, May 1952. So far, however, this method has been developed only for very small quantities of metal.



a *b*

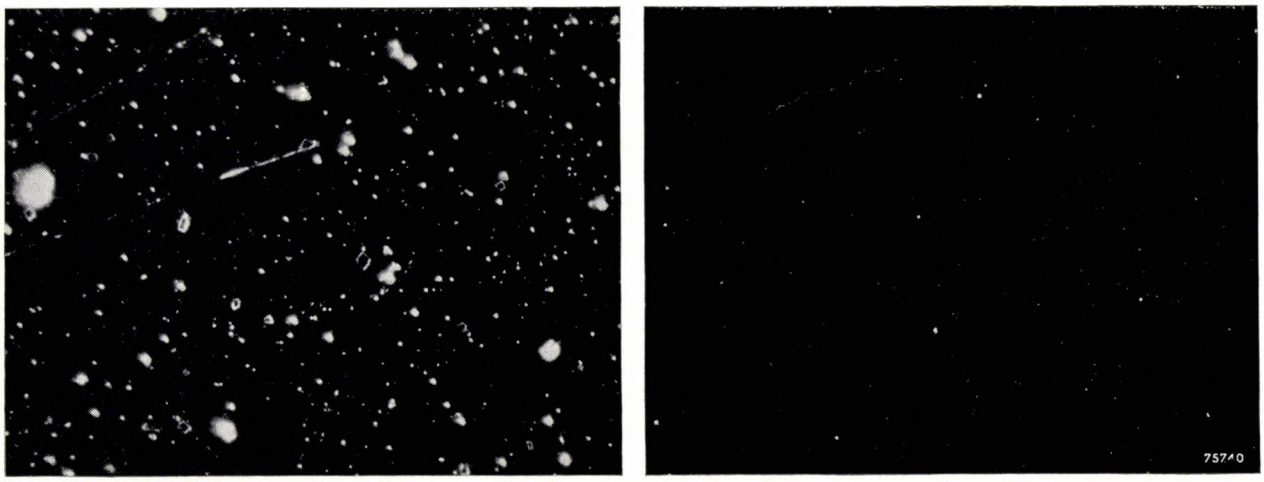
Fig. 5. *a*) Photo micrograph of an alloy of 80% Ni + 20% Cr prepared in accordance with conventional methods.
b) Photo micrograph of the same alloy prepared in the vacuum apparatus. It is at once evident that this contains much fewer inclusions than the specimen in (*a*). Owing to the greater purity of this specimen compared with (*a*), etching for a longer period was necessary to render the grain boundaries visible. The longer etching is evidenced by the more pronounced scoring of the polished surface. Magnification of (*a*) and (*b*) 400×.

the characteristics of the iron; the effects of certain combinations of impurities have also been studied. The results of these experiments have already been published in this journal and elsewhere [6]), and further results will be given in the near future.

Many nickel alloys of exceptional purity have also been prepared in the vacuum apparatus. Some examples are given below.

1) *Figures 5* and *6* are photo micrographs of alloys of 80% Ni + 20% Cr, the one alloy prepared in accordance with conventional methods and the other in the vacuum apparatus. It is clearly seen that the last-mentioned alloy contains much fewer inclusions than the first.
2) A high degree of purity, and in particular the absence of inclusions, is essential in alloys which are to be magnetically "soft". *Fig. 7* shows the

[6]) J. D. Fast, Investigations into the impact strength of iron and steel, Philips tech. Rev. **11**, 303-310, 1950; J. D. Fast, Ageing phenomena in iron and steel after rapid cooling, Philips tech. Rev. **13**, 165-171, 1951; J. D. Fast and L. J. Dijkstra, Internal friction in iron and steel, Philips tech. Rev. **13**, 172-179, 1951; J. D. Fast, Strain ageing in iron and steel, Philips tech. Rev. **14**, 60-67, 1952; J. D. Fast and J. L. Meijering, Anelastic effects in iron containing vanadium and nitrogen, Philips Res. Rep. **8**, 1-20, 1953 (No. 1); J. D. Fast and M. B. Verrijp, Diffusion of nitrogen in iron, J. Iron and Steel Inst. to be published shortly.



a *b*

Fig. 6. Photo micrographs taken with "dark ground" illumination (light at grazing incidence) of the same alloys as in figs. 5*a* and *b*. By this method of illumination the polished, specularly reflecting surface of the metal is invisible, but the irregularities caused by the inclusions appear as bright points. *a*) Alloy prepared in the usual manner; *b*) alloy made in the vacuum apparatus. Magnification 550×.

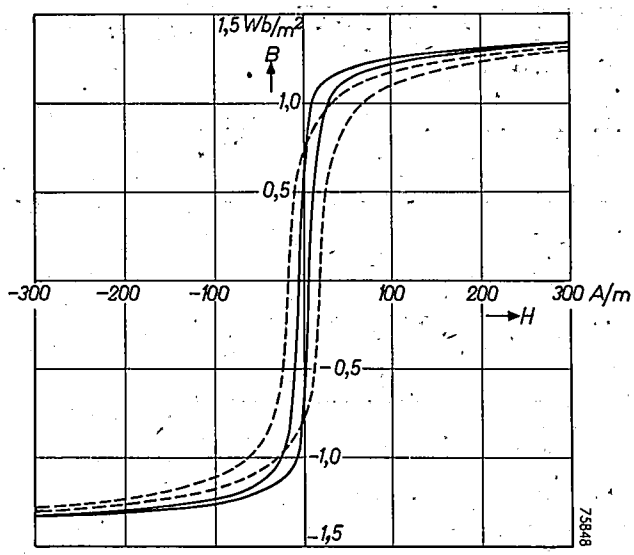


Fig. 7. Magnetization curve of two alloys of 40% Ni + 60% Fe. The broken lines relate to an alloy carefully prepared along conventional lines; the coercive force is as low as any value for such alloys mentioned in the literature [7]). The full line refers to an alloy prepared in the vacuum apparatus.

BH curves of two alloys of 40% Ni + 60% Fe; the full line refers to an alloy made in the vacuum apparatus, and the broken line to an alloy prepared with the greatest care along conventional lines. The coercive force of the last-mentioned metal is 16 A/m (0.2 oersted), which corresponds to that of the best alloys of this composition mentioned in literature on the subject [7]). The alloy made in the vacuum apparatus has an appreciably lower coercive force, viz. 3.2 A/m (0.04 oersted). In each case the heat treatment preceding magnetic measurement was the same; the specimens were made in the form of a ring (produced by rolling the cast bar to a flat plate and then turning on a lathe) and were heated together for 5 hours at 1200 °C in pure hydrogen.

3) The purity of alloys prepared in the vacuum apparatus was also demonstrated in the making of Al-Ni alloys containing about 1 atom Al to 1 atom Ni, as well as in alloys of this composition, to which small quantities of another metal, e.g. iron had been added. These alloys were all of a bright metallic colour and did not exhibit variations in tint observed by other research workers [8]). This seems to point to the fact that the coloration is attributable to impurities which can be avoided by preparation in the vacuum apparatus.

In accordance with what has already been said in the preceding section, the pure chromium and aluminium used for the alloys in experiments (1) and (3) were first placed in the side-tube *19* of the apparatus (fig. 2). They were added only after purification of the molten nickel in a stream of hydrogen. In view of the great heat of reaction between aluminium and nickel, the former had to be added in small portions. Only when all the metal had been added was the hydrogen pumped off and the alloy cast in the mould.

The importance of the preparation of pure alloys for metallurgical research has already been demonstrated. It may, however, be interesting to note that the vacuum melting apparatus can also be usefully employed in the development and manufacture of certain new products. In many cases the importance of alloys prepared in this manner lies not only in their precisely determined and possibly unconventional composition, but also in the fact that alloys melted and cast under high vacuum liberate very little gas when subsequently heated in the solid state. This is a very essential property for many of the components used in electron tubes, especially the lead-in wires through the glass envelope. Gas liberated by the internal components tends to reduce the life of the tube, while gas produced by the lead-in wires or plates may result in leaky seals. During the development of new types of electron tube so many unpredictable factors are operative, that it is a great advantage to be able to exclude one of them, viz. the harmful effects of evolved gases. The apparatus described in this article, which will handle charges up to about 2 kg, can even be used for production on a factory scale for certain components. This is the case for certain components which require only small quantities of metal in each tube (for example in the form of very thin wire) the cost of the material is of minor importance.

Summary. By melting metals and alloys under a high vacuum it is in many cases possible to attain a very high state of purity. Risk of introducing impurities into the pure alloy during processing is minimized by casting, also under high vacuum, in a cooled mould of suitable form. This simultaneously avoids the blow holes which normally occur when charges of several kilogrammes are cooled slowly in the crucible. The apparatus designed for this method of making castings is described in detail. Some points of practical interest with regard to the choice of the ingredient metals and crucible are also given. A few examples are included to demonstrate the high purity of the alloys prepared in the apparatus described; such alloys are particularly useful for metallurgical research, as well as for the development of electron tubes and other products.

[7]) Cf. R. M. Bozorth, Ferromagnetism, Van Nostrand Co. New York, 1951, fig. 5-14, p. 115.
F. Pawlek, Magnetische Werkstoffe, Springer Berlin, 1952, fig. 158, p. 188.

[8]) A. J. Bradley and A. Taylor, Proc. Roy. Soc. A **159**, 56-72, 1937.

THE DETERMINATION OF DROPLET SIZE IN ARC WELDING BY HIGH-SPEED CINEMATOGRAPHY

by P. C. van der WILLIGEN and L. F. DEFIZE. 621.791.75:778.534.83

In arc welding, material is transferred from the welding electrode to the work in the form of droplets. To a certain extent the size of these droplets determines the welding characteristics of the electrode. In order to provide data regarding the droplet size, films have been made of the transfer of the weld metal, with a camera taking up to 3000 frames per second. With the same camera, colour films have also been made which present a very clear picture of the welding process as a whole.

Introduction

Investigation into the manner in which the metal is transferred from the welding electrode to the work in arc welding, has revealed that this usually takes place in the form of free droplets of molten metal. Sometimes the droplets are so large that they constitute a short circuit while flowing from the electrode into the pool — in this case the droplets can no longer be considered to be wholly free.

To a great extent the droplet size determines the welding characteristics of the electrode; it is known that small droplets result in a smooth bead, and that the smaller the droplet, the more stable the arc. This stability relationship is especially noticeable in A.C. welding, where the arc is established twice per cycle. Electrodes producing small droplets, moreover, melt more quickly than those which give the larger drops, because a large drop usually clings to the electrode for a relatively longer time, thus impeding the melting of fresh weld metal. On the other hand the welder will generally prefer to employ an electrode giving a large drop when it is required to span the gap between two plates.

The number of different kinds of electrode manufactured has gradually increased in the course of time, and the need has arisen for a quantitative determination of the droplet size occurring with the various electrodes. In order to meet this need, systematic measurements have been carried out in the Philips laboratories at Eindhoven, and a description of the method employed is now presented. To illustrate the results obtained, the oxide-silicate and gas-shielded electrodes will be compared with the "contact" electrode which was derived from these types [1]). The results obtained with a "basic" or low-hydrogen electrode are also mentioned.

[1]) For a review of the characteristics and composition of welding electrodes see for example J. D. Fast, The function of the coating of welding rods, Philips tech. Rev. **10**, 114, 1948.

Before low-hydrogen and contact electrodes were introduced (the first of these before, and the second just after the last world war), mild steel and low-alloy steels were welded almost exclusively with oxide-silicate or gas-shielded electrodes, of which the Philips 50 and Philips 48 are examples.

Basic electrodes are so named because the main constituents of the coating are basic oxides. The outstanding feature of this kind of electrode is that the coating gives off very little water vapour which otherwise produces atomic hydrogen in the arc; this is liable to be absorbed by the molten metal, with detrimental effects on the weld [2]). These low-hydrogen electrodes excel by reason of their good mechanical properties and low sensitivity to impurities (especially sulphur) in the metal to be welded. However, they have to be used with a very short arc if porous welds are to be avoided, and this imposes rather heavy demands on the skill of the welder.

Contact electrodes are noted for their easy welding, and more particularly, for their high welding speed [3]). They do not represent a new type of electrode in the sense that the coating yields a new kind of slag, i.e. one not met in other kinds of electrode. Their essential feature lies rather in a modification of existing types, to produce an electrode suitable for contact welding, i.e. the arc is struck automatically when the coating is brought into contact with the work; the electrode with its coating is then rested on the work while the weld is being made. To make this possible, a semi-conductive coating is used, which is also thicker than usual, owing to the fact that a large part of the core metal is included in the coating in the form of powder.

Method of measurement and apparatus used

In order to determine the size of the droplets, films have been made of the arc, a method often employed to demonstrate the transfer of the weld metal from the electrode to the work. It was already known that the processes taking place within the arc are so rapid that very high film speeds (> 1500 frames per sec) would be required to permit these processes to be followed [4]). Use

[2]) J. D. Fast, Low-hydrogen welding rods, Philips tech. Rev. **14**, 96, 1952 (Nos. 3-4).

[3]) P. C. van der Willigen, Contact arc-welding, Philips tech. Rev. **8**, 161, 1946 and **8**, 304, 1946.

[4]) A. Hilpert, Werkstoffübergang im Schweisslichtbogen, Z. Ver. dtsch. Ing. **73**, 798-799, or Welding J. **8**, 21, 1929 and **12**, 4-8, 1933.

was made of an Eastman high-speed camera type III, which is capable of taking a maximum of 3000 frames per sec. on 16 mm film; exhibited at the normal speed of 24 frames per sec. the pictures are seen slowed down by a factor of up to 125. The camera does not of course operate at full speed from the moment of starting; this speed is developed in accordance with the curve shown in *fig. 1*.

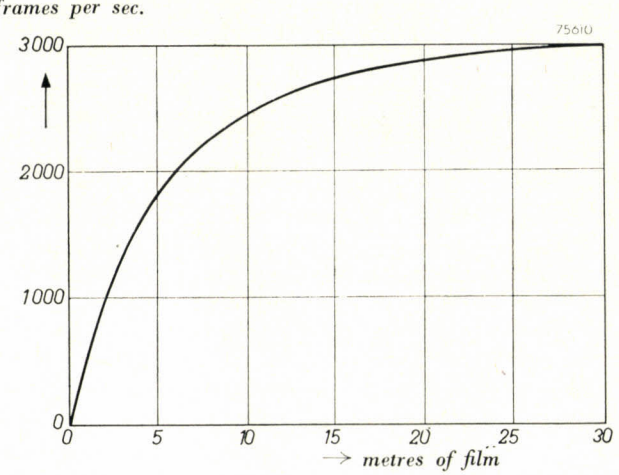


Fig. 1. To determine the diameter of the droplets transferred from welding electrode to work, a cine camera taking up to 3000 frames per sec was used. The curve in the diagram shows how the speed develops after the camera is started.

To obtain an indication of the speed relating to every part of the film, 50-cycle A.C. was used for the welding arc; the current thus passes through zero 100 times per second, so that the arc extinguishes that number of times per second. Another argument in favour of filming an A.C. arc is the increasing use of the A.C. welding technique, which involves much less difficulty due to magnetic arc blow and requires less expenditure on equipment.

The fumes, as well as the light emitted by the arc, tend to make filming difficult. The fumes can be blown from the field of view by means of a current of air directed across the arc. To minimize the effect of too much light from the arc itself, a powerful light source was placed behind it (*fig. 2*); originally a 60 A arc lamp from a cinema projector

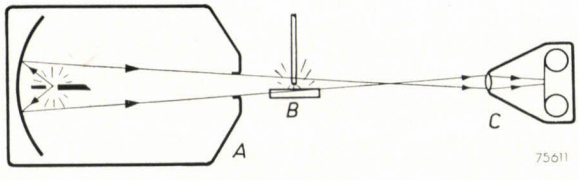
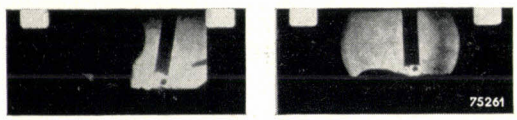


Fig. 2. Diagram of set-up for the filming in silhouette of the weld metal droplets in the welding arc. The light from a carbon arc lamp *A* is concentrated in the welding arc *B* by a concave mirror, so that the droplets can be seen against a bright background. *C* camera.

was used, the light being concentrated around the welding arc by a concave mirror, to produce a background of high and constant brightness [5]). The welding arc thus became almost invisible against the background, and the droplets of molten weld metal could be filmed in silhouette.

However, this light source was cumbersome and proved rather unmanageable in use, and it was subsequently replaced by a water-cooled mercury vapour lamp, this being much more convenient to handle. To concentrate the light, a lens was mounted between the lamp and the arc. Another advantage of this type of light source is that it introduces no interfering objects in the beam; the carbon arc and mirror had to be very carefully watched, to ensure that no shadow from the carbon-holder fell across the welding arc (see *fig. 3*).

Fig. 4 depicts the actual apparatus used, with the mercury vapour lamp. The welding electrode was mounted in such a way that the arc remained in the same place while the electrode was being consumed, with the electrode-holder lying in the direction of the electrode itself and supported by four guide rollers. The work was pulled along below the arc at the required speed by an electric motor. It was not found necessary to use automatic feed for the electrode-holder to maintain a constant arc length.



a *b*

Fig. 3. Increase in the field of view obtained by replacing the arc and mirror by a high-pressure mercury vapour lamp in the apparatus depicted in fig. 2.
a) Field with carbon arc and mirror. The shadow of one of the carbon-holders is visible on the left.
b) With the mercury vapour lamp the shadow is eliminated. The silhouette of the small bead already deposited can be clearly seen.

The arc length was checked by means of an optical system with calibrated eye-piece, and the camera was started as soon as the arc attained the desired length. The operator hat no difficulty in keeping the arc length constant for the few seconds needed to take the shot (15 m of film).

Another very attractive method of photographing the welding process is by means of X-rays [6]), the advantage of

[5]) A similar arrangement was first employed by L. Bull; see M. Lebrun, La soudure électrique à l'arc et ses applications, Bibl. off. centr. acétylène et soud. autog., Paris, 1931, pp. 39-44.

[6]) J. Sack, A new method for the investigation of the transfer of material through the welding arc, Iron and Steel Inst. Symp. on the welding or iron and steel, part 2, 553-559, London, 1935. J. Sack, Welding and welding rods, Philips tech. Rev. **2**, 129, 1937.

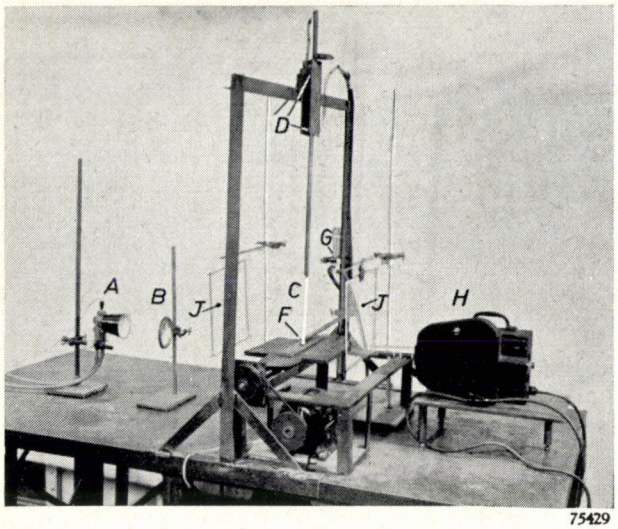


Fig. 4. Actual apparatus employed for filming the transfer of weld metal in the welding arc. *A* water-cooled mercury vapour lamp, the light from which is concentrated on the welding arc by a condenser lens *B*. *C* welding electrode (whitened with chalk to distinguish it from the holder). *D* insulated rollers between which the electrode-holder passes. *E* motor for drawing the work slowly under the arc. *F* jet (16 × 5 mm) through which 20 litres of air per minute is blown, to disperse the smoke laterally. *G* gauge indicating air velocity. *H* cine camera (up to 3000 frames/sec). *J* vertically mounted glass plates for protecting camera and condenser from spattering metal.

this being that it is thus possible to distinguish between the coating or slag and the metal. Further, no difficulties are then encountered owing to the light from the welding arc, or with fumes, so that it is not necessary to take steps to counteract these. As far as we are aware, however, no equipment has so far been designed that will give cinematographic X-ray pictures of sufficient frame speed. At the same time, efforts are being made to increase frame speed in X-ray filming techniques, for instance, Slack and co-workers [7]) describe a camera that will take 150 frames per second.

In order to obtain comparative data of the droplet size, it would be sufficient to have available a large number of instantaneous pictures, not necessarily in the form of film; the slow-motion film, however, does provide an opportunity for studying the whole process of droplet formation, the transfer of the weld metal and the form and motion of the arc.

Determination of characteristic droplet size

The individual pictures of the film are projected one by one on the screen, from which the diameters of the droplets as well as the diameter of the electrode are measured. As the actual electrode size is known, the real diameter of the droplets is easily ascertainable.

It is first necessary, however, to define the size of the droplet to be regarded as *characteristic* of an electrode. The transfer of metal is determined by the *volume* of the droplets, so it is desirable to work in terms of droplet volumes as obtained from the observed diameters. Let n_i be the number of droplets found of which the volume is v_i. It will be clear that the arithmetic mean droplet volume $\Sigma n_i v_i / \Sigma n_i$ will not be suitable for representing the volume of the characteristic droplet. As against the very large number of small droplets, the influence of the larger drops, which, after all, are those that contribute most towards the transfer of metal, would thus remain obscured [8]). This objection is removed by counting each drop a number of times proportional to its volume, i.e. to the contribution of the droplet to the transfer of weld metal. The characteristic droplet volume is then defined as:

$$v_c = \frac{\Sigma n_i v_i^2}{\Sigma n_i v_i}.$$

Now each droplet is encased in a protective coating of slag [9]), and the volumetric percentage of slag will certainly vary between one droplet and another. However, in order to take the average volume of slag into account in the calculation, we reduce the characteristic droplet volume by the (known) volumetric percentage of slag in the molten bead, taking the result to be the characteristic volume of the iron core of the droplets. From this we then obtain the *characteristic diameter* d_c of the droplet core.

As an example, *fig. 5* shows a histogram of the number of observed droplets as a function of their

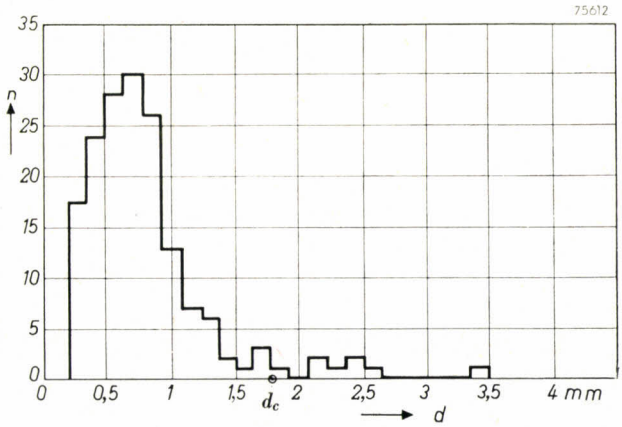


Fig. 5. Diameter of droplets transferred through the arc of contact electrode C 18-4, used with free arc. Vertical axis: number of droplets whose diameters lie between the limits defined by the intervals (1/7 mm in width) on the horizontal axis. (Here the characteristic diameter d_c of the core of the droplet is 1.8 mm). For the measurement, the film was projected with an enlargement of 7×. Droplets of smaller diameter than 3/14 mm (1.5 mm on the screen) could not with any certainty be distinguished from blemishes in the film, and are therefore not included.

[7]) C. M. Slack, L. F. Ehrke, D. C. Dickson, and C. T. Zavales, High speed cine-radiography, Non-destr. Test., Spring 1949, pp. 7-11.

[8]) Some of the metal evaporates. Strictly speaking, all the metal atoms in the vapour should also be taken into account as separate droplets, but this number of atoms is so large that the average droplet volume would be almost equal to the volume of a single atom.

[9]) See articles referred to in footnote [6]).

diameter, for the contact electrode [10]), C 18-4 (with which 38% of the total volume of the bead of molten material is slag). Another complication is encountered with contact electrodes in that the size of the droplet cannot be determined from the silhouette if the weld is made in accordance with usual practice. Normally, the cup of the electrode rests on the work, and the space between the coating and the work is then so small that the formation of the droplets cannot be observed. To enable the droplet size to be determined, the welding with the contact electrode was therefore done using an open arc. It remains a moot point whether the size of the characteristic droplet in the normal contact welding procedure would be the same.

The values obtained for contact electrodes C18-4 and C20-4, as well as for the oxide-silicate electrode Philips 50-5 (from which contact electrode C20 was derived), are shown in the *table*. A few of the results of measurement on the basic (low hydrogen) electrode Philips 56-5 are also included.

designed to give a large droplet, and the film in fact showed that metal was transferred from this electrode only when the arc was short-circuited (*fig. 6a* and *b*). It is therefore somewhat surprising that with the C18 electrode, which is the contact version of the Philips 48, the material is transferred only in the form of small, free, droplets (fig. 6*c*). The conversion to contact electrode has thus resulted in an appreciable decrease in the size of the droplet. Just the opposite has taken place, however, in the oxide-silicate electrode Philips 50 and its contact counterpart, the C20; the table shows that the ratio d_c/D' of the C20 is almost twice the ratio d_c/D of the Philips 50. No explanation has as yet been found for these facts.

The ratios d_c/D' of the C18 and C20 electrodes are practically the same. It is remarkable that a much larger percentage of very small droplets was found for the C18 than for any other of the electrodes measured. In the calculation of d_c, however, these smaller droplets are of minor importance.

Table. Diameter of the iron core of the *characteristic* droplet for several types of welding electrodes.

Electrode	Current (amp.)	Arc voltage (volt)	Arc length (mm)	Drops per sec.	d_c (mm)	d_c/D	d_c/D' (for contact electrodes)
C 18-4	180	45	4.8	185	1.9	0.47	0.41
,,	180	47	6.6	133	2.0	0.50	0.44
,,	165	42	6.1	145	1.8	0.44	0.38
C 20-4	190	35	6.1	91	2.3	0.58	0.44
Philips 50-5	200	35	4.8	162	1.3	0.26	
Philips 56-5	230	37	8.0	29	2.7	0.54	
,,	230	37	6.2	32	2.8	0.56	
,,	230	38	5.3	42	3.3	0.66	
,,	200	38	5.5	24	2.6	0.53	

In general, larger droplets are to be expected to result from larger core diameters D, so that, to characterise the type of welding rod, the value of d_c obtained should be divided by the value of D. For the contact electrode, moreover, the ratio d_c/D' is also given, where D' is the core diameter which the contact electrode would have if all the iron in the coating were included in the core. (In the C18, 24% of the iron is in the coating as powder; in the C 20, 41%.)

Discussion of results

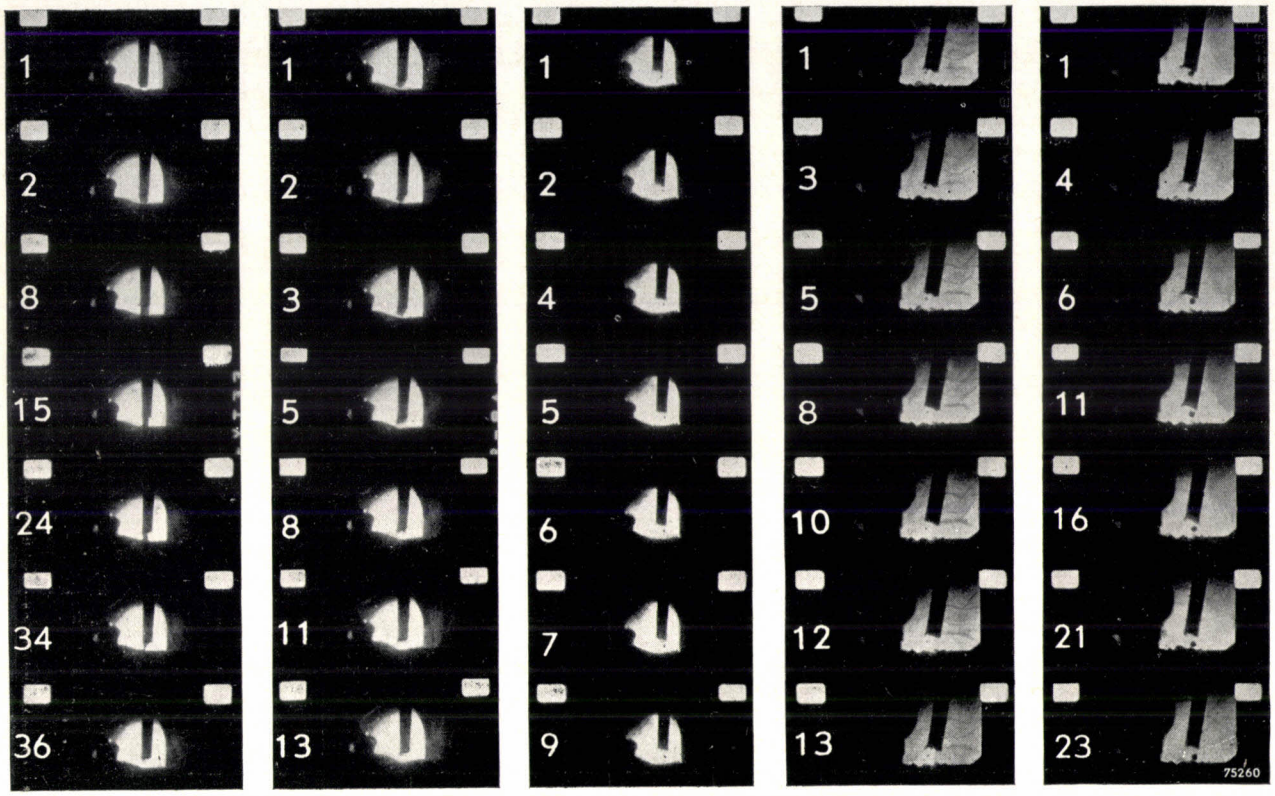
The gas-shielded electrode Philips 48 was specially

The molten metal at the extremity of the Philips 50 electrode, as well as that in the metal pool, is very mobile. This fact, coupled with the smallness of the droplet (the smallest measured) is related to the comparatively high FeO content of the molten metal (approx. 0.04%), which greatly reduces the viscosity and surface tension of the molten metal. The droplets transferred are irregular in shape.

The Philips 56, low-hydrogen electrode gives the largest droplets (figs. 6*d* and *e*); the spherical form of the droplet is characteristic of this electrode. It is known that the FeO content of the weld metal of this basic electrode is very low (0.001%); the viscosity and surface tension are accordingly high.

[10]) The second figure in the designation of an electrode in every case indicates the core diameter in mm.

Sack [11]) has also measured the droplet size of the Philips 50-5 and 48-5 electrodes, in this case by welding rapidly over a copper plate, so that the droplets were caught separately. These droplets, from which the slag was removed, were sorted according to weight, and the characteristic diameters of the droplets computed along much the same lines as those described above. For the Philips 50-5 electrode, Sack obtained a value of $d_c/D = 0.7$, which is roughly twice the value found by us. It is very probable, however, that by this method the droplets obtained are not droplets in the true sense, but larger masses formed by the merging of several droplets before solidification. For the Philips 48-5, Sack obtained droplets for which $d_c/D = 0.9$, but the method employed by him does not really lend itself to revealing whether the transfer of material is accompanied by shorting of the arc.



a *b* *c* *d* *e*

Fig. 6. Silhouette pictures of droplet transfer in arc welding. Some of the intermediate frames have been omitted from each strip, as will be seen from the numbering. These reproductions do not do justice to the film as exhibited; for example, it is barely possible to see from the above whether the arc is burning or not, whereas in the actual film the arc can be clearly seen to extinguish and re-ignite.

a) Metal transfer from Philips 48-5 gas-shielded electrode. The short-circuiting of the arc is clearly visible, starting at the moment the alternating current passes through zero and lasting almost one whole cycle ($^1/_{50}$ sec.). Film speed 2000 frames/sec.

b) Further pictures of the Philips 48-5 electrode. In frames 1, 2 and 3 the arc is extinguished (current = zero). In 2 there is a momentary short circuit, but no material is transferred. From 5, 8, 11 and 13 it will be seen that the arc exerts a force on the droplet, thus distorting it (see H. von Conrady, Der Werkstoffübergang im Schweisslichtbogen, Elektroschweissung 11, 109-114, 1943.)

c) With the contact electrode C18-4, which is derived from Philips 48 electrode, the metal transfer is in the form of small droplets. Weld made with free arc.

d) The low-hydrogen (basic) electrode, Philips 56-5. Two free droplets are seen, the upper one of which has the dimensions of the characteristic droplet ($d_c/D = 0.57$). The arc is extinguished only in frame 3.

e) The Philips 56-5 again. Here a large droplet is seen in process of being transferred. The arc is extinguished in 1 and 23. The 23 frames represent one half-cycle of the A.C. (1/100 sec).

The "optical efficiency"

In order to gain some impression whether the transfer of weld metal was actually recorded quantitatively on the film, the volume of material (iron and slag) transferred per second was first determined from the measured droplet diameters. This result was then divided by the volume of iron and slag per second ascertained from the actual weight of the quantities of material deposited and the known specific gravities. For the C18-4 this "optical efficiency" was found to be 100%, which of course it should be. The basic electrodes gave values of 150 to 200%, and the Philips 50 only

[11]) J. Sack, Overhead welding, Philips tech. Rev. **4**, 9, 1939.

a *b* *c*

Fig. 7. Extracts (not sequential) from three colour films of arc welding with A.C. The differences in colour between the films are due to the use of different filters. (*a*) and (*b*) were filmed with the set-up shown in fig. 8; (*c*) with that of fig. 9.

a) Philips 50-5. In the top and bottom frames the arc is just re-igniting after the current has passed through zero. The air-blast used for dispersing the smoke also tends to blow away the arc. In the last frame but one the arc is extinguished (current passes though zero). Note the spatter on the right-hand side.

b) Contact electrode C20-4 developed from the Philips 50. The electrode is now touching the work. The arc is almost entirely shrouded by the cup formed by the coating and therefore burns quietly. No spattering is seen, and the arc is not affected by the air-blast used for dispersing the smoke. The arc is able to withstand a much greater air velocity than Philips 50 before it begins to "stutter". The current passes through zero in the second frame from the top.

c) Contact electrode drawn across the work at an angle that affords a view into the cup. The arc is extinguished in the top picture and is almost out in the last but one. It is clearly seen that the arc originates in the core of the electrode and that, irrespective of the direction of the current, it fans out towards the work.

50%. By moulding the droplets from the basic electrodes in a synthetic resin and then cutting them through, it was shown that many of these droplets contained cavities. Hence the droplet volumes observed from the film were greater than their true volumes. (In normal welding, the droplets merge in the liquid pool and the cavities entirely disappear.)

In the case of the Philips 50 electrode, a number of droplets were apparently not observed; these were presumably quantities of very small droplets which had been rendered invisible by the smoke, or by the fact that their silhouettes were indistinguishable from blemishes in the film. Clearly, in this case, the value of d_c obtained for this electrode, which is already very low, is nevertheless too high.

Colour films

In addition to the black and white silhouette films, some films were also made which were intended to give an overall picture of the welding process. On the advice of Prof. J. Brillié of Paris,

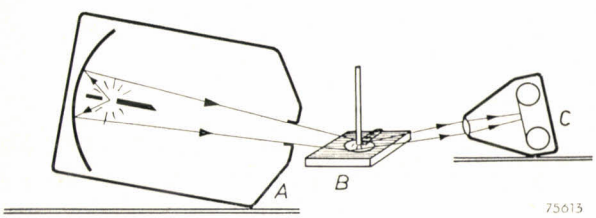


Fig. 8. Path of the rays in filming with reflected light using carbon arc lamp *A* as light source. This is the set-up used for taking the colour films. *B* welding arc. *C* camera.

these films were taken in colour (*fig. 7*). The light source was mounted so as to illuminate the work at slightly more than grazing incidence (*figs. 8* and *9*). The camera was similarly mounted, so as to photograph the work from an angle. In order to get as much reflected light as possible into the camera, the surface of the work was polished before each shot; in spite of this, so much light was lost that the arc was much brighter than the background.

The advantage of filming from an angle is that the pool of metal under the arc, with the layer of liquid slag on its surface, is then clearly visible. This gives a good perspective view of the welding processes. In addition to the advantage of more or less natural colouring, the use of colour film — owing to its laminated structure — gives pictures almost entirely free from the halation that otherwise occurs when the camera is directed towards the light source. With black and white films, halation is found to be a drawback at the beginning of the shot, when the camera has not yet attained its proper speed and the exposure time is accordingly too long.

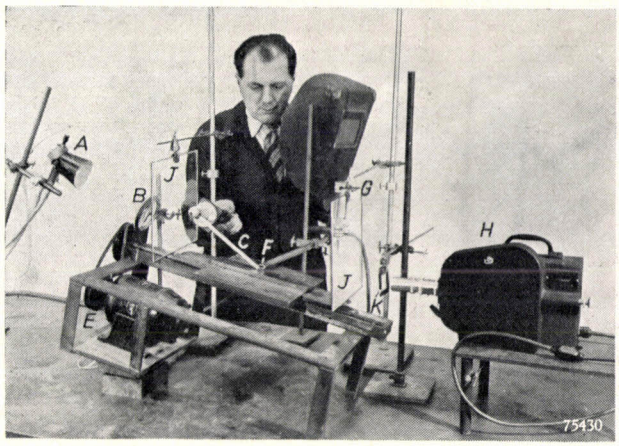


Fig. 9. Apparatus used for filming the processes in the crater of a contact electrode. The strip shown in fig. 9*c* was filmed in this manner. For meanings of letters *A* to *J* see fig. 4. *K* filter (Scott GG 3). The welding electrode has again been whitened. The weld is made in the plane defined by the light source *A* (here, a high pressure mercury arc) and the optical axis of the camera (instead of as in fig. 8, where the weld is shown along a direction perpendicular to this plane).

The series of pictures reproduced in fig. 7 are taken from three of these colour films. One series shows welding with an electrode with free arc (a), and the other two show contact welding with contact electrodes.

These films demonstrate the steadier arc and smaller spatter losses when the contact electrode is used; they also show the shape of the arc. The filming with the contact electrode at an oblique angle, giving a view into the cup (see figs 9 and 7c), confirms that the end of the core metal is always covered with a layer of slag during welding.

Summary. Systematic measurements have been carried out to ascertain the size of the droplets of weld metal transferred from welding electrodes to the work. For this purpose, films of the welding arc were made at a picture frequency o fup to 3000 per second. For the background a light of such brightness was used that the droplets were filmed in silhouette, notwithstanding the brilliant light emitted by the arc itself. Some of the measurements are discussed.

Colour films have also been made, at the same frequency, to give a more general impression of the processes taking place during welding. Certain features observed in the film are briefly reviewed.

BOOK REVIEWS

Fluorescent lighting: A review of the scientific and technical fundamentals and of the applications of the fluorescent lamp and its accessories. Editor-in-chief Prof. Dr. C. Zwikker, pp. 262, 187 illustrations. — Philips Technical Library — This book can be ordered through your technical bookseller.

The rapid evolution of the fluorescent lamp may be compared with that of the metal filament lamp (1907) and that of the gasfilled lamp (1913). On each of these occasions a new light source was introduced, of higher efficiency than its predecessor and of a colour making it suitable for general use.

The new tubular fluorescent lamp has been adopted so rapidly that nowadays practically all important new lighting installations for shops, restaurants, offices and factories use these lamps. The introduction of this new light source, however, has not been so simple as the changeover from the early incandescent lamp to its successor. The tubular shape demanded completely new fixtures. Control gear had to be built-in, and preferably in such a way that it would be unobtrusive. The usual method of increasing the lighting, namely by substituting a bigger lamp for a smaller one cannot be applied with fluorescent lamps, so that such lighting project has to be very accurately planned. In short, the development of the new lamp was not the end of the story. A new approach was also necessary with regard to fittings and to lighting schemes.

Anybody applying fluorescent lamps without taking into account these considerations, should be prepared for disappointing results.

For these reasons the publication of this book is highly appreciated. Under supervision of Prof. Dr. C. Zwikker, nine authors, all of the technical or scientific staff of Philips, have discussed the various aspects of fluorescent lighting. A summary of the subjects as given below.

Dr. J. L. Ouweltjes:	Luminescence, fluorescence and phosphorescence, and luminescent substances.
Dr. A. A. Kruithof:	Colour of the light emitted by tubular fluorescent lamps; colour rendering.
Dr. W. Elenbaas :	Gaseous discharges, lamp construction, and factors affecting the efficiency.
Drs. J. Funke :	Lamp types, starter switches, circuits.
Th. Hehenkamp :	Ballasts.
Ir. L. M. C. Touw :	Lighting fittings for fluorescent lamps.
Dr. D. Vermeulen :	Some fundamentals of lighting technique.
Ir. L. C. Kalff :	Applications of "TL" lamps.
Dr. R. v.d. Veen :	Uses of fluorescent lamps for agricultural purposes.

The book is amply and well illustrated; the picture of the colour triangle (fig. 19) and the energy-conversion graph on page 83 should be specially mentioned.

An extensive index of subjects renders valuable service to those interested only in certain parts of the subject.

The book is a very successful example of teamwork. There is a perfect balance as regards the various chapters, and overlappings and recurrences have been avoided. The book gives concise theoretical expositions as well as practical hints. It can be used not only by electrical engineers, but also by technicians, and by architects and physicists who wish to get acquainted with the present state of fluorescent lighting.

The reviewer's appreciation of this book can be best expressed by the remark that it is a great pity that a similar book has never been published on the gas-filled lamp.

N. A. HALBERTSMA.

Gas Discharge lamps, by P. J. Oranje, with a preface by M. Maurice Leblanc (President of the French Committee on Lighting and Heating and Vice-President of the International Lighting Committee), pp. 290, 161 figures and 32 photographs. — Philips Technical Library — This book can be ordered through your technical bookseller.

This is the English translation of the Dutch book by P. J. Oranje on discharge lamps. Since the publication of the Dutch edition, it has been brought up to date and supplemented by new material so as to correspond with the present state of the technique.

The book covers not only discharge lamps but also auxiliary gear, ballasts, fixtures and applications. As far as the last subject is concerned, the book contains some beautiful pictures of lighting installations. The lamps discussed are: sodium vapour lamps, mercury vapour lamps (high-

pressure and ultra-high pressure lamps with forced cooling by air or water), neon lamps, stroboscopic lamps and finally fluorescent lamps.

With the aid of numerous diagrams and characteristics the author discusses the fundamentals of discharge lamps and also matters concerning the auxiliary gear and installations. The construction of the lamps and the auxiliary gear is also explained with diagrams and pictures. The mutual influence of the lamp and its auxiliary gear is dealt with in detail. An interesting appendix gives dimensions and characteristics of the Philips products in this field.

As head of the discharge lamp testing department of Philips for many years, Mr. Oranje was enabled to become well acquainted with the problems of lamp construction and installation. The text represents the experience gained in that way, supplemented and arranged in order in a critical manner. This book has been written by a man who is not content with summarising previous publications, but who has gone through all the facts he describes and who has been obliged to *solve* many of the problems arising in this field.

The book gives an answer to numerous questions relating to the use of gas discharge lamps, and will greatly contribute to the elimination of wrong ideas, which are still only too prevalent.

C. ZWIKKER.

ABSTRACTS OF RECENT SCIENTIFIC PUBLICATIONS OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN

Reprints of these papers not marked with an asterisk * can be obtained free of charge upon application to the Administration of the Research Laboratory, Kastanjelaan, Eindhoven, Netherlands.

2024: A. A. Kruithof: Specification of light sources by the spectral band method (C.I.E. Proceedings, Stockholm 1951, Vol. 3).

The specification of light sources by the spectral band method, as introduced by Bouma, is examined. For 10 pigments, illuminated by standard illuminants A or C, the displacement r of the colour point in the colour diagram, caused by a 10 percent increase in energy in a band 5 mμ wide about the wavelength λ, is calculated as a function of λ. The ratio r/r_A (where r_A is the radius vector of the corresponding McAdam ellipse) is plotted against λ. The area under the envelope of these curves is divided into 7 or 8 parts thus obtaining the following bands:

380-445-485-525-560-595-625-750 mμ

and 380-445-470-500-530-560-595-625-750 mμ.

By a separate investigation it is shown that for a band between 380 mμ and x mμ, the displacement of the colour λ point by concentrating the luminance of the band into one wavelength (436.8 mμ) is a minimum for $x = 445$ mμ. The distributions found satisfy the condition that the mercury lines 546 mμ and 578 mμ shall not be too near the limit of a band.

2025: B. Verkerk: A possible extension of the density-difference method of spectrographic analysis (J. Opt. Soc. America **41**, 1017-1022, 1951).

An investigation into the use of Seidel's "transformed density" $D_t = \log [(d_0/d)-1]$, in which d_0 and d are the densitometer deflections of the unexposed plate and the spectral line, respectively. An empirical formula is derived for D_t as a function of the illumination E. Using this formula it is possible to estimate the error introduced if the D_t-log E curve is considered as a straight line for interpolation according to the density-difference method. This proves permissible under certain specified conditions in the density range 0.1—2. The influence of the wavelength on the shape of the transformed density curve is investigated and some experimental results obtained with the method are described.

2026*: G. W. Rathenau: Etude des phénomènes de récristallisation par microscopie électronique à émission (Revue Métallurgie **48**, 923-928, 1951, No. 12). (Study of recrystallisation by emission electron-microscopy; in French).

Description of a method for studying crystal growth by means of emission electron-microscopy. Results with face-centered NiFe alloys show, after intensive rolling, a cubic structure with (001) planes parallel to the rolling plane and the [100] direction parallel to the rolling direction. It is proved that, in general, crystal limits with highest energy are displaced, but it is impossible to predict whether crystals having a large surface energy will grow or dissappear. The former occurs in normal crystal growth, starting from small grains, the latter in secondary recrystallization. Most probably after primary recrystallization, large crystals are in the most favourable condition for growing. (See No. **1986**.)

2027: G. D. Rieck and J. M. Stevels: The influence of some metal ions on the devitrification of glasses (J. Soc. Glass Tech. **35**, 284-288, 1951, No. 167).

A study, at 1200 °C, by means of X-ray spectrometer diagrams, of the crystallization of glasses containing about 95 mole % SiO_2 and about 5 mole % of Li_2O, Na_2O, K_2O, BeO, MgO, CaO, BaO or ZnO. The univalent ions are shown to promote the formation of tridymite from initially formed cristobalite. The glasses with bivalent ions devitrify by formation of cristobalite. A number of these ions influence the crystallization in such a way that it starts with the temporary formation of quartz. Probably this is due to a direct influence of these ions on the glass, and not to an indirect influence on the transformation of the crystalline phases.

2028: W. Six: Het gebruik van koude-kathodebuizen als schakelelement (T. Ned. Radiogenootsch. **17**, 17-32, 1952, No. 1). (Cold cathode gas discharge tubes as switching element; in Dutch).

This paper points out that there is an electronic analogy for most circuits which can be built up with electromechanical parts, such as relays, switches, etc. A small cold-cathode gas-filled valve has been developed to be used as switching-element in electronic circuits.

Examples are given of equivalent electronic circuits for relays circuits, counting-and selecting-circuits, registers, busy-test and lock-out circuits and for talking-path circuits.

2029: M. Travniçek, F. A. Kröger, Th. P. J. Botden and P. Zalm: The luminescence of basic magnesium arsenate activated by manganese (Physica **18**, 33-42, 1952, No. 1).

In the system MgO-As_2O_5 there is a compound of the composition $6MgO.As_2O_5$. Products containing small concentrations of manganese show fluorescence upon excitation by cathode rays or by ultra-violet radiation. The band emitted depends on the valency of the mangenese. Divalent manganese causes fluorescence in a broad band in the green, with a maximum at 5050 Å. Tetravalent manganese causes a red fluorescence in five narrow bands at 6200 Å, 6300 Å, 6500 Å and 6580 Å. Both fluorescences decay exponentially; the time constant at room temperature is 130 sec^{-1} for the green band and 360 sec^{-1} for the red bands. For the red luminescent systems the spectral distributions of the absorption, excitation and fluorescence, and the temperature dependence of the intensity of the fluorescence, the fluorescence spectrum and the decay constant have been determined, and a quantitative explanation of several of these values has been given.

2030: J. A. Haringx: De wondertol (De Ingenieur **64**, O13-O17, 1952, No. 4). (The magic top (Tippe Top); in Dutch).

Starting from the analysis of a series of forced regular precessions the author has shown why and under what conditions the "tippe top" exhibits its typical behaviour. The limits of the ratio of its principal moments of inertia with respect to the eccentricity of its centre of gravity have been determined and the minimum initial rotation which is needed for reversal has been calculated.

2031: J. M. Stevels: Les pertes diélectriques des verres et les renseignements qu'elles donnent pour l'étude de leur structure (Silicates industrielles **16**, 325-328, 1951, No. 10; **17**, 15-18, 1952, No. 1). (Dielectric loss in glasses and its relevance to the study of their structure; in French).

The power factor ($\tan\delta$) of glasses is discussed as a function of the frequency (f), the temperature (T) and the chemical composition, both from a theoretical and a practical point of view. A three-dimensional model, showing the relation between $\tan\delta$, f and T is described. The influence of chemical composition on the power factor is discussed at some length.

R185: J. L. H. Jonker, A. J. W. M. van Overbeek and P. H. de Beurs: A decade counter valve for high counting rates (Philips Res. Rep. **7**, 81-111, 1952, No. 2).

A decade counter tube has been developed having a ribbon shaped electron beam and 1-dimensional deflection. The electrode configuration resembles that of a cathode-ray tube and is such that the beam can be fixed in any of ten discrete positions. Slotted electrodes permit part of the beam to fall on the appropriate part of a fluorescent layer, so that the instantaneous position can be read from the adjacent figure on the wall of the tube. The beam moves from one position to the next when an impulse is applied to one of the deflection electrodes. Counting and calculating devices constructed with these tubes permit a great economy in valves and components, in comparison to conventional counting circuits, The decade counter tube has small dimensions (the size of a normal receiving valve), low power consumption and low supply voltage (300 V). For

random pulses the tube has a mean resolving time of about 15 μsec when connected in a simple counting circuit. With the aid of one double triode this can be improved to 6 μsec. With a special circuit incorporating a triode-hexode and a secondary-emission tube it has been reduced to less than 0.2 μsec. For pulses occurring at equal intervals the resolving time is larger. See also Philips tech. Rev. **14**, 313-326, 1953 (No. 11).

R186: J. A. Haringx: Instability of thin-walled cylinders subjected to internal pressure (Philips Res. Rep. 7, 112-118, 1952, No. 2), See Abstract No. **1997**.

R187: P. Schagen, H. Bruining and J. C. Francken: A simple electrostatic electron-optical system with only one voltage (Philips Res. Rep. 7, 119-130, 1952, No. 2).

A simple electron-optical system is described for application in television pick-up tubes or image converters. This system consists of two concentric spheres, the sphere with the smaller radius, the anode, being provided with a hole. The electron paths in the field of force of this system are calculated and the influence of the hole in the anode is estimated. The calculations indicate that an inverted image of the cathode can be obtained at a distance equal to $6(n-1)R_a/(n-4)$ behind the hole, the magnification of this image being equal to $3/(n-4)$, where n is the ratio of the radii of the spheres and R_a represents the radius of the anode sphere. With experimental tubes, adapted to practical requirements, pictures of good resolution and geometry have been obtained.

R188: A. E. Pannenborg: On the scattering matrix of symmetrical waveguide junctions (Philips Res. Rep. 7, 131-157, 1952, No. 2).

The description of the electromagnetic behaviour of microwave circuits with the aid of the scattering matrix is systematically developed. Special attention is paid to resonators. The basic results obtained by Tomonaga on this subject are elaborated to the point of practical usefulness.

R189: J. M. Stevels: Some experiments and theories on the power factor of glasses as a function of their composition, III (Philips Res. Rep. 7, 161-168, 1952, No. 3).

It is shown that the power factor of borosilicate glasses (measured for widely differing frequencies) is a minimum on the straight line separating the accumulation region and the destruction region in the phase diagram, thus showing that on this line the rigidity of the network is a maximum. This maximum is less pronounced for the higher SiO_2 concentrations and vanishes at composition of 75% SiO_2-20.5% B_2O_3-4.5% M_2O (M = Na or Li). The experimental results are discussed from a theoretical point of view.

See abstracts **R 127** and **R 158**.

R190: A. E. Pannenborg: On the scattering matrix of symmetrical waveguide junctions, II (Philips Res. Rep. 7, 169-188, 1952, No. 3).

Continuation of **R 188**. General considerations about the consequences of structural symmetry of microwave circuits lead up to an extensive dicussion of junctions consisting of two parallel sections of rectangular waveguide having one of their sides in common.

R191: J. A. Haringx: Instability of bellows subjected to internal pressure (Philips Res. Rep. 7, 189-196, 1952, No. 3).

See No. **1998**.

R192: Th. P. J. Botden: Transfer and transport of energy by resonance processes in luminescent solids, II (Philips Res. Rep. 7, 197-235, 1952, No. 3).

Continuation of **R178**. A number of sensitized systems are dealt with. Of the several explanations for the phenomenon of sensitization, only a few are borne out by experiment. On the basis of some of these theories it has been deduced that the ratio of the quantum efficiencies of activator and sensitizer fluorescence will not increase with increasing sensitizer concentrations at constant activator concentration. This is found to be in contradiction with the author's results, obtained with the systems $Ca_3(PO_4)_2$-Ce-Mn, $Sr_3(PO_4)_2$-Sn-Mn, $Ca_5P_3O_{12}F$-Sb-Mn, $Ca_2P_2O_7$-Sn-Mn and also with Fröhlich's results on $Mg_2P_2O_7$-Ce-Th-Mn. An explanation of the experiments is given on the basis of: (1) a random distribution of the sensitizer and the activator ions, (2) direct transfer of energy from the sensitizer to a neighbouring activator (by quantum-mechanical resonance), (3) the assumption that if the excited sensitizer has not an activator in its neighbourhood, the excitation energy may be transferred via other sensitizer ions (by resonance) to a sensitizer that *has* an activator in its neighbourhood, followed by transfer of the energy to the activator. The numbers of cation sites around a sensitizer are calculated from the experiments on a few systems, for transfer both to an activator and to another sensitizer. The mechanisms of transfer and transport of energy are discussed on the basis of the theory of Mott and Seitz; the temperature dependance of sensitization may be explained in this way.

VOL. 15 No. 5, pp. 133-160 NOVEMBER 1953

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
THE PHILIPS INDUSTRIES

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, NETHERLANDS

MODERN CASTING TECHNIQUES

by H. J. MEERKAMP van EMBDEN.

621.74.045

Casting by the lost-wax process, an ancient craft practised with great skill already thousands of years ago by the Sumerians and the Egyptians, has retained its importance throughout the ages up to modern times. Many of the beautiful art bronzes of the Renaissance were cast by this method, and the same technique forms the basis of a modern industrial process for the mass production of metal articles, which has been adopted during the last decade.

The casting of molten metals to form articles of general utility is a very ancient craft that goes far back into the Bronze Age. If we examine the history of metallurgy, we find that the first metals to which primitive man had access were found in a native state (gold, silver, copper.) At a very early period, however, copper and bronze were separated from their oxidic ores in a primitive fashion, and the metals thus obtained were shaped by means of casting, sometimes followed by forging and other processes, into weapons and objects for religious and domestic use. Cast copper and bronze articles are known to have been already made in various regions in the third or fourth millennium before Christ.

The metallurgy of iron underwent an entirely different development. Native iron is very rare (meteorites), and the melting-point of iron was too high for the craftsmen of those days, so that extraction was restricted to reduction of the ore to metal in a solid or dough-like state. This metal, impregnated with slags, was hammered into bars by repeated forging, which also had the effect of expelling the slag. Only when the development of the furnaces was in an advanced state did it become possible to obtain liquid iron. The oldest known cast-iron comes from China and dates from about 500 B.C. In Europe the great technical development of cast-iron begins towards the end of the Middle Ages. Cast-steel, owing to its low carbon content, has a melting point lying 350 °C higher (above 1500 °C), and could be made only about the beginning of the 19th century, while the modern alloyed stainless and heat-resistant steels have only in the last few decades begun their remarkable development, as is the case with aluminium, magnesium and their alloys.

The number of alloys employed has greatly increased, especially in this century, and the same can be said with regard to the number of melting and casting methods. The crucible furnace, in which the material is melted in a crucible, the reverbatory furnace in which burning gases pass over the material, the blast furnace in which ore and coal are introduced together at the top while the air required for the blasting is blown in and the metal and slag drawn off from below: these are the furnaces which have been developed in the course of the centuries to the perfected equipment of to-day. Quite new are the electric furnaces. The introduction of the electric-arc furnace was greatly stimulated by the first World War. The high-frequency furnace, which made its début in industry during the 1930's plays only a very limited part in production as yet, being used mainly in the production of high-alloy steel. The low-frequency induction furnace, on the other hand, has won for itself a wide field of employment in the melting of

brass and bronze. Aluminium and magnesium alloys are very often melted in electric radiation furnaces.

Casting in sand-moulds

Now for the actual casting methods. Originally, a pattern was pressed in sand or loam and the hollow obtained was filled up with molten metal. At an early date, however, castings were made in closed moulds: two boxes filled with sand in which a pattern (usually a wooden one) was laid, so that a hollow was formed, half in the lower box and half in the upper box (*fig. 1*). The top box was then lifted off, the pattern removed and the top box replaced on the lower one. In the top box a channel ("sprue", "runner" or "gate") was made through which the metal was poured. After solidification of the metal, the mould is "shaken out", and gates and runners are cut off. During casting, the air in the mould and the water-vapour and gases liberated from the mould material by the hot metal, escape through the porous mass.

In this way, intricate castings could be produced. The technique has been more and more perfected in the course of the centuries, until we come to the modern sand or loam casting by which, with the aid of all kinds of contrivances, complicated parts can be cast, that form the basis of our tools and machines. No great dimensional accuracy can be expected, however, for two reasons. The first is that, after removing the pattern, it is not possible to replace the halves of the mould so accurately that there is absolutely no alteration in their mutual position. The second reason is that the pattern has to be taken out of the sand, which entails some knocking or vibrating to loosen it, with the result that the dimensions of the mould are slightly different from those of the pattern itself. A restriction on the shape of the moulding is that the pattern must be "detachable", that is, it may have no protruding parts that would prevent its removal from the sand. Yet it is astonishing what complicated castings can be made by the subsequent insertion of loose cores of sand fixed in the hollow of the mould. The mould for the cylinder block of a motor-car, for instance, is built up of a large number of sand parts (sometimes more than fifty). In this case various kinds of binding agents are added to the "sand" (clay, oil, cement and numbers of others) in order to meet all requirements as to strength, porosity (for the escape of gases etc., see above) and the quality of the metal surface.

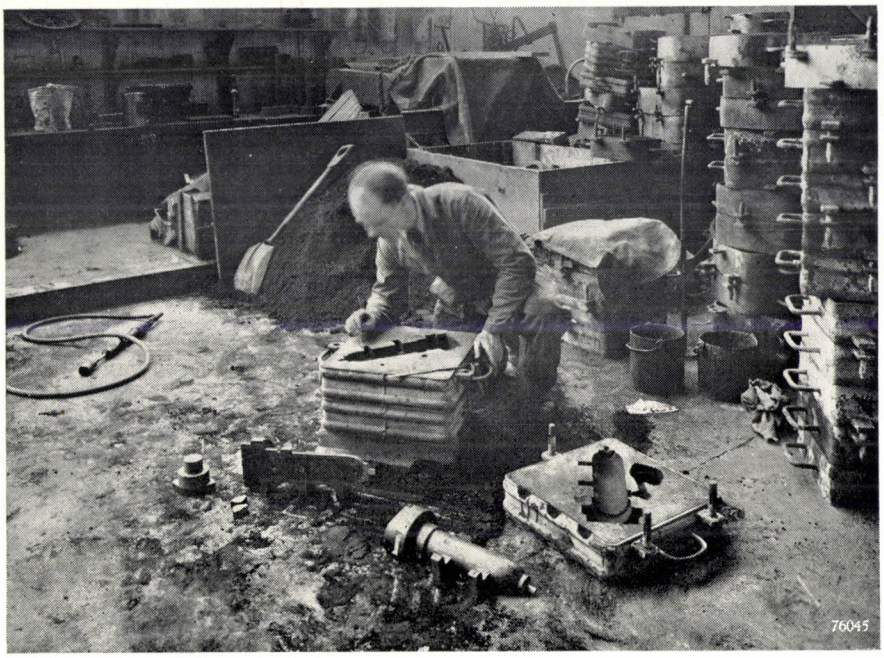


Fig. 1. Upper and lower boxes of a sand mould. The wooden pattern can be seen in the foreground. In the upper box the "gate" or "sprue" can be seen through which the metal is poured. (This photo, taken in the Reineveld Engineering Works at Delft, was supplied by the Casting Research Centre, National Council for Industrial Research, Netherlands.)

Permanent moulds

Interesting as the subject may be, we shall not go further into "sand-casting" but turn our attention to other methods. One of the oldest employs permanent moulds that can be used more than once.

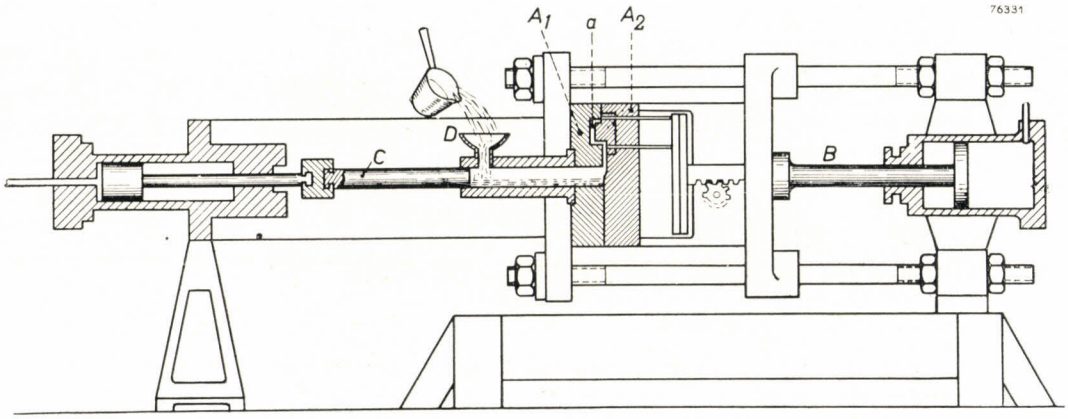


Fig. 2. Machine for pressure die casting. A_1 and A_2 are the two parts of the steel die. The moving part is pressed against the fixed part by the hydraulically moved piston B, while the piston C forces the liquid metal, poured into funnel D, into the cavity a of the die. (From "Designing for Alcoa die castings", Alum. Co. of Amer. Pittsburgh 1948.)

Even primitive peoples made use of such moulds hewn in stone and this method developed into metal moulds that are still employed on a large scale. They are, for example, used for the casting of ingots. These are afterwards worked up into sheets, bars, strips and wire by hammering, forging, rolling etc., thus forming the raw material of our entire metal working industry. Quite 90% of the metals at present consumed by industry is worked up in this way and feeds the metal working machines, and only a small part of the total supply of metal is cast directly to shape.

Permanent moulds are also used for the direct manufacture of utility articles. We speak of "gravity die casting" if the metal is poured in with a ladle, and of "pressure die casting" if the metal is forced into the die under (usually high) pressure (*fig. 2*). Gravity die casting is an old process; pressure die casting is an invention of this century of mass production [1]). This method is especially employed for the mass production of small parts made of low melting-point material. An example is the zinc alloy, from which all kinds of automobile parts are made, from the intricate carburettor to the radiator grill and the doorhandles. In addition, a great deal of pressure die casting is performed in aluminium and brass. Metals with higher melting-points cannot in practice be cast in this way, as the die cannot stand the sudden change in temperature during the injection of the metal, and is soon damaged.

Besides these methods, two others have recently come to the fore: the first is the so called "Croning process" (C-process or shell-mould casting) and the second is the lost-wax process, also called "precision casting" and "investment casting" (see later).

Croning process

The Croning process [2]) is new and originates from Germany where it was invented during the last World War. It it based on the following principle (cf. *fig. 3*). A metal pattern is divided into two

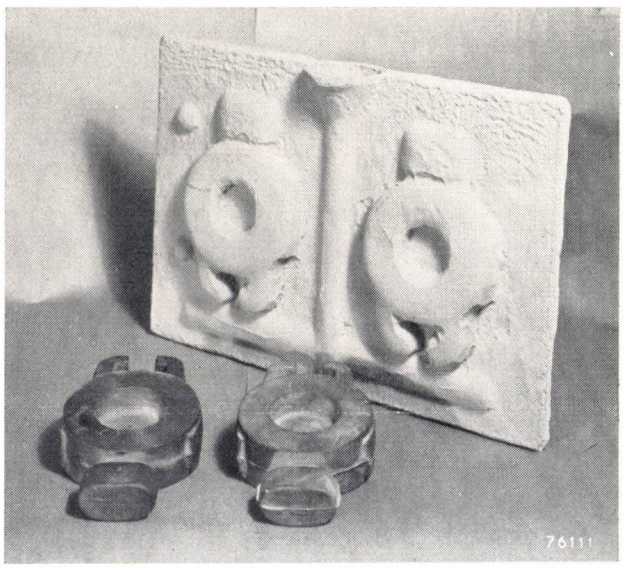


Fig. 3. Half of a double mould for the Croning process. The two halves of the mould are made by dropping a mixture of fine sand and powdered plastic on a heated pattern. In the foreground are pieces cast from such a mould. (From K. Rose, Materials and Methods, Jan. 1953.)

[1]) See for example. "Die casting for engineers", The New Jersey Zinc Co., 1946.

[2]) Field Information Agency (Technical), Final Report 1168, "C"-process of making moulds and cores for foundry use.

halves and each half is mounted on a flat plate. The plate is heated to about 250 °C and then placed face downwards on a box containing a small quantity of moulding powder. This powder is a dry mixture of fine sand and powdered resinous plastic. The box with the plate is then turned over, the powder falls on the hot mould, the plastic melts and causes the grains of sand to cohere. After a few minutes the box is turned up again and the pattern plate, on which a "shell" of coherent sand about $^1/_2$ cm thick is formed, is taken off the box and heated. The resin-sand shell is "cured" by the heating and after some time can be separated from the pattern. Two of these shells are fitted together and in this way form the mould into which the metal can be poured. During casting the mould becomes so hot that after the solidification of the metal the resin will for the greater part be burnt out, so that the remains of the mould can easily be removed from metal. The fine sand gives a clear impression with good detail. The process has proved itself to be very suitable for the mass production of small articles. The United States in particular are attempting to perfect the technique [3]). Many industries are developing the process to extend its use to the most varied products of cast-iron and other metals. The croning process is even being developed for certain high-grade steels, e.g., for gas turbine parts.

Lost-wax casting

The last process that we shall examine here is lost-wax casting. This will be dealt with rather more thoroughly, in view of the investigations which have been made in the Philips laboratories into the possibilities of its utilisation. The process

76001

Fig. 4. Ancient Egyptian bronze (from the collection of the late Dr. A. F. Philips, at Eindhoven). The bronze is about 15 cm high. These bronzes were cast by the lost-wax method. In excavated workshops, the remains of wax patterns have been found.

[3]) R. W. Tindula, Foundry, Cleveland, **80**, 201-210, 1952 (No. 7).

itself is very old. The Sumerians and the Egyptians practised this method already three thousand years before Christ [4]). Decorative objects were first fashioned in wax. These were then covered with wet clay in which a hole ("sprue") had been left. The lump of clay was then dried and afterwards heated, causing the wax to melt away. The molten metal was poured into the cavity thus obtained. Many bronze images and objects that accompanied the Pharaohs on their journey to the realm of the dead were made in this way [4]), and exhibit a great sense of art and great skill (*fig. 4*).

The great value of this process for the plastic arts is obvious if we compare it with sculpture in hewn stone. The free modelling in wax, the ease with which smooth surfaces and fine details can be reproduced, the mechanical strength of work cast in metal and finally the possibility of forming the wax pattern (which is lost during the process — hence the name) by means of a mould, whereby the process is potentially repetitive: all this induced artists in many cases to prefer metal casting to working in marble or stone.

The technique has therefore remained in use throughout the centuries for the making of bronzes. It reached a culmination in the Italian Renaissance. Vasari, in the prologue to his well-known "Lives of Artists" (1550), gives a detailed description of the technique [5]). Benvenuto Cellini (1500-1571) the goldsmith-sculptor, renowned for the splendid bronzes, some of them of great dimensions, which he created by this means (*fig. 5*), has written an extensive record of the methods he employed [5]).

During the last 50 years, however, except for sculptors in bronze, it was only the dentists who used wax impressions, to serve as a pattern for the making of the gold castings for false teeth! Not until about 1930 did the process find any other employment (viz., in the jewellery industry) but about 1940, when the gas turbine made its début in aircraft, the process was introduced on a large scale. The blades of gas turbines (*fig. 6*) have to come up to the highest mechanical and thermal requirements, and the newly developed materials of which they were made were difficult to machine.

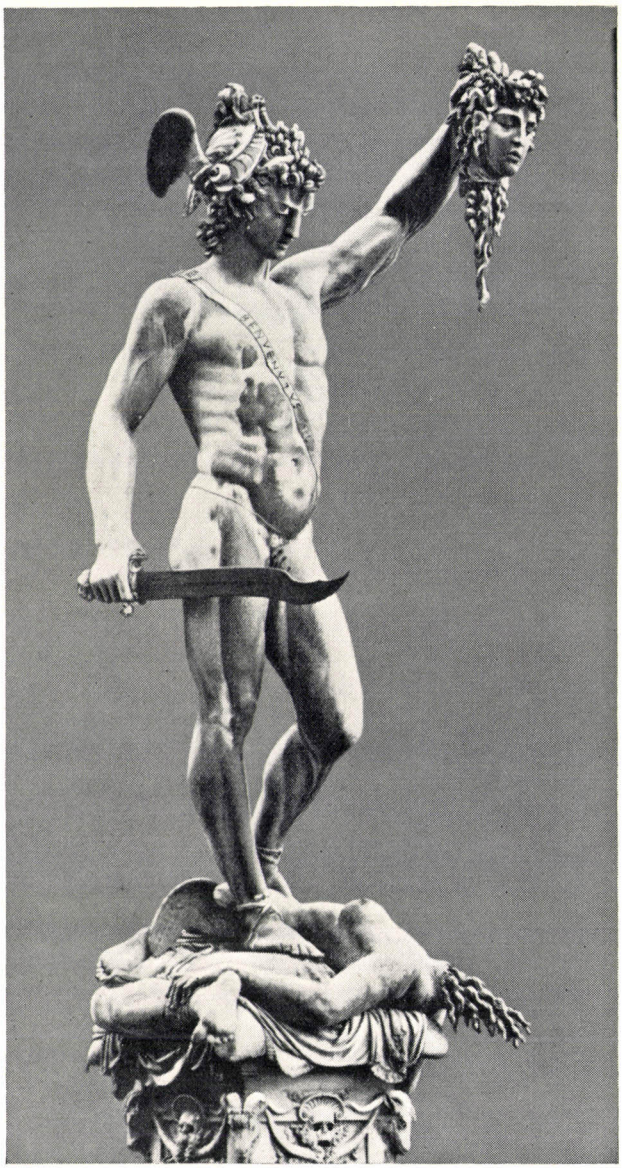


Fig. 5. Benvenuto Cellini's bronze Perseus with Medusa head, in the Loggia dei Lanzi at Florence (completed 1545). The statue displays a great fineness of detail, made possible by the lost-wax casting method. (Block kindly supplied by A. Oosthoek Publishing Co., Utrecht.)

Lost-wax casting was the solution of this problem. At the end of the war there was an American factory with a monthly production of two million blades! It is small wonder that after its successful application to the gas turbine, the process was tried for other products as well, and is now employed in many countries, though still on a limited scale [6]). Some years ago Philips set up a research group to

[4]) See for example: H. Garland and C.O. Bannister, Ancient Egyptian Metallurgy, Griffin, London 1927.
G. Roeder, Die Herstellung von Wachsmodellen zu ägyptischen Bronzefiguren, Z. ägypt. Sprache und Alt. **69**, 45-67, 1933.
V. Gordon Childe, The technique of prehistoric metalwork, Antiquity **22**, 29-33, 1948.

[5]) G. Vasari, Le vite de' più eccellenti architetti, pittori e scultori, Florence 1550.
B. Cellini, Trattati dell' oreficeria e della scultura, Panizzi and Peri, Florence 1568.

[6]) L. G. Daniels, Great savings in replacing complex assemblies by a single investment casting, Metal Progress **56**, 490-491, 1949.
R. L. Wood and D. V. Ludwig, Precision investment castings replace parts produced by other methods, Materials and Methods **32**, 49-53, 1950.
R. Miller, Investment casting, money saving mass producer, Steel, 14 Jan. 1952.

study this casting technique and put it into practice. The process has now been in use for some time on a factory scale. The photos *fig. 7a* and *b* give an impression of some of the machine parts made here.

We shall now discuss further the present industrial form of this process, in connection with the manufacture of machine parts, particularly from a mass production point of view.

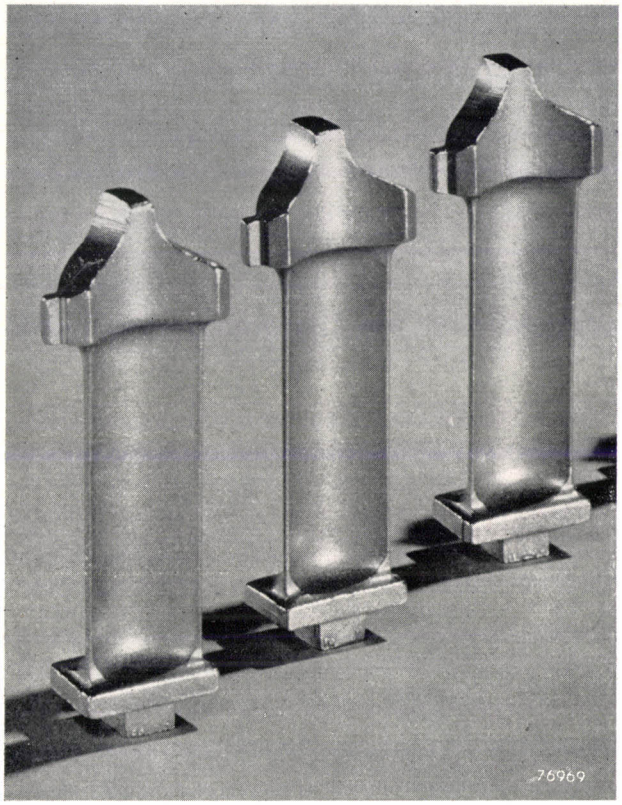


Fig. 6. Blades for gas turbines cast from a heat-resistant alloy by the lost-wax casting process. The blades are about 4 cm long. (By the courtesy of Haynes Stellite Co., Kokomo, Ind., U.S.A.)

The procedure is as follows: a wax pattern is first made of each article to be cast. These patterns are then mounted on a "wax-tree", the trunk of which will serve later on to pour in the metal. The "tree" is mounted vertically on a wax covered plate, and a box or cylinder without ends is placed on the plate and melted to it. The lidless box formed in this way is filled with a "slurry", a ceramic mass that solidifies after some time. (This operation is called investment: hence the term "investment casting".) The box and its contents are then heated, the bottom plate loosens and is removed, and the wax melts and runs out of the mould, which is then dried and baked. The mould is now ready, and the cavity is filled with molten metal. After cooling, the mould is shaken out, the castings are cut off the "tree", cleaned and, if necessary, machined.

After this introductory description we shall examine the various steps in the process more closely and then consider what the advantages and limitations of the process are.

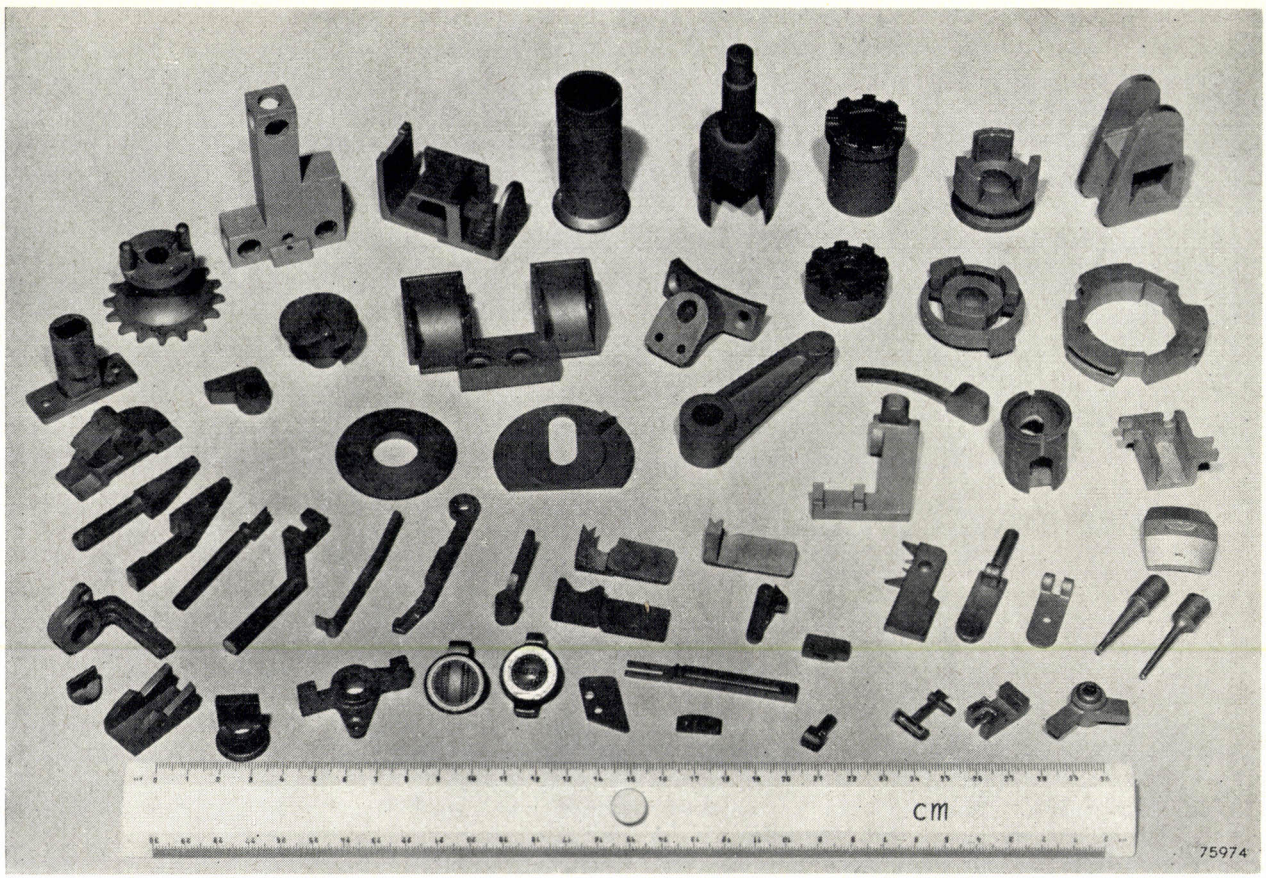
Making the wax pattern

For making the wax pattern we need a metal die. This is filled with liquid wax, and opened, after the wax has set, to allow the pattern to be removed. The wax pattern can also be composed of various parts made by means of separate dies and then simply joined together with a not-too-hot soldering iron. The stipulation that the pattern must be detachable (from the die) now applies only to each part, and by a suitable subdivision of the wax die, this last restriction on the shape of the object can be eliminated. (Figs. 8 and 9 serve to illustrate this.)

A die, consisting usually of two halves (for intricate wax patterns, it may consist of several parts), can be made of steel or brass by the normal manufacturing methods. Another method is to start from a so-called master pattern, a hand-made metal pattern of the object to be cast. This is embedded in a metal of low melting-point (preferably an alloy of bismuth and tin, which undergoes practically no change in dimensions when solidifying). This is carried out in two parts, first the lower half and then the top half, so that the die can be opened (*fig. 8*). The two halves are then pegged and an inlet is drilled for the wax. This method is especially attractive for irregular patterns such as gas turbine blades, but the bismuth-tin alloy is rather brittle, so that damage often occurs and the sharp edges of the die soon crumble. For this reason the dies are often made according to the former method, that is of brass or steel (*fig. 9*). The first is easily machined but rather soft, while steel has the disadvantage of rusting and is more difficult to machine. The better the finish of the die the better the wax pattern will be and hence the better the subsequent casting. It is therefore better to pay greater attention to the making of the dies if quality products are required.

One of the difficulties that does not arise in casting art objects but is at once obvious in the case of industrial castings, lies in the fact that the ultimate product does not possess the exact dimensions of the wax die. There is a certain shrinkage. The wax pattern must therefore be about 1 to 2% larger than the finished product. Unfortunately the shrinkage is not equal in all directions and must therefore often be determined empirically, which entails corrections to the die or to the master pattern. If the dimensions have to be very accurate, as in the case of turbine

a



b

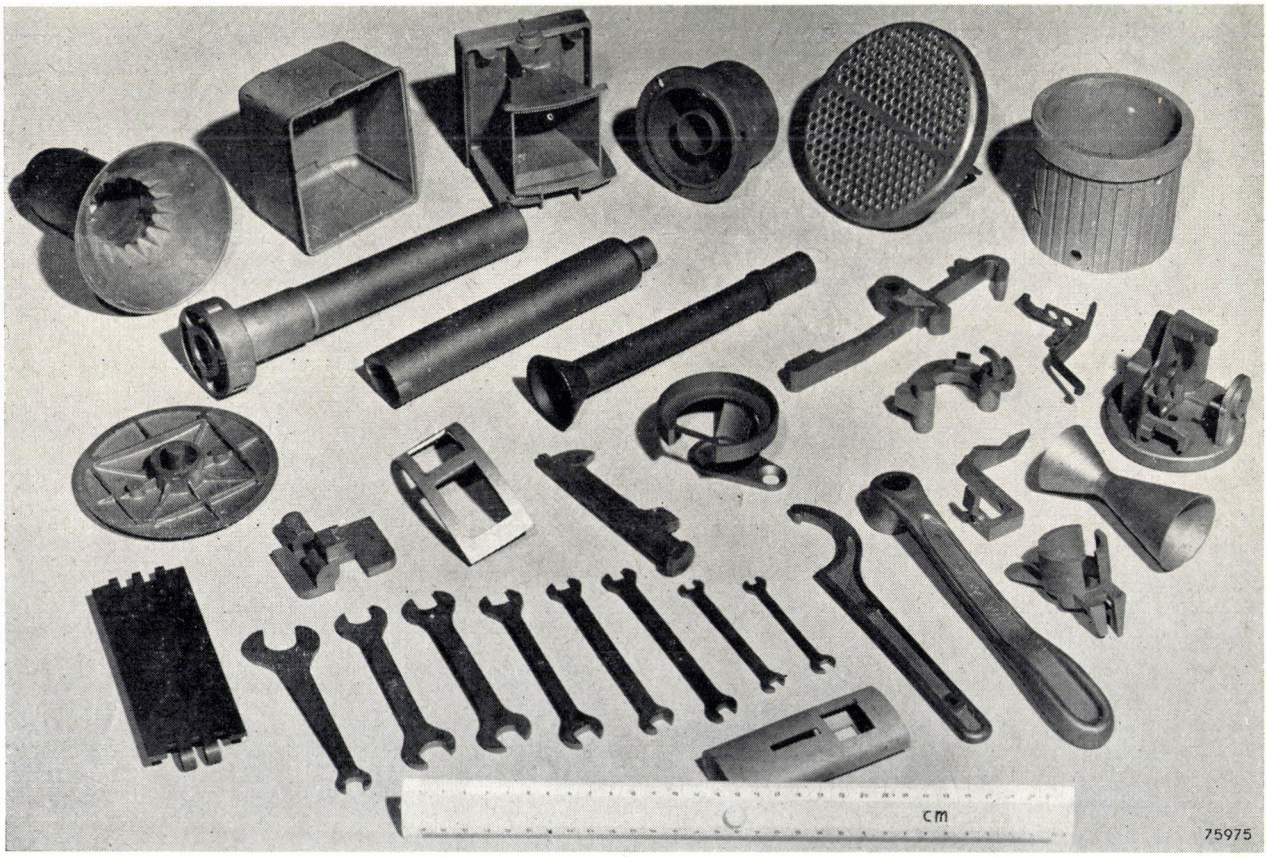


Fig. 7. *a*) and *b*) A seleetion from the series of tools and parts cast in the precision casting department of the Philips laboratories at Eindhoven. None of the pieces shown has been machined. Note the wide variety of shapes and the smooth surfaces. These articles are cast from a number of different alloys.

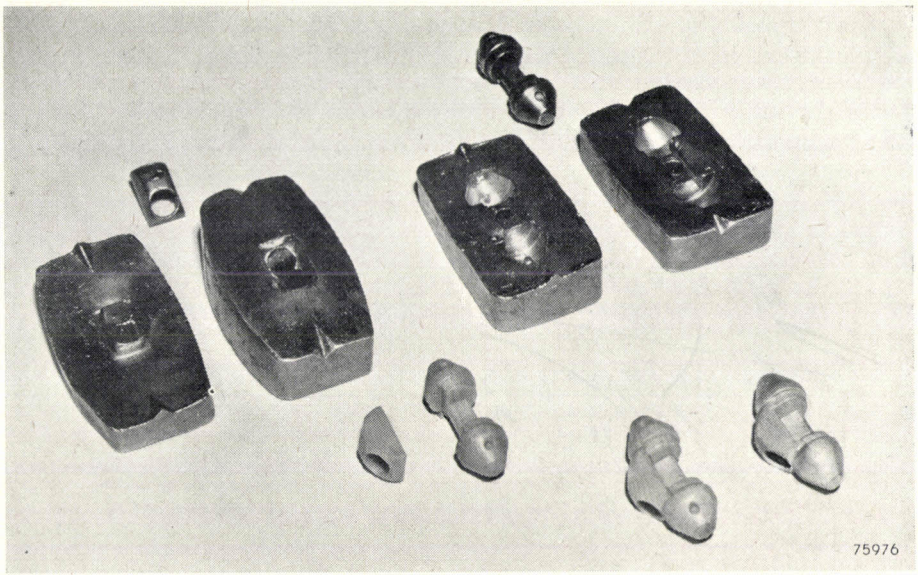


Fig. 8. Wax die of bismuth-tin alloy. Above: the hand-made master pattern, composed of two individual patterns. Below: bottom and top half of each of the two dies made by embedding the two master patterns in the molten bismuth-tin alloy. Further below are the wax patterns made with these two dies, and a complete wax pattern obtained by "soldering" two of these patterns together. At the bottom right is an aluminium casting made with such a wax pattern.

blades, the following procedure is followed: a master pattern is made, from this a wax die and from this latter a trial series of wax patterns which are finished normally into blades. From the results obtained, it can be decided whether the master pattern can be corrected or whether a new pattern must be made. Then a second and, if necessary, a third trial is made, until all requirements are fulfilled and production can begin.

The press which extrudes the wax from a reservoir into the dies (*fig.10*) usually has electric heating elements, automatically controlled to ensure uniform temperature. It is often operated by a compressed air supply at a pressure of 4 to 5 atmospheres.

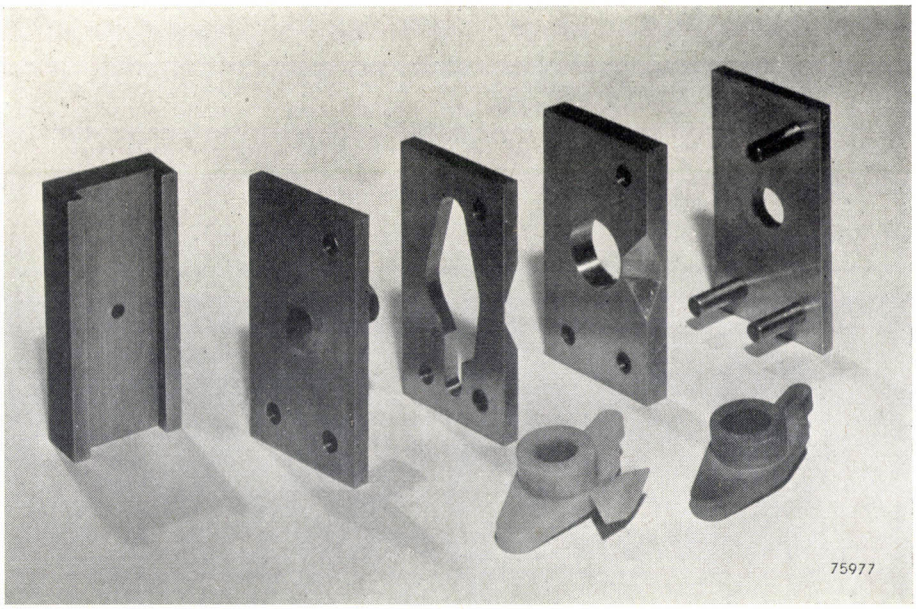


Fig. 9. Wax die of several parts made of brass ("exploded view"). In the foreground is a wax pattern (with gate) made from the wax die, and a casting (chromium steel).

The wax used is usually a mixture of various species of wax and paraffins. The exact composition is chosen with some care so as to obtain a reproducible product. The wax, when solidifying and cooling, must not undergo much shrinkage and must be hard when cold, otherwise the patterns would sag

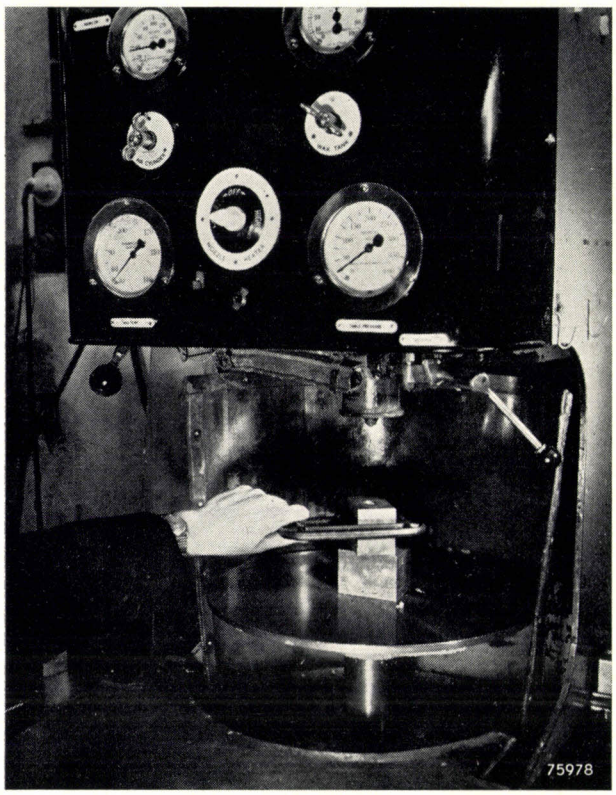


Fig. 10. Wax press. The wax die, its two halves clamped firmly together, is placed on the hydraulically moved table; the die is about to be raised and pressed with its opening against the head of the nozzle, after which it will be filled with wax from a heated reservoir under a pressure of several atmospheres.

even before casting. Instead of wax other materials are also used. Certain plastics (chiefly polystyrene) are often used if a large number of castings is to be made [7]). This material is much stronger, so that the patterns made of it do not easily break, but it has to be pressed in a die under very high pressure. Instead of the simple wax dies that withstand a pressure of five to fifty atmospheres, dies of hardened steel and pressure die machines capable of handling some thousands of atmospheres have to be used. After investing, the plastic patterns are removed from the mould by burning them out.

Frozen mercury is sometimes used instead of wax or plastic. This requires temperatures of — 40 °C (below the melting-point of mercury), the technical problems concerning which have apparently been solved. The advantage of this method, according to the inventor, is a very fine sharp impression of the (steel) die, because the mercury remains liquid when poured in and fills the die well, and has a low shrinkage when solidifying [8]).

Making the mould

The "treeing" of the wax patterns is done fairly simply with the aid of a soldering iron (*fig. 11*). In many cases the wax die can be so constructed that the gate of the mould is already formed, and for small, simple articles, the wax die can be provided with a number of cavities, so that with one injection of wax a number of wax patterns can be made. Generally the same requirements obtain here as in the moulds for sand castings: the metal must flow into the mould as uniformly as possible without sudden strictures or changes of direction that might hinder the steady filling of the mould with metal.

After the tree has been stuck on to the waxed plate and enclosed by the lidless box, the latter is filled (*fig. 12*) with the moulding mass. This is a slurry poured round the wax patterns, which after some time solidifies or "petrifies". When the mould is petrified, the loose bottom is removed and the mould warmed. The purpose of this is to dry the mould and to melt out the wax. A relatively low temperature is sufficient for this. Later on, when the greater part of the moisture has evaporated, the temperature is gradually raised up to 1000-1100 °C to drive out the remaining moisture and to burn out the remnants of the wax in the mould.

For the preparation of the moulding mass or slurry, various materials have been used. Plaster of Paris can be employed, sometimes mixed with quartz powder or other ceramic materials. Often aqueous suspensions of quartz powder and ground refractories, with silicic acid as a binding agent, are used. This transforms the liquid after some hours into a gelatinous mass, binding the suspended particles to a fairly firm mass. Ethyl silicate is usually the basic material, which on decomposition is converted into silicic acid and ethyl alcohol. Sometimes the cheaper alkali silicates are used. Use is also made of the Sorel reaction, based on the action of magnesium oxide (practically insoluble in water) on phosphoric acid or acid phos-

[7]) W. H. Sulzer, Das Präzisionsgussverfahren, Neue Giesserei **37**, 557-565, 1950

[8]) W. I. Neimeyer, Precision casting with frozen mercury patterns, Iron Age **163**, 94-97, 1949.
H. Chane and L. T. Schakenbach, New precision casting process, Materials and Methods **29**, 52-56, 1949.

phates. A slow reaction takes place, and in the course of some hours, basic magnesium phosphates are formed, which take up large amounts of water, so that the slurry solidifies after several hours.

With these materials a rather thin slurry is obtained, which entirely encloses the wax patterns and which, after the escape of the enclosed air bubbles, solidifies. The water disappears in the drying process and now it is important to keep shrinkage and distortion to a minimum. For this reason the box patterns. When the metal is poured in, it would otherwise penetrate the fairly large pores of the baked moulding material, which is to be avoided. If the whole mass were made of so fine grain that no metal can penetrate, the porosity would be too low and the enclosed air and gases would be unable to escape in time. For this reason the wax patterns are dipped or sprayed with a thin coating of a very fine fireclay, dense enough to prevent the metal from penetrating and yet sufficiently permeable to allow

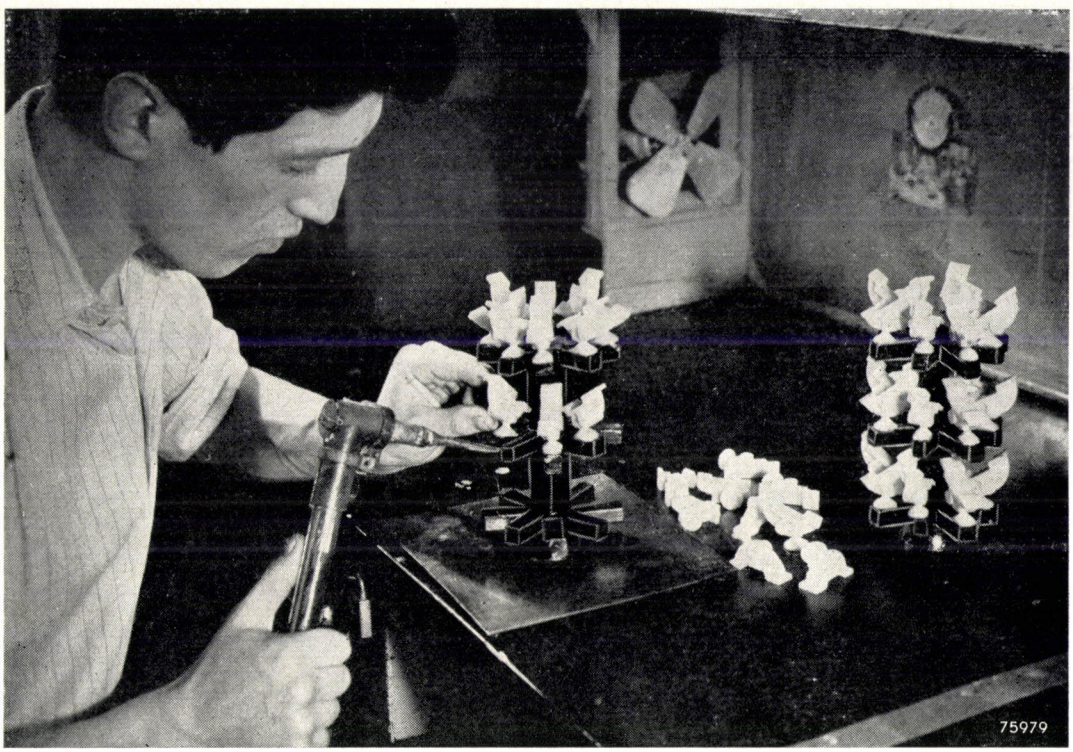


Fig. 11. "Treeing". The "sprues" through which the metal will flow have been formed beforehand (skeleton of dark wax). A great many wax patterns are "soldered" on to this. On the right is a completed "tree".

with the still-liquid mass is usually placed first of all in a vacuum (*fig. 13*), thus enlarging the enclosed air bubbles so that they rise more easily and rapidly. It is then vibrated on a vibration table (fig. 12). This causes the solid particles in the mass to sink, and a dense packing of the grains is obtained before the liquid becomes thick and gelatinous. When the mould is subsequently dried and heated (baked), the coarse grains touch each other and the mould remains practically free from shrinkage. There are of course cavities in the spaces between the grains, so that the mould remains porous when hard.

Before the moulding mass is poured into the boxes, a thin, not very porous coating is applied to the wax the gases to escape rapidly. When this is done the mould is filled, vibrated and allowed to solidify and then, as already mentioned, the wax can be removed and the mould baked.

Filling the mould with metal

After baking, the mould can be filled with metal. It is desirable to keep the mould as warm as possible, so that the inflowing metal cannot solidify against the wall of the mould before it is full. The result is that castings can be made with far thinner walls than if the metal is cast in cold moulds. For castings of thick section, high mould temperatures have disadvantages, for the solidification

takes too long and a casting is obtained of coarse structure, which generally has poor mechanical qualities.

To fill the mould, the molten metal can simply be poured in, but it is better to speed up the process, again with an eye to possible premature local solidifying of the metal. Various methods can be applied for this purpose, e.g. applying pressure on the column of liquid metal, suction on the porous mould or centrifuging the mould. The method used is closely bound up with the method by which the metal is to be melted. For metals with a low melting-point, such as aluminium and magnesium, practically any type of furnace can be used, as the temperature of the melt does not exceed 700 °C. A large quantity of metal can be melted at once and the various moulds filled with a ladle, as is usual with gravity die casting. In the case of iron and steel, the melting temperature is so high that special furnaces have to be used. Generally two types are employed: the electric arc furnace and the high-frequency furnace.

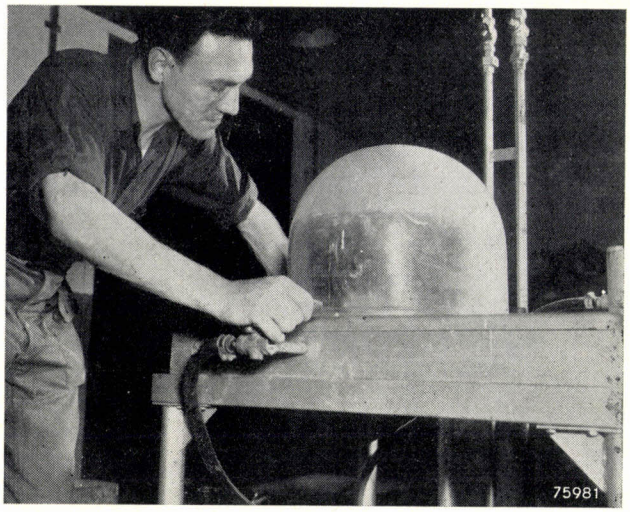


Fig. 13. Before vibration the air bubbles are removed from the moulding mass by placing a vacuum bell over the box and connecting it for a time with a vacuum pump.

Fig. 12. The wax tree has been placed in a box and is being invested with the slurry. A number of boxes already filled, stand together on a vibrating table.

With the first type (*fig. 14*) the melting is performed in a refractory crucible enclosing a bottle-shaped melting space. The metal is fed through the neck. In the melting space, two graphite electrodes protrude, between which an electric arc is drawn. The radiation is reflected by the walls to the material to be melted, and the neck must be narrow to minimize the loss of radiation to the outside. One to three kg of metal can be melted in such a furnace, just enough for one mould as used in the present practice of lost-wax casting. This mould is then placed upside down on the furnace and clamped to it, after which the furnace and mould are inverted. The molten metal runs into the mould and can be forced in by connecting the furnace with a compressed air line.

The high-frequency furnace is slightly simpler in form. In this the crucible is surrounded by a water-cooled copper coil through which a high-frequency alternating current is conducted (*fig. 15* and *16*). The metal in the crucible is heated and fused by induction currents. The apparatus is on the whole more expensive than an electric arc furnace but the crucible can be cylindrical and can therefore be more easily filled with coarser pieces than the rods passing through the neck of the electric arc furnace. Scrap material and cut-off

"gates" from former castings can be fed into this furnace, and in this way save material. Moreover, the electric arc furnace is sometimes undesirable, for the melt is polluted by sparks from the carbon tips. With some materials, for instance stainless steel, this may be prohibitive, as even small quantities of carbon can have a very unfavourable influence on its resistance to corrosion.

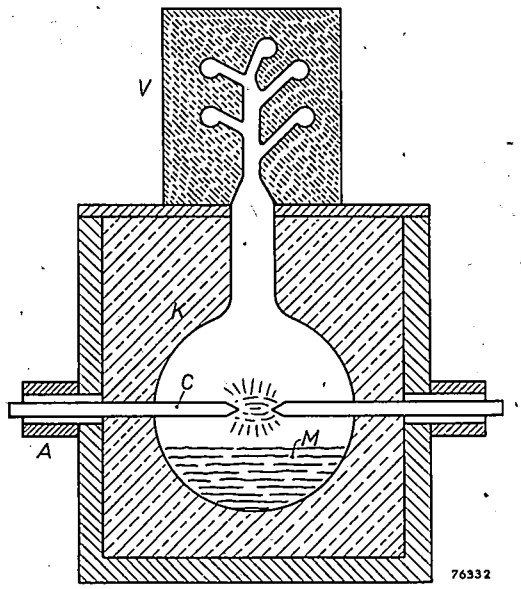


Fig. 14. Schematic section of an electric arc furnace. The metal M at the bottom of the bottle-shaped crucible K is heated to melting-point by the radiation of the arc between the two horizontal carbons C. The mould V is placed on the furnace and clamped to it. The mould is filled by inverting the furnace about the axis A.

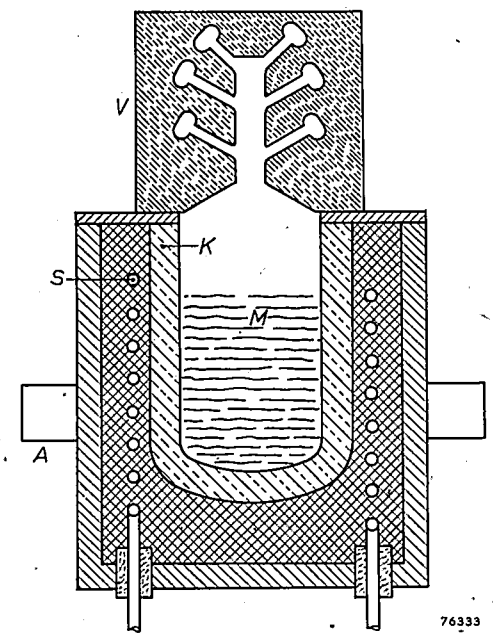


Fig. 15. Schematic section of a high-frequency furnace. The crucible K is here cylindrical. The load melts by the eddy currents induced in it by the high-frequency current sent through coil S. The wide mouth makes this furnace easy to feed.

After casting and cooling, the mould is shaken out, the "gates" are sawn or cut off and the casting cleaned and finished. Usually the finishing is done by sand or shot blasting and, if necessary, the metal may be heat-treated.

Advantages and limitations of the lost-wax process

The articles made in this way are not always ready for use. They often have to be machined in the usual way by turning, milling, drilling or polishing: the dimensional tolerances for modern machine parts are often much closer than can be attained by the lost-wax process. It is therefore a pity that the term "precision casting" has been invented for it, for "precision" is a very relative idea. Compared with normal sand casting, the dimensional fidelity and the quality of the surface are certainly much better, which may be considered great advantages over conventional casting methods. For small dimensions up to 50 mm (2″), tolerances are generally given of ± 0.2 or 0.1 mm (± 0·008 or ± 0·004″), while some manufacturers guarantee ± 0.05 mm (± 0·002″), but this is mostly effected by means of an expensive method of selection in which the samples not falling within the tolerances are rejected. For larger sizes, above 100 mm (4″), a tolerance of ± 0.3 mm to ± 0.5 mm (0·012″ to 0·02″) must generally be accepted. The limitations on the attainable precision are not surprising, having regard for the fact that the ultimate deviation is determined by the sum of a considerable number of factors. The wax pattern shrinks in the die, the dimensions of the cavity are affected by the expansion of the wax in melting and by the changes in dimensions during the drying and solidifying of the mould. The temperature of the mould when casting and the temperature of the cast metal, the shrinkage during solidification and cooling down to room temperature, sometimes obstructed by the resistance of the ceramic mould, all these co-operate in causing variations in dimensions. These can be reduced by working under the most constant conditions with regard to temperature and composition, but the tolerances just mentioned are, in practice, the utmost that can be expected.

In the mechanical workshop much smaller tolerances are implied by the term "precision work", viz., ± 0.005 mm, (0·0002″), or in special cases ± 0.003 mm (0·00012″); in this case, one is inclined

to reject even lost-wax castings with contempt, as the work of a blacksmith! But if the matter is examined more closely it will be noted that these high demands are rarely necessary for *all* the dimensions; a little machining will usually make the casting quite acceptable without much trouble. This is a matter of mutual arrangement between the designer, draughtsman and caster,

Fig. 16. High-frequency furnace, with casting mould attached, being tilted for casting. In the background is the high-frequency generator.

and a great deal of expense can often be saved by making a casting in such a way that it can take the place of a combination of machine-made sub-parts. This is especially so since extra-high demands are made for each of the separately machine-made sub-parts, in order to ensure that the total tolerances are not too great on assembly. (The total tolerances are equal to the sum of the tolerances of the component parts.) It is often cheaper to make one intricate casting and finish it by machining to the required final dimensions.

A second important advantage of the casting method described becomes evident when articles are required with thin walls. Making such pieces from bar stock, from which large portions have to be removed by machining, is generally an expensive business, especially if materials have to be used that are difficult to machine. This is generally the case with wear-resistant and also with many corrosion-resistant alloys, and such materials are becoming more and more used in modern fabrication.

The most important advantage of lost-wax casting and of the modern casting methods in general lies, however, in our opinion, in the fact that now many kinds of alloys can be used which were hitherto out of the question. In general engineering the tendency is, rightly, to try to limit the number of alloys as much as possible. This will be readily understood, for if a certain metal is to be used as a constructional material in industry, there must be a certain stock of various sizes available, from which to draw. It is little wonder that no use can be made — at least not for general factory use — of the hundreds of known aluminium alloys, not to mention the iron and copper alloys of which even more have been described, and the stream of new combinations in which practically all metals can be alloyed. A special alloy can certainly be taken for a special purpose, for instance, cadmium copper for the overhead conductors for electric trains; but for general use, many alloys are not directly available and a compromise is effected by making use as far as possible of those that are in stock. This situation can now be regarded differently. For making castings we can start from the *elements* — or from a limited number of master alloys. A casting can therefore easily be given the composition that best meets the requirements. This means that many alloys that have been developed during years of metallurgical work, but which could not easily be put to general use, can now be exploited. Many alloys with special mechanical, physical, magnetic or electric properties are now available for general engineering.

An instructive example, in which not only the advantage just discussed, but also the "precision" of lost-wax casting plays a part, is seen in the employment of the stellites: hard, very wear- and corrosion- resistant alloys that can be machined only by grinding, which were developed and described about

1910 by Haynes and others [9]). Hitherto it had not been possible to make much use of these, because pieces cast in this material were so rough and inaccurate that machining became far too expensive. By application of the lost-wax casting process, a great future awaits them, as machine parts cast in this way from stellite need practically no machining.

To sum up: it may be stated that the lost wax casting method opens up the following possibilities:

1) The making of small but intricate castings in large or small numbers. In many cases, parts of production machines can be made that can combine various functions in one piece. For metals with a low melting-point, this method is rivalled by pressure die casting for large numbers (i.e., runs of more than 2000), provided the shape is not too complicated; for metals with a high melting-point the Croning process (shell-mould casting) is useful for turning out simple articles in large numbers.
2) It can be said that practically any alloy that is fusible can now also be used for small castings. As casting eliminates the necessity of keeping a wide range of metals and alloys in stock, attention can be paid to selecting those alloys that best meet the requirements of the work. These requirements may be of a mechanical nature (resistance to wear, corrosion and heat, and hardness) or of a physical nature (electric and magnetic properties, thermal expansion, thermal conductivity).

[9]) E. Haynes, Amer. Pat. 873745, December 1907 (!).
E. Haynes, Alloys of cobalt with chromium and other metals, Ind. Eng. Chem. **5**, 189-191, 1913.
W. Oertel and E. Pakulla, Beitrag zur Frage der Kobalt-, Chrom-, Wolfram- (Molybdän)- Legierungen, Stahl und Eisen **44**, 1717-1720, 1924.

Summary. After a brief survey of the development of metal casting, and a description of sand casting and casting in permanent metal moulds (gravity die casting and pressure die casting), two modern processes for the casting of small objects are described, namely the Croning process (shell-mould casting) and the lost-wax process ("precision casting"). A number of details are considered of this last mentioned process, which has been investigated at Eindhoven and used for some years. The process is actually a very ancient one, but has found its way into industry only during the last decade. It is very suitable for making intricate castings and for the employment of many alloys with special properties, hitherto impracticable to use on a factory scale. Added to this is the advantage that the accuracy of dimensions and the quality of the surface of the cast parts are considerably better than, for example, in the case of normal sand casting.

BOOK REVIEWS

Television, by F. Kerkhof and W. Werner, pp. 475, 360 illustrations, 36 photographs — Philips Technical Library — This book can be ordered through your technical bookseller.

Writing a book on television is a rather hazardous enterprise, for in various aspects television technique is still undergoing rapid development. There is a real danger that some parts of a book on this subject will quickly become out of date.

The authors have been well aware of this danger and therefore aimed at focussing attention as much as possible on the physical principles, for these will remain the same — the chance that future television apparatus may be built on other principles is remote. The authors have succeeded in writing a book of lasting value, which will still be useful even when television techniques undergo further development.

The authors' intention has been to address themselves to as wide a group of readers as possible, i.e., to persons of varying education ranging from radio electricians to university graduates. In the opinion of the reviewer, the authors have succeeded in their aim.

The text can be easily followed. The essential properties of circuits are set forth. In many instances, calculations have been inserted in small characters for the benefit of those trained in mathematics; these calculations serve sometimes as the mathematical basis for the main text and sometimes to establish the circuit parameters more precisely. These passages can be skipped without making it difficult to understand the whole.

When reading the book, it becomes clear that the problems of television reception have been gone into more extensively than those on studio- and transmitting apparatus. This is not surprising, since both authors are concerned with television receiver development at Philips. The consequence is that the book is most up to date in those chapters in which reception is discussed; this does not mean, however, that the studio and transmitting apparatus have been neglected. The discussion on these subjects is more simplified and the various associated problems are discussed less exhaustively.

A short survey of the subject-matter is as follows. The first chapter gives a general survey of the method in which the image is transmitted. In the second chapter the behaviour of electrons in magnetic and electric fields is described, while the third chapter treats camera-tubes and picture-tubes.

In the next chapter various transmitting systems are described and an analysis is made as to how the various "informations" needed in television transmission are transmitted and separated at the receiving end. The two chapters which follow, contain discussions on relaxation-oscillators and on deflection-generators. A whole chapter is devoted to H.T. generation for the picture-tube.

The problems relating to the amplification of wide frequency bands are discussed extensively. They are followed by treatments of cable-transmission and antennae, in two different chapters. The next chapter contains a description of the various optical problems which arise in projection television receivers. There is also a chapter on colour television, in which the various systems proposed in the USA by C.B.S. and R.C.A. are discussed. Finally, there is a chapter containing two complete diagrams of television receivers, one for a transmitting system with positive modulation, and one for a transmitting system with negative modulation.

At the end of the book are some appendices, including a list of definitions of terms used in television technique and an extensive bibliography. Sixteen photographs illustrate the most common faults in tuning a TV receiver; the appropriate remedial measures are also given.

The book forms a valuable addition to contemporary television literature. This is particularly the case for the chapters treating modern reception techniques.

J. HAANTJES.

Television Receiver Design, Monograph 1: I.F. Stages, by A. G. W. Uitjens, pp. 177, 114 illustrations. Philips Technical Library, Volume VIII A, Electronic Tubes series, 1953.

This book is the first volume of a specialist series, in which the various problems of television receivers design will be discussed. This series of books in the Philips Technical Library will form a welcome addition to the existing literature. The present volume deals with the intermediate-frequency amplifier of a television superhet receiver. The subject matter is, of course, equally applicable to the high-frequency amplifier of non-superhet receivers.

As may be expected in a book of the Electronic Tube series, the characteristics of the amplifier valves used are extensively dealt with.

In the first chapter the product *GB* (product of stage gain and band width) is introduced, and the correction factors to be applied in practical circuits are deduced.

Two-pole interstage coupling impedances are discussed in the second chapter. The reader is introduced to the cascade coupling of simple *L-C* circuits tuned to the same frequency and to amplifiers with *L-C* circuits tuned to slightly differing frequencies (staggered circuits).

In the third chapter the influence of frequency and phase characteristics on the step function response is discussed. The treatment includes not only symmetric side-band systems but also the asymmetric sideband systems generally used in television.

Four-pole interstage couplings are described in Chapt. IV; author introduces the term, "transfer admittance", which permits the use of the same mathematical treatment as for two-pole coupling impedances.

A very important feature of an amplifier is the signal-to-noise ratio. More than 20 pages of the fifth chapter are devoted to this subject. The various noise sources are discussed and some methods of calculating the noise factor are presented. The means by which this ratio may be improved is also discussed.

The author devotes the thirty odd pages of the sixth chapter to the discussion of the effects of stray couplings in the interstage networks and in the valves. The measures taken to minimize these effects are outlined.

The final chapter deals with the design of an I.F. amplifier with staggered circuits, making use of the formulae given in the previous chapters.

A number of appendices contain both the derivation of certain important general formulae and some convenient tables giving valve and circuit data.

W. WERNER

THE TROPOSPHERE AS A MEDIUM FOR THE PROPAGATION OF RADIO WAVES - I

by H. BREMMER.

621.396.11:551.510.52

*The hypothesis of Kennelly and Heaviside (1902) that an ionized layer existed in the atmosphere, which reflected radio waves and in this way explained satisfactorily why these waves are able to follow the earth's curvature — was confirmed experimentally in 1924 by Appleton. His experiments may thus be regarded as the beginning of research on the ionosphere. This has proved to be very important both theoretically and practically and it is now being pursued vigourously in many places all over the world. The methods used in this field of research have been reviewed in this journal by Prof. C. J. Bakker *).*

*Now that waves shorter than 10 metres are so widely used (for television, radar, etc.) it has become clear that for the propagation of these "microwaves" **), it is not the ionosphere that is important but the lowest layer of the atmosphere. This layer extends from the earth's surface to about 10 km above it, and is called the troposphere. Its behaviour in respect of radio propagation is closely related to meteorological conditions.*

Research on the troposphere is in full swing. Although it has not yet achieved a satisfying whole, as is the case for ionospheric research, it has gone far enough for us to be able to give a review of what is now known, especially in relation to radio wave propagation. The first part of this review follows.

Introduction

The simplest model for describing the propagation of radio waves around the earth is a perfect sphere surrounded by a uniform atmosphere. Radio waves can then only be propagated by means of so-called ground-waves. Even in this simple case, the calculation of the electromagnetic field is a complex problem. (It has already been discussed in this journal [1]). The results show that even for the shortest wavelengths used in practice — centimetre waves — the ground-waves extend to beyond the horizon of the transmitter [2]), and their range increases with the wavelength, but except for very long waves it is much less than the actual transmission range attained in practice. This is a consequence of the fact that all waves longer than about 10 metres are refracted in the ionosphere, between 70 and 400 kilometres above the earth's surface, and so may be deflected back towards the earth. However, the still shorter waves which have become extremely important for radar, television and transmitters with frequency modulation, are unaffected by the ionosphere. Instead they are refracted in the *troposphere:* any such waves reaching a receiver on the earth's surface have travelled exclusively through the troposphere. This is the name given by meteorologists to the lowest layers of the atmosphere, where the temperature decreases with height, down to a minimum of about —50° C at the "tropopause", at an altitude of about 10 km.

Microwave transmission can be effected only via the troposphere. In this respect microwaves behave like medium waves (about 200 m $<\lambda<$ 1000 m) during daylight. The ionosphere absorbs the medium waves during daylight, whilst it is always transparent to microwaves, so that in both cases, the troposphere is the only medium through which communication is possible.

In spite of this similarity, microwaves differ in behaviour from medium waves in two important respects.

1) Microwaves are much more closely related to optical waves, and like these, have only low intensities in shadow regions, such as beyond the horizon or behind a mountain. These are regions that the energy cannot reach if propagated according to geometrical optics, but can do so by diffraction. It would be exceptional, however, for the range of microwaves to extend to double the distance from the transmitter to its horizon.

2) Microwaves are far more sensitive to irregularities in the troposphere. This is understandable, for these irregularities are now more likely to be of the same order of magnitude as the wavelengths. In contrast, the behaviour of the longer waves depends rather on the average value, over one wavelength, of the irregularities, and is therefore less sensitive to them.

---

*) C. J. Bakker, Radio investigation of the ionosphere, Philips tech. Rev. 8, 111-120, 1946.

**) For the purposes of this article, the usual interpretation of the term "microwaves" has been somewhat extended to include all waves of $\lambda < 10$ m.

[1]) B. van der Pol and H. Bremmer, The propagation of wireless waves round the earth, Philips tech. Rev. 4, 245-253, 1939.

[2]) Called "the horizon" throughout, for short.

Irregularities in the troposphere are always important in the propagation of microwaves. It is therefore not surprising that the behaviour of these waves is closely bound up with meteorological conditions. Their behaviour depends both on mean climatic or seasonal variations over a long time, and on short time variations (those of synoptic meteorology). Hence the properties of microwaves are not so easily surveyed as those of longer waves, for which the distribution of land and sea, for example, and daily fluctuations of temperature and humidity are not so significant in their effects. Many observations on both commercial and experimental transmissions have been made in order to gain more insight into these effects. It now seems clear that fluctuations in the atmosphere over minute distances — of the order of centimetres — are important, with which meteorologists were little concerned until very recently. Thus the investigation of microwaves has stimulated new meteorological research.

It is proposed to review here the theories put forward to explain the widely varying propagation characteristics of microwaves. The physical properties of the troposphere as medium for microwave propagation are first discussed.

The troposphere as a dielectric

The only physical quantity that is important for radio propagation is the refractive index n, which can equally well be written as $\sqrt{\varepsilon}$, where ε is the dielectric constant with respect to a vacuum, as conduction currents in the troposphere are negligible. We shall not yet consider how n varies from point to point, but only explain how n can be calculated from the physical properties of the gases in the troposphere.

The theory allows for the molecular polarization of each of the gases under the influence of electromagnetic waves. The theory states that not n itself, but $(n^2-1)/(n^2+2)$ is additive, i.e., the refractive index n_m of the mixture is given by:

$$\frac{n^2_m-1}{n^2_m+2}=C_1\frac{n_1^2-1}{n_1^2+2}+C_2\frac{n_2^2-1}{n_2^2+2}+\ldots$$

where C_1, C_2, ... are the volume concentrations of the component gases $(C_1+C_2+\ldots=1)$, having refractive indices n_1, n_2, ... Now, for each of the gases, n differs only slightly from 1; the factor $(n+1)/(n^2+2)$ is therefore almost constant. Hence, for the mixture, a good approximation to n is given simply by summing values of $n-1$, which is very small compared with 1. Moreover, it appears that for wavelengths $\lambda>3$ cm, absorption effects can be neglected for all important constituents of the atmosphere (nitrogen, oxygen and water vapour). The refractive index is then real, and it is found that for microwaves with $\lambda>3$ cm it can be represented by:

$$n=1+79\times10^{-6}\frac{p_a}{T}+0.38\frac{p_w}{T^2}. \quad \ldots \quad (1)$$

Here T is the absolute temperature, and p_a and p_w are the partial pressures in millibars of dry air and water vapour respectively. This formula was derived theoretically and has been verified experimentally.

It is worth noting that the refractive index is rather sensitive to the amount of water vapour present, since the dielectric constant of water vapour is relatively high — the strong polarizability of the molecules of water falls off rapidly only at wavelengths shorter than those of microwaves. Further we note from (1) that n is independent of λ, so that the troposphere, unlike the ionosphere, causes no dispersion.

For wavelengths shorter than 3 cm, deviations from equation (1) begin to come in through absorption effects. Ultimately the absorption renders the use of very short wavelengths impossible for telecommunication, even though it is technically perfectly practicable to produce such waves.

For decreasing wavelengths, the absorption first perceptible is due to water (as a liquid) present in the atmosphere as rain, cloud or mist. The disturbance due to the water droplets can be calculated from the scattering and absorption in individual droplets. It is not so large as for an equal mass of water vapour. Both the absorption and the scattering can be derived, to a good approximation, from the dipole moment that is induced in each droplet by the field of the radio waves reaching it.

For water droplets in clouds and mist the calculations can be greatly simplified, for the droplets (max. diameter about 0.2 mm) are small compared with the wavelength. The theory is then the same as that for the scattering and absorption of light due to inhomogeneities in the atmosphere, whereby Rayleigh explained the blue of the sky. The amount of scattering depends mainly upon the *total volume* of the inhomogeneities, and not on their distribution as individual droplets, always provided they are all small compared with the wavelength.

Unlike light waves, radio waves are attenuated more by *absorption* in water droplets than by scattering. This absorption appears to be less than that for a coherent mass of water of equal volume.

When centimetre waves are absorbed by raindrops, the absorption is no longer independent of the size distribution of the drops, for they are not then small enough compared with the wavelengths. When this distribution is known, the absorption can still be calculated, although less simply. Rain can disturb the transmission of wavelengths shorter than 3 cm. For longer wavelengths than this, hail and snow absorb much less than the same amount of rain, but this is not necessarily so for shorter wavelengths.

For $\lambda<2$ cm there is an appreciable absorption not only by water but also by the gases of the atmosphere. There is a resonance band near $\lambda = 0.5$ cm, and a resonance line near $\lambda = 0.25$ cm, both due to oxygen, and water vapour shows a single line close to $\lambda = 1$ cm. Both the band and the lines have been calculated theoretically [3]) and demonstrated experimentally. The propagation of the still shorter millimetre waves is rendered well-nigh impossible by strong absorption lines of water-vapour, but of course they can be used for communication over very short distances. The resonance band of oxygen has even been exploited to prevent a message sent over a short distance from being overheard by unauthorized persons further away.

In what follows, only waves longer than 3 cm will be discussed. For these waves the refractive index is given accurately enough by the simple equation (1).

Survey of idealized cases of radio propagation through the troposphere

The all-important refractive index at any point in the atmosphere depends, broadly speaking, on the altitude. The average values of n at different heights depend on the season and are closely related to the climate. At any one time and place, the change with height deviates from its mean value owing to the day-to-day changes in the weather, and these cause the index to vary from place to place at any one height. The result is that widely varying situations are possible, and these can influence the wave propagation to the good or the bad. We can simplify our description of what happens by enumerating several idealized types of wave propagation. Actually the waves will always be propagated in several ways at the same time.

In all these cases except the last, it is assumed that the refractive index depends only on the height.

The waves, then, may reach the receiver from the transmitter in the following ways:

a) By paths determined by laws of optical reflection at the earth and refraction in the atmosphere (*fig. 1a*).

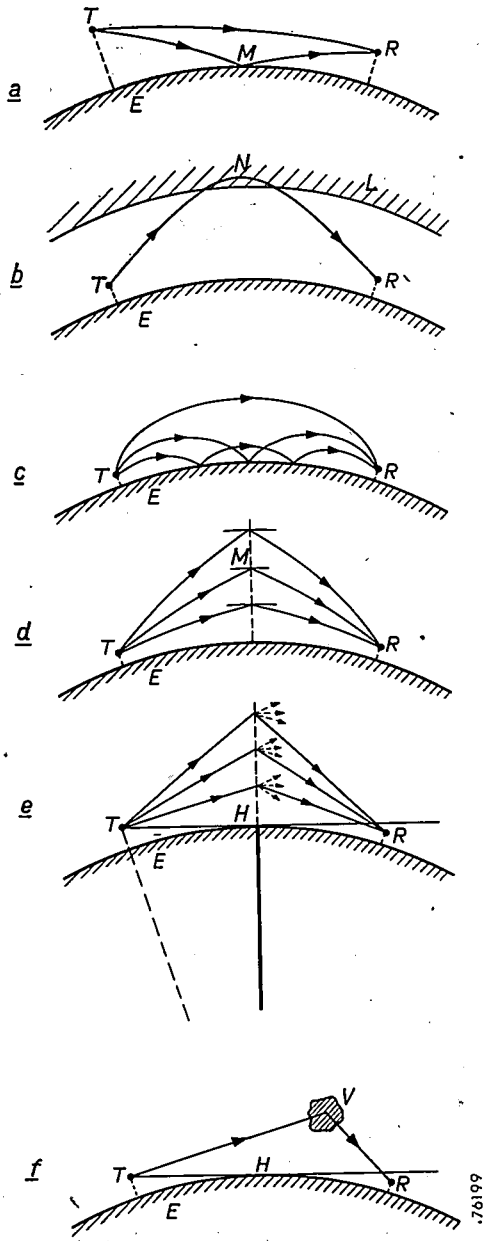


Fig. 1. Schematic representation of the different ways in which microwaves can reach a receiver R. T transmitter. E earth (curvature relatively greatly exaggerated).
a) Propagation along two paths, one direct (TR) and one reflected from the earth's surface.
b) Rays being continuously deflected in a transition layer (L).
c) Propagation by very many rays, as in a waveguide (super refraction).
d) Propagation via gradient reflections.
e) Although the receiver is below the horizon (H), rays reach R by diffraction, for example at the vertical plane through H.
f) R again below the horizon: rays reach R as a result of scattering at an inhomogeneity V in the troposphere.

[3]) See for example J. H. van Vleck, The absorption of microwaves by oxygen, Phys. Rev. **71**, 413-424, 1947, and The absorption of microwaves by uncondensed water vapor, Phys. Rev. **71**, 425-433, 1947.

b) By paths determined by the laws of geometrical optics, but occurring only under exceptional meteorological conditions (transition layers, fig. 1*b*); only a limited number of rays can reach the receiver.

c) By paths also determined by a transition layer, but which is due to other unusual conditions, very many rays may end at the receiver: they are propagated as if in a waveguide (fig. 1*c*).

d) By reflection at the thin "layers" of which the troposphere is composed, owing to its continuously changing refractive index (gradient reflections, fig. 1*d*).

e) By such paths that the field at the receiving end can be accounted for only by diffraction (fig. 1*e*).

f) By scattering at inhomogeneities in the troposphere, i.e., short-distance fluctuations in the refractive index both in a vertical and in a horizontal direction (fig. 1*f*).

Before discussing these cases in full, a few more general comments may be made. The transition layers in (b) and (c) may be regarded as surfaces of discontinuity if the refractive index changes within a height-interval less than one wavelength; reflection then occurs. In the more common case of a transition layer thicker than one wavelength, a ray progressing upwards (or downwards) can be bent continuously throughout its path within the layer, and so much so that it leaves the layer on the same side (*fig.* 2). In this case we speak of

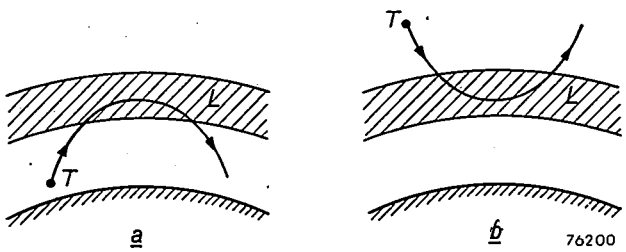


Fig. 2*a*) "Continuous reflection" in a transition layer *L*; rays reaching the layer from a transmitter *T* below *L* are bent back towards the earth.
b) *T* is above the layer *L*, which now bends the waves away from the earth.

continuous reflection at a layer, although the effect is in fact produced by refraction. Similar reflections occur when short waves are deflected from the ionosphere, and when sound travels a great distance (the sound waves are deflected at the ozone layer (ozonosphere), at an altitude of about 30 km).

An important question is — how large is the horizontal area over which the refraction index must be constant if it is not to disturb the paths based on a refractive index varying only with altitude? Suppose that the paths between transmitter *T* and receiver *R* are determined by geometrical optics. From the wave theory, whether or not a disturbance is caused by a local deviation in n (at *P*, say) depends on the path-length *TPR*; a disturbance is caused only if this path-length differs from the direct path (not necessarily straight) by less than $\frac{1}{2}\lambda$. The points *P* that satisfy this condition, lie within the "first Fresnel zone" with respect to *TR*.

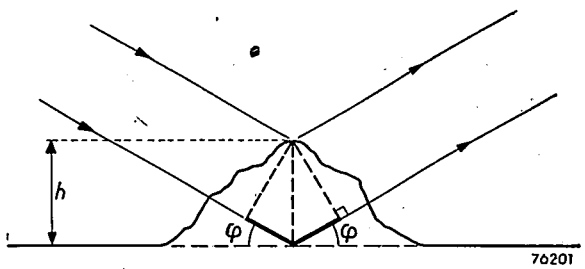


Fig. 3. The heavy line is the path difference between rays reflected at the top and the foot of the hill respectively, equal to $2h \sin\varphi$, when φ is the angle between the rays and the surface.

A geometrical study shows that if the maximum height of the ray is small compared with the distance *D* between *T* and *R*, the zone is concentrated within a narrow region about the vertical plane through *TR*; it extends only a distance $\frac{1}{2}\sqrt{D\lambda}$ from it. Thus, at points further from this plane, fluctuations in the refractive index are unimportant: for $D =$ 100 km, for example, and $\lambda = 5$ cm, the critical distance is only 354 metres.

This narrowness of the first Fresnel zone is important, too, in relation to rays reflected at the earth's surface. In unbroken country we can regard the reflection as occurring on a plane, provided that over this plane the properties of the ground are uniform within the Fresnel zone. This zone, again, is a narrow one along the shortest path between *T* and *R* along the earth's surface, and extends for distances of about $\sqrt{D\lambda}$ on either side of the path. However, suppose a hill or other object is near the line *TR*, with a height *h* (*fig. 3*). Rays reflected from its foot and top respectively have a path difference equal to $2h \sin\varphi$, where φ is the angle of reflection (i.e. between the rays and the earth's surface). If this path difference exceeds the order of $\frac{1}{2}\lambda$, i.e. $\lambda < 4h \sin\varphi$, the reflection will be considerably modified — usually diminished.

a) Communication along the two normal paths

Whenever the receiver is above the horizon of the transmitter, it receives rays from both normal paths — the direct ray and the ray reflected at the

earth's surface. Both are curved since the troposphere is not homogeneous. If we now allow for the curvature of the earth and still assume that the refractive index n depends solely on the altitude, then n is a function of the distance r to the centre of the earth. It then follows from Snel's law of refraction that the product $r\, n(r) \sin \tau(r)$ is constant along each ray, where τ is the angle between the tangent to the ray at the point in question and the vertical through that point (*fig. 4*).

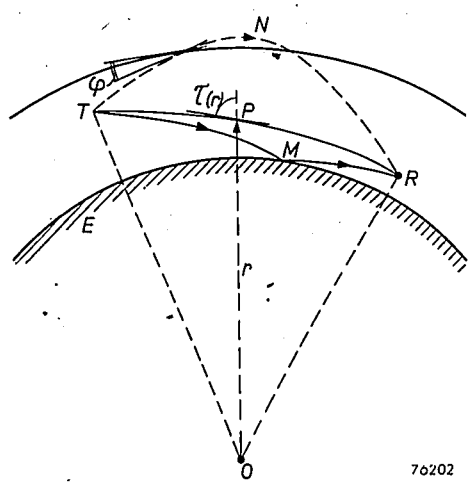


Fig. 4. Rays corresponding to fig. 1*a* (full curve) and fig. 1*b* (dashed curve). r distance between an arbitrary point P on the path and the centre of the earth. $\tau(r)$ angle between the tangent to the path at P and the vertical through P. φ = angle between an incident ray and the transition layer L.

Under normal conditions, $n(r)$ decreases with increasing height much more slowly than r increases, so that r dominates in the product $r\, n(r)$. According to Snel's law, therefore the angle $\tau(r)$ decreases along a rising path, and increases along a falling one. This is illustrated schematically in fig. 4. The propagation along TR is very much as it would be in free space, although the path is slightly curved. The field strength at R is therefore about inversely proportional to the length along the curve, and if the distance D along the earth's surface is much greater than the altitudes of the transmitter and receiver it is also inversely proportional to D.

The curvature of the shells with constant n, which are of course concentric to the earth's surface, causes slight deviations from this simple rule, and these make the paths diverge less for downward rays and more for upward ones, than would be the case in a homogeneous troposphere.

The field along the ray reflected at the earth, TMR, is also about inversely proportional to the distance along the path from T, if we ignore the effect on the intensity of the reflection at M. For all but extremely short waves, this reflection, too, can be regarded as coming from a flat surface (see end of previous section). The corresponding reflection coefficient can be derived from Fresnel's formula, which is based on Maxwell's equations. For microwaves, this reflection coefficient proves to be very close to —1, since the waves almost graze the earth's surface. This means that the reflection is as good as total, but it is accompanied by a phase change of close to π radians. However, for waves shorter than 10 cm, the earth's surface is not usually smooth enough to be regarded as flat, and the theory can no longer be applied. The reflection is more like scattering at a rough surface, giving an effective reflection coefficient of the order of 0.2.

The total field intensity at the receiver is found by adding together the fields produced by TR and TMR and allowing for their difference in phase. The latter is partly due to the difference between the two path lengths, and partly to the phase shift of nearly π where the second ray is reflected (when this can be regarded as reflected at a flat surface). The phase difference means that, at a fixed height, the field intensities vary with distance. These interference effects are of course much weaker when the reflecting surface at M is "rough", for then the field due to TMR is quantitatively unimportant.

b) Communication via transition layers

The presence of a large gradient of the refractive index in the transition layers is contingent on the meteorological conditions. Such a layer is usually accompanied by an inversion (i.e. the temperature *rises* with height), or at least by a much smaller fall of temperature with height (lapse rate) than the average. It is not so much the temperature itself, but the corresponding changes in pressure of water vapour that are mainly responsible for an abnormal change in the refractive index (see equation 1). These transition layers are found especially at heights between 500 and 2000 metres, and the change in n across the layers is often of the order of 5×10^{-6} (while n—1 is round about 3×10^{-4}, see above). Slowly-moving transition layers often occur in settled weather; on the other hand, layers may occur that extend along meteorological fronts and move with them.

When transition layers are present, propagation along paths like the dashed line TNR in fig. 4 is possible. This kind of path is especially important when the normal paths TR and TMR are impracticable because the receiver R is below the horizon; apart from this, R can be reached only by diffraction, and the corresponding field at R is then weak. In

particular, it has been established that the field is temporarily strengthened when a meteorological front passes between transmitter and receiver. Several transition layers can be operating at the same time [4]).

The field along a path via a transition layer is, once again, about inversely proportional to the distance from the transmitter, but the effect of reflection at the layer must be allowed for. Two limiting cases will be discussed in which it is easy to give a general idea of this effect: these are, when the layer is much *thinner* than the wavelength and much *thicker*.

Layer thickness small compared with the wavelength

For thin layers the conditions are nearly those of reflection at a surface of discontinuity of the refractive index, so that the coefficient of reflection may be calculated from Fresnel's theory. The angle of incidence φ on the layer (see fig. 4) is small, yet it is larger than the critical angle (below which the reflection is total), and because of this the reflection is partial: part of the energy of the incident ray gets through to the space above the layer. Further, the angle φ is small compared with Brewster's angle (that is the angle for which there is zero reflection of waves with electric vectors in the vertical plane through the incident ray). Because of this, a good and simple approximation to the reflection coefficient R proves to be

$$R \approx \frac{\Delta n}{2n \sin^2 \varphi}.$$

Here Δn is the difference between the refractive indices on the two sides of the layer. Thus the reflection coefficient does not depend on how n changes inside the thin layer. When the distance D is large compared with the height h of the layer above the earth, φ and $\sin \varphi$ can both be replaced by $2h/D$, and then

$$R \approx \frac{\Delta n}{8n} \cdot \frac{D^2}{h^2} \quad \ldots\ldots \quad (2)$$

from which it is clear that a given layer reflects best when the transmission is over great distances. This is understandable, for then the incident angles may approach the critical angle for total reflection.

Layer thickness large compared with the wavelength

The reflection at thick layers can be said to be "continuous" and appears to be total, i.e. no rays pass through the layer. However, according to the theory, the continuous reflections are associated with an extra phase shift of $-\frac{1}{2}\pi$, which can be compared with similar phase shifts in optics.

Actually, all intermediate layers between the two extreme ones are possible. Their reflection coefficients have been calculated by making assumptions about the way n changes through the layer, and in this way Epstein [5]) and Rawer [6]) have demonstrated that the reflection is generally better for the longer wavelengths.

The fields due to reflections at one or more transition layers must be compounded vectorially with the field due to the two normal rays (discussed under *a*), provided of course these exist, i.e. the receiver is above the horizon. What is observed is often so complex that it is difficult to say whether or not transition layers play any part. At great distances they can be recognized by the increase of the reflection coefficient with distance under eq. (2), which can partly compensate for the normal weakening of the field. Apart from this, the transition layer may be revealed by an increase of the field for the longer wavelengths, due to the greater reflection coefficient of the layer, but this can be masked by less reflection of the same wavelengths at the earth's surface. Total reflections at transition layers can be observed only for low-altitude layers or by observations from the air, in which cases the angles of incidence φ can be very small.

c) Cases where the troposphere acts as a waveguide

In general, ascending waves escape into space, but this does not necessarily happen to all rays. For example, some rays may be deflected back by continuous reflection, and then remain confined to the troposphere by reflection between the earth's surface and the highest "continuous reflecting" surface. The space in which the rays travel to and fro between the earth's surface and the continuous reflecting surface has all the typical properties of a waveguide [7]). The theory of this effect was given by Booker. The conditions for the existence of such a waveguide can be derived as follows. A ray undergoing continuous reflection must attain a maximum height somewhere, and there its tangent is horizontal and τ is 90°. Since by Snel's law, $rn \sin \tau$ is constant, τ can become 90°

[4]) For a survey of the effect of transition layers, see an article by J. A. Saxton and others, Proc. Inst. El. Engrs. **98**, III, 360-378, 1951.

[5]) Proc. Nat. Acad. Sci. **16**, 627-637, 1930.

[6]) Ann. Physik **35**, 402, 1939.

[7]) For the properties of waveguides, see e.g. articles by W. Opechowski in Philips tech. Rev. **10**, 13-25 and 46-54, 1948.

only at that point on the path where rn has the smallest value. There must therefore be a region where rn decreases over a rising path. Such a region is called a duct, and it is not often met with, for usually r increases more than n decreases. In the exceptional cases where n decreases strongly enough for rn to decrease and produce a duct, we speak of super-refraction.

The presence of a duct does not mean that *all* rays going upwards are bent downwards again by continuous reflection. To consider a ray that is deflected back, let r_0 and n_0 be respectively the values of the distance to the earth's centre and the refractive index at the highest point in its path, and a and n_a the corresponding values at the earth's surface. Then for a ray leaving the earth's surface at an inclination τ_a to the vertical, Snel's law gives

$$a\, n_a \sin \tau_a = r_0 n_0.$$

When rn is a minimum (denoted by $(rn)_{\min}$), i.e. on the upper side of the duct, τ_l, the inclination, to the vertical of the steepest ray to be reflected, is given by

$$a\, n_a \sin \tau_l = (rn)_{\min}.$$

A similar limiting angle is also known in propagation through the ionosphere. Only the less steeply rising rays can be continuously reflected back towards the earth from within the ionosphere. But continuous refraction in the ionosphere and super-refraction in the troposphere differ in several respects, as follows:

1) In the ionosphere, continuous reflection and a critical angle for it are quite normal, whilst in the troposphere, super-refraction only occurs under exceptional meteorological conditions [8]). Over the sea, conditions are often favourable for super-refraction, due to the relative humidity falling off quickly with height. Over the land, the cooling of the surface at night is very favourable for super-refraction, when a layer of cold air is formed over the ground. Here also the decrease of humidity with increasing height is very important: in dry air super-refraction is only possible in a temperature inversion with rise of more than 11.2 °C in 100 m, which is highly improbable, whilst a much smaller temperature gradient suffices in damp air. (In contrast to this, acoustic super-refraction can occur with nearly every temperature inversion.)

2) In the ionosphere, because of dispersion, the limiting angle depends very much on the frequency, but this is not so for super-refraction.

3) In propagation via the ionosphere there is a dead zone (skip distance) around the transmitter, and rays can reach the earth through continuous reflection only outside this zone. The receiver R can be reached only by a few rays in this way. For super-refraction in the troposphere, on the other hand, when the duct lies just above the earth's surface, very many rays (in theory infinitely many) go from the transmitter to the receiver according to Snel's law. The rays differ in the number of reflections undergone. This great number of rays is possible because there is no homogeneous region between the earth and the duct — as there is effectively between the earth and the ionosphere. As a result, the distance from the transmitter to the point where the ray first returns to the earth's surface can decrease only when the angle τ_a increases (*fig. 5a*) — in contrast to the case of ionospheric reflection.

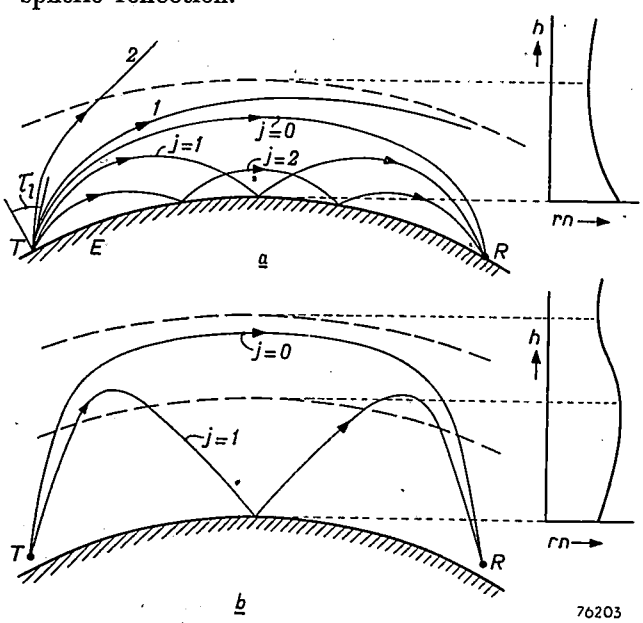


Fig. 5. Cases of super-refraction. There is sometimes a region in the troposphere ("duct") within which rn decreases with rising height, either (*a*) from the ground upwards ("surface duct"), or (*b*) only above a certain height ("elevated duct"). Many rays can now go from the transmitter to the receiver, and they undergo j reflections at the earth's surface and reach maximum heights $j + 1$ times.

The ray *1*, sent out in the direction of the critical angle τ_l at T, has for its asymptote a circle concentric with the earth. The steeper ray *2* does not return to the earth.

On the right rn (horizontal) is drawn as a function of h (vertical).

The typical characteristics of a waveguide in the troposphere are not so pronounced when the duct, (that is the region where rn decreases with rising height) only begins well above the earth's surface. In such an elevated duct, the number of possible connecting rays is limited since the limiting angle

[8]) These conditions have been treated in detail, in an article by H. G. Booker, Elements of radiometeorology, J. Inst. El. Engrs. **13**, IIIa, 69-78, 1946.

τ_l is larger (fig. 5*b*); there is then only a vague distinction between an elevated duct and a transition layer producing total reflections. But for non-total reflection, the longer the wavelength the better the reflection at a transition layer, whereas for an elevated duct the contrary is the case.

So far only the geometrical properties of super-refraction have been discussed. Of these, the most notable is that the range is much larger than normal. Something analogous might be expected with light waves, but their absorption is generally too strong for super-refraction phenomena to be observed. Optical visibility can sometimes be greater than is normally possible, but it can never extend to several times the distance to the normal horizon, as radio waves can. We may mention that mirages (for example, the well-known "fata morgana") constitute a counterpart to super-refraction, as they are generally seen when the refractive index decreases with height less than normally (or even increases); such cases of under-refraction can also occur with radio waves (see section e, on diffraction).

In the space between the earth and the duct the ray paths described exist for all frequencies; however, as in a conventional waveguide, the interference of the fields due to the various rays can only lead to an abnormally strong field if their frequencies exceed a certain limiting one. In the next section we shall discuss the limiting frequency corresponding to a surface duct more fully. This is the simplest case, since the duct is not then merely one boundary of the waveguide, but coincides with it.

Existence of a limiting frequency for a tropospheric waveguide

The waveguide with which the troposphere may be compared when it is super-refracting has the peculiarity of having no firm wall for its upper boundary, whilst its lower boundary is always the earth's surface. The rôle of the upper boundary is played by the concentric sphere (to the earth) from which a certain ray can be continuously reflected, although the level of this sphere depends on which abnormal ray is being considered (fig. 5*a*). The vagueness of the upper boundary has the result that the theory of the tropospheric waveguide is much less simple than that of a normal waveguide bounded by solid walls. Despite this, the theory can be developed in two different well-known ways, namely by geometrical optics or from the complete wave equations.

Geometrical-optical explanation

In the geometrical-optical way of regarding the problem, the field at the receiver R is interpreted as the vector sum of infinitely many rays from T. The path lengths of the rays are all about the same, so their separate contributions would be about equal if we ignored the reflections at the earth's surface (j of them for each ray) and the continuous reflections above ($j+1$ in number). For both rays, both sorts of reflection are nearly total and the accompanying phase shifts are π and $-\frac{1}{2}\pi$ radians respectively (see above). From this it follows that although the amplitudes of the separate components are scarcely changed by the reflections, the mutual phase differences of the fields due to successive rays (for which the parameters j differ by unity, see fig. 5), become about $\frac{1}{2}\pi$; and it appears that for not-too-short wavelengths the differences between path lengths are less important. For each ray one can therefore always find a second ray with a j value differing by only a few units (and hence nearly enough the same amplitude), such that the second ray will more or less extinguish the first one. This interference effect weakens the resulting field very markedly. There is a way out of this difficulty, but only when the phase shift due to differences in path lengths can contribute an extra $\frac{3}{2}\pi$ for a number of successive paths. The total phase difference of $\frac{3}{2}\pi + \frac{1}{2}\pi$ for two consecutive rays then effectively corresponds to a whole wavelength. In practice this can only happen for very short waves, so that only from them can we expect to get a reasonable field strength. This explains in principle why a limiting frequency exists.

On the basis of this model the limiting frequency can be calculated if we assume a suitable function for the change of refractive index within the duct with height; the function must of course be consistent with the condition that rn decreases with height. From a very simple model (r^2n^2 a polynomial of second degree in $1/r$) the wavelength λ_0 corresponding to the limiting frequency comes to

$$\lambda_0 = \tfrac{4}{3} h_0 \cos \tau_l = \frac{4\sqrt{2}}{3} h_0 \sqrt{\frac{\delta(rn)}{rn}} . \quad (3)$$

Here h_0 is the altitude of the top of the duct, i.e. the height where rn is a minimum, and so, by the definition of the limiting angle τ_l, the height approached by a ray leaving the transmitter at the limiting angle; $\delta(rn)$ is the difference between the values of rn at the top of the duct and at the earth's surface.

Equation (3) explains numerically why the waveguide effect in the case of super-refraction is

perceptible only for very short wavelengths. The value of cos τ_l is very small in practice and h_0 is of the order of several tens of metres or several hundreds; hence λ_0 is usually in the region of centimetre waves. In any event, super-refraction was not discovered until the second world war, when centimetre waves were used on a large scale for the first time (to observe aircraft and ships by radar).

Equation (3) also demonstrates the consequences of having an indefinite upper boundary for the duct; the factor cos τ_l is a measure of the indefiniteness, for the factor would be unity if every ray directed upwards was bent back to the earth in one way or another. This does happen, more or less, with reflections of longer waves in the ionosphere, and in this case the whole space between the earth and the upper boundary of the ionosphere can be regarded as a wave guide. The limiting wavelength of this wave-guide is about equal to its width, that is roughly the height h_0 of the ionosphere above the earth. This height is of the order of 100 km, which is greater than any radio wavelengths ever used in practice, and that makes it clear why the ionosphere is a propagating medium for all radio waves long enough to be refracted by it.

Explanation by wave theory

The treatment of radio-propagation in the troposphere with a refractive index depending only on the altitude, generally amounts to the following, whether super-refraction is present or not. From the three-dimensional wave equation, solutions are determined that consist of products of a function of r only and a function of ϑ only, where r and ϑ are polar co-ordinates in a vertical plane through the transmitter with the centre of the earth as origin. $\vartheta = 0$ denotes the vertical through the transmitter (*fig. 6*). From the infinite

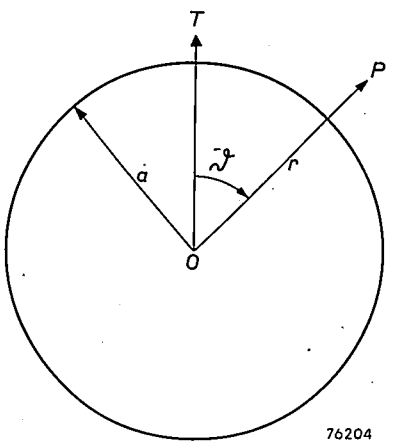


Fig. 6. Polar co-ordinates r and ϑ of a point P, with the centre of the earth as origin O. Angle ϑ measured from the vertical through the transmitter T.

number of solutions of this form for a given frequency, particular solutions are sought that are zero at infinity and satisfy the boundary conditions imposed at the earth's surface. These solutions constitute the modes, which are infinite in number but form a discrete assembly. Each mode represents a condition in which no energy is brought into the system from outside, while the field strength along the vertical through the transmitter is infinite. By choosing the right combination of these modes not only is a field obtained that is infinite at the transmitter — regarded as a point source — but also the singularity there can be chosen to agree with the known data concerning the transmitter (radiated power, polar diagram.)

The limiting frequency now comes into the theory in the following way. The ϑ-dependent factor in the mode-solutions of the wave equation is approximately proportional to

$$\frac{e^{\alpha_r \vartheta}}{\sqrt{\sin\vartheta}},$$

where α_r is different for every mode. The exponents α_r are generally complex with a negative real part, so that for increasing angular distances the wave function falls off exponentially. Only in the case of super-refraction, and even then, only for high enough frequencies, are there one or more modes with purely imaginary values of α_r, corresponding to a wave-function not decreasing exponentially with ϑ. The limiting frequency is then the lowest frequency that has at least one of the latter modes; for them the amplitude of the field is proportional to $(\sin^{-1/2}\vartheta)$ and so decreases as ϑ increases even more slowly than for waves spreading out in a vacuum (where the field is inversely proportional to the distance, i.e. for short distances proportional to ϑ^{-1}). The duct action of super-refraction for frequencies above the limiting one can thus be explained in terms of the contribution to the field strength from modes with purely imaginary values of α_r: we shall denote them by $j\beta_r$.

The wave function for these modes — each is then a function $(\sin^{-1/2}\vartheta)\exp(j\beta_r\vartheta)$ multiplied by a function of r — may be studied more fully, and it is found that they are geometrically-optically equivalent to a system of curved rays, which, like the rays of fig. 5, travel up and down more than once and are alternately reflected discontinuously at the earth's surface and continuously in the troposphere. All rays belonging to one such mode remain in one beam and neither converge nor diverge, and are derived from one another by rotating the whole curve through an arbitrary

angle in the direction of ϑ. In fact these rays are mutually congruent (*fig. 7*). Further it can be deduced that the inclination $\tau_{a,r}$ to the vertical of the ray at the earth's surface is given by

$$\sin \tau_{a,r} = \frac{\lambda}{2\pi a} \beta_r . \quad . . . . . \quad (4)$$

What is especially significant about a system of rays and modes with $\alpha_r = j\beta_r$ is that energy is transported along these rays almost without loss; whilst it is much less simple to describe what happens

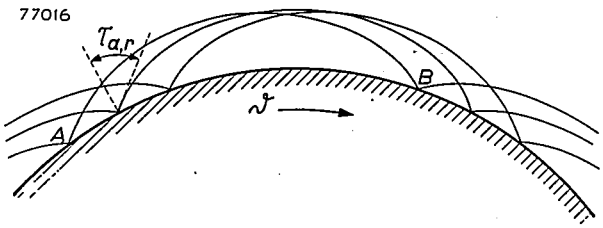


Fig. 7. In the case of super-refraction, the rays belonging to one mode, with $\alpha_r = j\beta_r(\beta_r$ real), can be obtained from one another by rotating the whole of a curve representing one ray about an arbitrary angle in the direction of ϑ. $\tau_{a,r}$ is the inclination of the ray to the vertical where it is reflected at the earth's surface.

when α_r is complex. In the latter case, the energy travelling outwards in the ϑ direction diverges, and part of it leaks away upwards and does not remain concentrated within a layer just above the earth's surface, as it does when α_r is purely imaginary. These two types of modes have been named "leaking modes" and "trapped modes" respectively.

Given the exponent $\alpha_r = j\beta_r$ for a trapped mode, equation (4) shows that a system of rays that can actually return to the earth is only possible if $\lambda < 2\pi a/\beta_r$. This establishes that each of these modes has its own limiting frequency, while the general limiting wavelength λ_0 then corresponds to the smallest value of β_r. The boundary conditions at the earth's surface impose severe restrictions on the possible values of β_r and hence of $\tau_{a,r}$, and this can be interpreted as a quantizing of directions for a given frequency. This must not be confused with the quantizing in many other problems involving characteristic functions, where unknown *frequencies* are quantized and not unknown directions. The direction quantization can also be regarded as arising from the condition that the total phase change between two similarly situated points A and B, due to a rising and falling ray path, must be a multiple of 2π.

To sum up the two treatments: using geometrical optics, we start with the transmitter and the rays sent out from it, and we are only later concerned with the boundary conditions at the earth's surface, when calculating the reflection coefficient. On the other hand, in the wave-theory method we start with the boundary conditions, and from these deduce the modes, and afterwards, when combining the modes, allow for the transmitter.

d) Gradient reflections

The transition layers discussed in a previous section were rather vaguely defined, in that we spoke only of a rapid change in the refractive index over a small interval of height. In fact, every infinitesimally thin spherical shell between heights h and $h+dh$ can be regarded as a layer in which n deviates from its value in the adjoining layers, so that the layer can act as a partial reflector for rays coming from both below and above [9]). For a layer to reflect, it is only necessary that n should not vary much horizontally within the Fresnel zone corresponding to the reflection. Apart from this, these reflections exist in theory whenever the refractive index varies with the height; thus the gradient of n is the origin of the reflections, which may be termed gradient reflections. The ray changes its direction discontinuously, in contrast to what happens with the continuous deflections shown in fig. 1*b*.

In the field at the receiver, contributions are to be expected not only from single gradient reflections (as drawn in fig. 1*d*), but also from multiple gradient reflections, whereby the ray travels to and fro between the different shells several times. The quantitative significance of gradient reflections is still an object of discussion.

A feature of gradient reflections is that they too can contribute to the field far beyond the horizon. The receiver R can always be reached (in the case of *single* gradient reflections) by means of reflections in the plane of symmetry of the arc TR (fig. 1*d*), provided the distance TR is not so large that the point of reflection lies in too rarified air. In any case it is understandable that the differences in field strength obtained for homogeneous and inhomogeneous tropospheres (in which n depends only on the height), can be formally ascribed to gradient reflections, as Feinstein [10]) has shown. Usually, however, this variation in the field is explained by diffraction theory (see next section).

In general, the field resulting from gradient reflections is strongly diminished through the extinguishing interference between the contributions

[9]) H. Bremmer, Physica **15**, 593-608, 1949.
[10]) J. Feinstein, Trans. Inst. Rad. Engrs., Professional group on Antennas and Propagation, Dec. 1952, pp. 2-13.

arising from each of the extremely thin individual layers. These contributions reach the receiver with widely differing phases. Because of this it is assumed that fields observed far beyond the horizon that are too strong to be explained by diffraction, do not come from gradient reflections but from scattering (see Part II of this article). It has to be remembered that the variations in the refractive index, which in any case are strongest in the vertical direction but also exist horizontally, make it difficult to define a sharp distinction between gradient reflections and scattering. In the region above the horizon of the transmitter, the effect of gradient reflections is entirely drowned by the field due to the normal rays discussed above.

e) Fields due to diffraction

Under normal conditions (super-refraction absent) a point on the horizon of the transmitter is the point of contact of a ray just touching the earth. This ray is always bent to some extent through the ever-present refraction, usually downwards, as the refractive index normally increases with approach to the earth. The horizon therefore lies rather further away than it would do under a perfectly homogeneous troposphere. The region beyond the horizon beneath the tangential rays is inaccessible to geometrical-optical rays, if we ignore gradient reflections and transition layer effects (if any). The field in this region can be explained only by diffraction, in the same way as the penetration of light into the edges of shadows.

The deflecting field is usually investigated with the help of the method of the modes, already discussed. The wave functions of the modes were proportional to $(\sin^{-1/2}\vartheta)\exp(a_r\vartheta)$, wherein each exponent a_r has a negative real part $-c_r$ (leaking modes) if we ignore super-refraction. The decrease of the wave function at increasing angles from the transmitter is apparently mainly determined by the factor $\exp(-c_r\vartheta)$.

The field is given by the infinite series that combines all modes, and the series converges everywhere, including above the horizon, although there very slowly. The situation is simplest for distances so great that only the first term (that with the smallest c_r) of the series is numerically important. It is therefore this smallest c_r that determines the range of the transmitter. The most important c_r-values have been found numerically for many physical models, in which the dependence of n on the height is chosen according to meteorological data. A general theory for this has been given by Pekeris [11]).

The results of calculations on any one model for a non-homogeneous troposphere can now be compared with those for a homogeneous troposphere. For the latter, and a given state and type of ground, the field substantially depends only on the following dimensionless parameters:

$$\frac{D}{a^{\frac{2}{3}}\lambda^{\frac{1}{3}}},\quad \frac{h_1}{a^{\frac{1}{3}}\lambda^{\frac{2}{3}}},\quad \frac{h_2}{a^{\frac{1}{3}}\lambda^{\frac{2}{3}}},\quad \ldots\ldots \quad (5)$$

where a is the radius of the earth, and h_1 and h_2 are the heights of the transmitter and the receiver respectively, above the ground. When the gradient of the refractive index near the earth's surface varies only slightly, it appears that the results for an inhomogeneous troposphere can be obtained from those for the homogeneous one by replacing a in the three parameters in (5) by an "effective radius":

$$a_{\text{eff}} = \left[\frac{r\times n(r)}{\frac{\mathrm{d}}{\mathrm{d}r}\{r\times n(r)\}}\right]_{r=a} = \frac{a}{1+a\dfrac{n'(a)}{n(a)}}. \quad (6)$$

Usually, the result of changing the parameters (5) is expressed in another way, namely that in the formulae for the field in a homogeneous troposphere the horizontal distances have to be multiplied by $(a/a_{\text{eff}})^{\frac{2}{3}}$ and the heights by $(a/a_{\text{eff}})^{\frac{1}{3}}$. Normally $n'(a)$ is negative, so that $a_{\text{eff}} > a$; the apparent distances are therefore smaller than the real ones. This agrees with the fact that under average atmospheric conditions, and when the transmitter and receiver are on the ground, a field is observed which is stronger than would occur if the troposphere were homogeneous.

One can obtain a picture of the effective earth-radius in geometrical-optical terms. Consider a ray that starts from a height $h = r-a$ and moves horizontally. The initial curvature of the ray is found to be

$$\frac{1}{\varrho} = -\frac{n'(r)}{n(r)}.$$

If $\varrho = r$, the ray has the same curvature r as the concentric sphere through its starting point, i.e. the ray "follows" the curvature of the earth. In this case, the ray remains at a constant height above the earth, as would also be the case with a *homogeneous* troposphere above a *flat* earth. It is understandable that in all other cases the field depends primarily upon the difference between the curvatures $1/\varrho$ and $1/r$. On the earth's surface, this difference is exactly equal to $1/a_{\text{eff}}$, whereby the

11) C. L. Pekeris, J. appl. Physics **17**, 1108-1124, 1946.

significance of a_{eff} becomes evident. To obtain the field in hilly country, it is reasonable to replace the actual curvature of the earth by that of a circle that approximates as nearly as possible to the profile of the ground between the transmitter and the receiver.

It has been mentioned that the fall in temperature and humidity with height is, under most meteorological conditions, such that $a_{\text{eff}} > a$; a_{eff} is often about $\frac{4}{3}a$. It can happen, however, that $a_{\text{eff}} < a$, especially with a low hanging mist. The conditions for propagation are then worse than in a homogeneous troposphere.

Finally it may be remarked that, even for microwaves, the field never changes abruptly at the horizon for either a homogeneous or a non-homogeneous troposphere.

Limiting cases of very favourable conditions for propagation

To illustrate the conditions that govern the all-important variation of the refractive index, a few idealized limiting cases may be mentioned.

Using the conception "effective radius" of the earth, super-refraction is represented by a negative value of a_{eff}, for the differential coefficient $\partial(rn)/\partial r$ is negative in a duct. When a_{eff} changes its sign, it passes through infinity, and the product rn then has a stationary value at the earth's surface. When close to the surface, rn then varies only slightly with height, and properties not very different from those in an atmosphere with a constant rn can be expected.

In this kind of atmosphere it is easy to see that any one ray must have the same inclination τ to the vertical everywhere, in order to satisfy the path equation $rn \sin \tau = \text{constant}$ (Snel's law). In particular, a horizontal ray remains horizontal, and in theory goes round the earth and comes back to its starting point. In such a medium, one could see the back of one's head if light absorption did not exist! Fleming calculated quite some time ago [12]) that this condition would be very nearly realized if the earth was surrounded by an isothermal atmosphere of krypton with the same temperature and pressure at the earth's surface as the real atmosphere. (In such a krypton atmosphere the pressure would change with height differently of course from our atmosphere.) This is a nice illustration of how the molecular composition of the atmosphere is involved in the refraction index and its changes.

The situation with a constant value of the product rn is actually nearly achieved if super-refraction occurs, near the upper boundaries of the ducts ($r = r_0$); since such a boundary is characterized by a minimum value of rn. Near the level $r = r_0$, rn thus changes very little. Hence at this level, a horizontal ray does indeed remain "floating" and follows the curvature of the earth (ray *1* in fig. *5a*). This is certainly not the most favourable condition for propagation over great distances, for a ray very slightly inclined to the spherical surface $r = r_0$ and leaving the sphere will get a curvature that deflects it still further away. However, when rn is not a minimum but a maximum, rays slightly diverging will be deflected back to the direction parallel to the earth's surface; and this happens on the lower side of an elevated duct. A similar situation exists for acoustic refraction in seawater because the refractive index is influenced both by decreasing temperature and by increasing pressure at greater depths. Exceptionally favourable conditions for acoustic propagation in the sea are realized at a depth of about 1 km. At that depth the product rn (the factor r of which is scarcely important) goes through a maximum. The principle has been put to use in the so-called "Sofar" technique to detect very distant explosions of depth charges: the charges are set to go off at the depth of the duct, and microphones are placed at the same depth. In this way acoustic signals have been sent over many thousands of miles.

We have now discussed five of the six characteristic mechanisms of radio propagation. The sixth, scattering, will be the subject of the second part of this article.

Summary. This article forms the first part of a review of what is known about the troposphere in relation to radio transmission. The most important physical magnitude that influences propagation of electromagnetic waves through the atmosphere is the refractive index n. A formula for n is given as a function of temperature and of the partial pressures of dry air and water vapour in the atmosphere, valid for waves longer than 3 cm. For shorter waves there are deviations due to absorption, by water droplets, and for waves shorter than 2 cm by water vapour and by oxygen. The rest of the discussions are confined to waves longer than 3 cm (and shorter than about 10 m).

Six idealized possibilities are distinguished (which can arise in many different combinations) whereby radio communication can be made using microwaves:

a) along two rays, one direct and one reflected at the earth's surface, both according to the laws of optics (the normal case),
b) along one or a few rays via a transition layer,
c) along many rays that are propagated as if in a wave guide (super-refraction),
d) via reflections brought about by a continuous change in the refractive index with altitude (gradient reflections),
e) by rays which are diffracted, and
f) by rays which suffer scattering.

Cases (*a*) to (*e*) have been considered in some detail. The appearance of a duct is discussed under (c), and the existence of a limiting frequency is explained. Scattering (case *f*) will be dealt with in the second part of the article.

[12]) J. A. Fleming, Proc. Phys. Soc. London, **26**, 328, 1914.

ABSTRACTS OF RECENT SCIENTIFIC PUBLICATIONS OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN

Reprints of these papers not marked with an asterisk * can be obtained free of charge upon application to the Administration of the Research Laboratory, Kastanjelaan, Eindhoven, Netherlands.

2032: J. D. Fast: De invloed van verontreinigingen op de eigenschappen van metalen (Metalen **7**, 2-12, 23-27, 48-50, 1952, Nos. 1, 2 and 3). (The influence of impurities on the properties of metals; in Dutch).

The deleterious influence of impurities on the properties of metals and alloys is discussed. Inclusions may, however, also have a favourable influence on certain mechanical properties of metals and alloys, by inhibiting the growth of crystals. Very useful effects (mechanical hardening, magnetic hardening and resistance against creep at high temperatures) may be obtained with inclusions formed by precipitation from a supersaturated solution. The same effects are deleterious if mechanically or magnetically soft materials are wanted. From a study of the effects of carbon and nitrogen in pure iron, and in iron containing 0.5 per cent manganese, it is possible to give an explanation of the phenomenon of magnetic ageing of steel. The interaction between impurities and dislocations is discussed in connection with Cottrell's theory. Several facts in the domain of strain ageing can thus be explained, but other facts do not fit very well in this theory. (See also Philips tech. Rev. **14**, 60-67, 1952).

2033: F. A. Kröger, H. J. Vink and J. van den Boomgaard: Absorption and fluorescence of solid solutions MgO-NiO (Physica **18**, 77-82, 1952, No. 2).

The absorption spectrum of MgO-NiO solid solutions, between 2500 Å and 8000 Å, shows five discrete bands with maxima at 2900 Å, 4000 Å. 4650 Å, 6700 Å and 7200 Å. At liquid air temperatures the 4000 Å and 4700 Å bands show a fine structure with $\Delta\sigma = 235\text{-}240\ \text{cm}^{-1}$. With cathode-rays MgO containing 10^{-3} mole NiO shows a green luminescence at liquid air temperature, in a band consisting of at least eight equidistant sub-bands with $\Delta\sigma = 194\ \text{cm}^{-1}$.

The bands are attributed to electronic transitions between various states of the Ni^{2+} ion, broadened by coupling with lattice vibrations. This explains why NiO exposed to radiations of wavelength $2500\ \text{Å} < \lambda < 8000\ \text{Å}$ does not show photoconductivity.

2034: W. J. Oosterkamp and J. Proper: Free-air and thimble ionization chambers for Grenz-ray dosimetry (Acta Radialogica **37**, 33-43, 1952, No. 1).

A free-air ionization chamber is described for roentgen radiation with a half value length (H.V.L.) between 0.02 and 1 mm Al. A dose rate of 50 000 r/min can easily be measured.

Rigid thimble ionization chambers can be made with relatively thick windows (0.2 mm and 0.45 mm respectively). Their sensitivity is independent of wavelength between 0.02 and 1 mm Al and between 0.04 and 2.5 mm Al H. V. L. respectively. The independence is obtained by using for the wall a material that has an effective atomic number lower than that of air. This is arranged by a suitable combination of window thickness, depth of the ionization chamber and effective atomic number of the wall material. The variation, with the radiation quality, in the attenuation of the beam due to the chamber window and in the contribution to the ionization from secondary electrons released in the chamber wall, compensate each other.

Dose rate and absorption curves in aluminium, measured with these ionization chambers, are given for a tube with mica-beryllium window (inherent filter equivalent to 40 μ Al) for tube voltages of 11 kV, 16.5 kV, 21.5 kV, 31.0 kV and 49.5 kV (DC).

2035: N. Warmoltz and E. Bouwmeester: An easily degassable ionization gauge with a simple and stable circuit (Appl. sci. Res. **B 2**, 273-276, 1952).

An ionization gauge of the tetrode type is described, with two grids and a thin ion collector on the glass wall. The tube is very easily degassed by heating it in an oven and passing a current through the grids.

The electron current is stabilized by the first grid. For measuring the ion current a balanced bridge D.C. amplifier with one tube is used.

VOL. 15 . No. 6, pp. 161-188 DECEMBER 1953

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
THE PHILIPS INDUSTRIES

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, NETHERLANDS

THE EQUIVALENT CIRCUIT OF A GAS DISCHARGE LAMP

by C. ZWIKKER *).

621.327.4.012

The gas discharge lamp, so far as its electrical behaviour is concerned, differs in many ways from other circuit elements, amongst other things in the non-linearity of its current-voltage characteristic. The behaviour of such a lamp in a circuit is, consequently, often difficult to predict in detail. It is sometimes considered that the gas discharge lamp can be replaced, schematically, by a simple resistance, but as discussed in this article, this may easily lead to false conclusions. Nevertheless for testing purposes, use may often be made of such a resistance, provided that an additional small self-inductance is connected in series. For the purposes of more exact calculations, however, it is necessary to employ a more comprehensive equivalent circuit which exactly reproduces the behaviour of the lamp.

The unambiguous representation of the electrical properties of a gas discharge lamp; for example, the behaviour of the lamp in an A.C. circuit, presents considerable difficulty. In the first place, the properties depend on the type of ballast resistor employed. It is thus incorrect to speak of the properties of the *lamp*. In the second place, all sorts of non-linear phenomena arise as a consequence of the fact that the relation between the current through the lamp and the voltage across it, i.e. the characteristic, is non-linear.

Expedients have, of course, been used to facilitate practical calculations for installations of gas discharge lamps. It is quite common, for example in the testing of ballasts, to consider a lamp as being replaceable by a simple resistance. It will be clear that this can give only a very rough approximation to the truth. Before going further into the considerations leading to a better equivalent circuit for a gas discharge lamp, we will first discuss briefly a few fundamental properties of a gas discharge [1]).

Fundamental properties of the gas discharge lamp

Under the influence of an electric discharge, a proportion of the gas atoms present in the lamp are split into positive ions and electrons.

Let us consider for a moment the behaviour of a single electron. The electron moves under the influence of the electrical field and collides with atoms and ions. During the time interval between two successive collisions it acquires a certain amount of energy. This energy increment is proportional to the field-strength F and also to the distance travelled, i.e. to the so-called mean free path, or more correctly, to the projection of the mean free path in the direction of the field. The mean free path is, however, inversely proportional to the number of particles per unit volume, and hence to the gas pressure p_0 at a given temperature, for example 0 °C. The increase in energy between two collisions is thus proportional to the ratio F/p_0.

The kinetic energy of the electron does not increase, however, without limit; the electron loses energy repeatedly in collisions with the atoms or with the wall. In the equilibrium state, the electron must lose on the average exactly as much energy per collision as it has acquired in its flight. The energy loss during the collision with an atom can take place in two ways. The collisions can be elastic; the atom then acquires part of the energy of the electron in the form of kinetic energy, i.e. the gas becomes warmer. The collision may, however, be inelastic; the atom then reaches an *excited state* or is perhaps even *ionized*. In the latter

*) Now Director of the National Aeronautical Research Institute, Amsterdam; previously with N.V. Philips, Eindhoven.

[1]) For a general survey of the fundamental processes in a gas discharge, see M. J. Druyvesteyn and F. M. Penning, Rev. Mod. Phys. **12**, 87-174, 1940.

case — ionization — a new electron is released into the "plasma". The number of free electrons in the tube is kept up to a certain level in this way, since electrons are continually being lost by recombination. The energy of the colliding electron is, however, lost from the point of view of the purpose for which it is desired to use it, i.e. as radiation energy. This is because the fate of all ions is to be attracted to the negatively charged glass wall and to re-unite there with an electron. The energy thereby set free is lost to the glass wall as heat. The energy employed in the *excitation* of the atoms, however, does fulfil a useful purpose. The excited atoms always return, sooner or later, to the ground state, thereby emitting light or ultraviolet radiation.

The frequency of this process of excitation and emission is independent of the dimensions of the vessel. The rate of neutralization of the ions at the wall is, on the other hand, dependent on the tube diameter, and will occur more rapidly according as the tube is made narrower. It is then said that the wall losses are higher. Thus, to keep the number of electrons at a given level, relatively more ionizations must take place in a narrow tube than in a wide one. The average energy of the electrons in the first case must therefore be greater, i.e. the field-strength must be higher. The narrower the tube, the stronger the electric field and the higher the lamp voltage necessary to achieve an equilibrium condition, i.e. the arc voltage.

Whether a tube is to be considered as "narrow" or "wide" depends on whether the electrons and the ions undergo only a few or many collisions with neutral atoms before reaching the wall. It thus depends on the mean free path of these particles, or, more precisely, on the ratio of the tube radius R to the mean free path. In view of the fact that the mean free path, for ions as well as for electrons, is inversely proportional to the density, i.e. to p_0, the criterion of "narrow" or "wide" is the value of Rp_0, and this product is to be regarded as the decisive parameter in determining the magnitude of the wall losses.

The fact that in the equilibrium condition, the energy taken from the field must be equal to the energy lost by the electrons in collisions with the atoms and with the wall, may be expressed in the equation:

$$F/p_0 = f(Rp_0) + S,$$

in which S represents the contribution made to this balance by the energy radiated in the form of light and heat. S may be considered to be a constant in so far as it is practically independent of the tube diameter and the pressure.

If, now, the lamp voltage is measured and by dividing this by the arc-length the field-strength F is determined, it is found that for a given gas or gas-mixture, at a given temperature, a functional relationship such as that indicated in the above equation, does in fact exist between the variables, p_0 and R. This relationship is depicted in fig. 1, as experimentally determined for the mixture of argon and mercury employed in low pressure fluorescent lamps ("TL" lamps) *). If the argon is omitted from the mixture, p_0 would be so small that the wall losses would become too large. The argon is therefore sometimes called the collision-gas, while the mercury vapour is termed the light-gas, owing to the radiation by the latter of ultraviolet of 2537 Å wavelength.

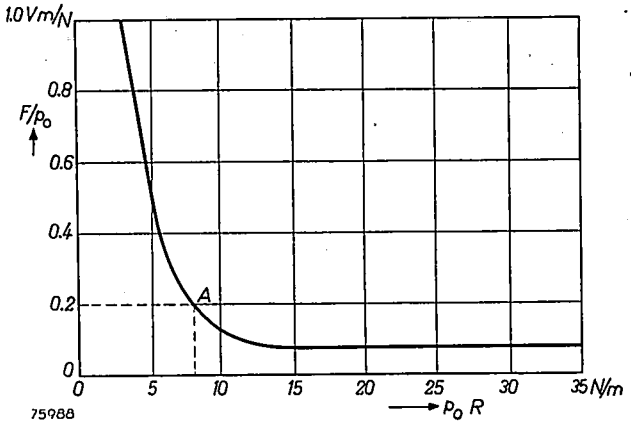


Fig. 1. Relation between the product of gas-pressure [2]) p_0 (reduced to 0 °C) and radius R of the lamp, and the ratio of field strength F to gas-pressure. This relationship is fixed for a given gas-mixture and temperature. The working point of the lamp is chosen preferably in the region of the point A.

The curve of *fig. 1* is useful in the calculation of lamp dimensions. The working-point of the lamp is chosen at the bend in the curve. There is no advantage in working much to the right of the of the bend, since the field-strength decreases only slightly with increasing R, i.e. with increasing weight of glass. A working-point much further to the left has the disadvantage that the small value of R gives rise to considerable wall losses. The working-point having been selected, any one of the three quantitites p, R or F may be varied according to choice, the remaining two then being fixed. The working-point [2]) can be for example (A in fig. 1): $p_0R = 8$ N/m, $F/p_0 = 0.2$ Vm/N; if we choose $R = 0.02$ m, then $p_0 = 400$ N/m², and $F = 80$ V/m.

*) In Gr. Britain "MCF/U"

[2]) The pressure is expressed in newton per m², the unit of pressure in the M.K.S. system: 1 N/m² (= 10 dynes/cm²) is approximately 10^{-5} atmosphere; 133 N/m² is about 1 mm Hg.

Characteristic of a gas discharge lamp

The reader may have observed that in this treatment of the mechanism of the gas discharge the number of electrons has played no direct rôle. The number of moving electrons, i.e. the magnitude of the current passing through the lamp, is determined by the ballast connected in series with the lamp. This number has no influence on the field-strength, i.e. on the lamp voltage; a relationship has been derived above for F with which each electron must comply, without any influence being exerted by the other electrons. In this sense, therefore, the lamp voltage is independent of the current; the characteristic (current plotted as a function of voltage) of the gas discharge lamp is therefore a straight line parallel to the current axis.

Only on deeper consideration is it seen that the current exerts a small influence on the lamp voltage. At greater current strengths, i.e. at greater electron densities, the excited atoms have less time to return to the ground state before colliding again with an electron. There is thus a greater chance that they are brought into a higher excited level or even ionized by a second collision, even if the second colliding electron does not in itself possess the full ionization energy. The ionization is thus made easier, and more charged particles are produced. This shows itself in a lowering of the lamp voltage necessary to maintain the discharge. The characteristic thus exhibits a tendency to fall off slightly.

This shape of the characteristic has the important consequence that the lamp cannot burn in a stable manner without a ballast; this could only be possible with a sufficiently rising characteristic. It would be digressing too far to discuss the deeper origins of this phenomenon, and the reader is therefore referred in this connection to an article [3]) on the stability problem of gas discharges.

In the following we shall assume the necessity of a stabilizing element such as a resistor or choke, with or without a capacitor in combination. We shall, however, for convenience assume that the lamp voltage is independent of the current.

As already observed, the properties of a gas-discharge lamp depend on the ballast employed. This is also true of the characteristic. If, however, we wish to speak of *the* characteristic, we imply thereby the direct current characteristic; by the use of direct current, a resistor is the only form of ballast coming into consideration. The characteristic is, however, also dependent on the temperature of the surroundings, which determines to a considerable extent the lamp temperature, so that measurements may only be commenced after a period long enough to ensure that an equilibrium condition has been reached.

[3]) F. M. Penning, Phys. Z. **33**, 816-822, 1932.

Operation on alternating-current; equivalent circuit

We will consider now the case where the lamp is supplied with alternating current at 50 cycles. The voltage across the lamp, v_{arc}, then varies to a first approximation, as shown in *fig. 2*; the voltage

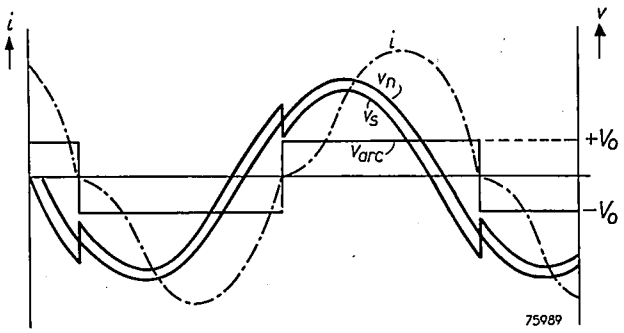


Fig. 2. Form taken by the various current and voltage curves in a gas discharge tube; v_n mains voltage, v_{arc} lamp voltage, which varies discontinuously between $-V_0$ and $+V_0$, v_s voltage across the choke, i current through the lamp.

changes sign each half-period, and is then constant during the half-period. We tacitly assume here that the time needed to reach the equilibrium condition is very small in comparison with a half-period. The fact that this is really not the case has certain consequences, for example in connection with the restriking of the arc at the change of polarity. This refinement is not important, however, for our further treatment. The lamp voltage changes discontinuously, then, during operation, from a positive to a negative value, and vice versa; let the amplitude of this rectangular waveform be V_0. The polarity reversals occur each time that the current i passes through zero (see fig. 2). We assume further, that the lamp is stabilized by a loss-free choke of constant inductance L. At the time t, then, if v_n represents the mains voltage:

$$v_n = v_{arc} + L\frac{di}{dt}. \quad . \quad . \quad . \quad . \quad . \quad (1)$$

If v_n changes truly sinusoidally, the voltage across the choke, $v_s = L\,di/dt$, must display the same discontinuities, though in the opposite sense, as v_{arc}. This voltage is also shown in fig. 2. A discontinuity in di/dt implies, however, a kink in the curve of i plotted as a function of t. The i-curve thus contains kinks at the point where it cuts the horizontal axis. Such a current wave-form is in fact confirmed by measurements [4]). It deviates from a pure sinusoidal form, the current thus containing harmonics.

The behaviour of the lamp can be completely investigated by analysis of the solution of the

[4]) See, for example, the oscillograms shown in the article by E. D. Dorgelo, Philips tech. Rev. 2, 103-109, 1937.

differential equation (1). In this article, however, an alternative, more graphic method will be followed. The lamp in the circuit may be considered as the equivalent of an electromotive force, the value of which jumps discontinuously between $+V_0$ and $-V_0$. This waveform, sometimes termed a *signum* [5]) function (sgn ωt, in which ω represents the angular frequency of the mains voltage), can be analysed in a Fourier series, as follows:

$$v_{arc} = V_0 \operatorname{sgn} \omega t = \frac{4}{\pi} V_0 \left(\sin \omega t + \frac{1}{3} \sin 3\omega t + \frac{1}{5} \sin 5\omega t + \ldots\right). \quad (2)$$

The equivalent circuit of the lamp is thus constituted by a number of A.C. generators connected in series, having angular frequencies ω, 3ω, 5ω etc., and giving voltages of respectively: $V_1 = 4\,V_0/\pi$, $V_3 = 4V_0/3\,\pi$, $V_5 = 4\,V_0/5\,\pi$, etc. Since the voltage at the lamp terminals is considered to be independent of the current, these generators must be considered as being of zero internal resistance. Certain phase relationships also exist between the generators; each time that the total lamp voltage displays a discontinuity, all the partial voltages pass through zero.

In practice, the term containing sin $5\omega t$ and higher terms play no further rôle, partly because of the appearance of the coefficients $\frac{1}{5}$, $\frac{1}{7}$, ..., and partly because the signum function employed is only an idealised form of the voltage curve actually obtained. The latter has somewhat rounded corners, and therefore contains the higher harmonics to a lesser degree than expressed in formula (2). We will confine our attention, therefore, to the first two terms; the equivalent circuit of the lamp is thus reduced to two A.C. generators in series, with frequencies ω and 3ω.

From formula (2) it is seen that the second harmonic, and in general all even harmonics, do not appear in the lamp voltage. The fact that they do appear in practice, albeit to only a slight degree, is due to the asymmetry of the lamp (cathodes not exactly equal). The asymmetry becomes more pronounced as the lamp nears the end of its life.

Calculation of the current through the lamp

We shall now, on the basis of this equivalent circuit, make some simple calculations regarding the shape of the current curve in fig. 1, investigating particularly the fundamental and the third harmonic of this current. This is important in practice: it is desirable that the third harmonic is kept as small as possible in the lamp circuit, since such a distortion unfavourably affects the economy of the electricity supply system.

The fundamental

We can also make a Fourier analysis of the lamp current. The component i_1, with frequency ω, makes up by far the greatest part of the current through the lamp (as can be seen from fig. 2). In the first approximation we can neglect all the higher components, and thus omit the suffix 1. We assume further that the fundamental is in phase with the rectangular voltage v_{arc}. We shall discuss a better approximation at the end of this article.

The fundamental term i can be represented in the usual way, together with the mains voltage v_n and the component v_1 of the lamp voltage having frequency ω, in a vector diagram, since all these quantities vary with the same frequency ω. *Fig. 3* shows this vector diagram for the effective values of the various quantities, given in capital letters. The effective value V_1 of the voltage v_1, is put equal to αV_0; this follows from the peak value $4V_0/\pi$ (see formula (2)) by dividing by $\sqrt{2}$, so that obviously $\alpha = 4/\pi\sqrt{2} = 0.90$. The voltage across the choke, $V_s = I\omega L$, is also shown.

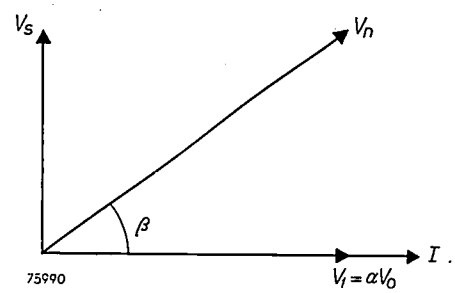


Fig. 3. Vector diagram of the effective values (given in capital letters) of the currents and voltages in the lamp and choke. V_1 is the Fourier component, with frequency ω, of the lamp voltage v_{arc}. It is equal to αV_0, where α is the distortion factor.

As assumed, the current I is, to a first approximation, in phase with V_1; the mains voltage is displaced in phase by an angle β in relation to the current. The voltage across the choke of course differs in phase by $\pi/2$ with reference to the current.

The power taken up by the lamp is (since I and V_1 are in phase), according to this approximation, equal to $\alpha V_0 I$. The quantity α is called the distortion factor of the lamp. It is less than unity because the current and voltage curves differ from each other in *shape* and not because of the existence of a phase difference (in this approximation zero phase difference is assumed).

[5]) The latinized form of *sign*. Strictly, the term applies only to a square waveform of amplitude *unity*. Thus sgn $\omega t = \pm 1$, alternating with frequency $\omega/2\pi$.

Although the only correct equivalent circuit for a gas discharge lamp is the above-mentioned series connection of voltage generators, it is very useful for practical purposes to have available a more convenient substitution element with which measurements can be carried out. This could be employed, for example, in the testing of ballasts.

On the basis of the above treatment, it is sometimes attempted to achieve this end by replacing the lamp by a simple resistance ϱ, of magnitude:

$$\varrho = aV_0/I. \quad \ldots\ldots\ldots \quad (3)$$

Such a resistance certainly dissipates the power aV_0I. It is clear, of course, that this treatment ignores the presence of harmonics in the lamp current. The value of the equivalent resistance depends on the current that is to be allowed to pass through the lamp: the term "equivalent resistance" must therefore be employed with some caution.

From fig. 3 we see that

$$I \doteq \frac{aV_0}{\omega L} \tan \beta.$$

For various reasons, which will not be discussed further here, a value of 60° is chosen for β. This is obtained by a suitable choice of V_0, and thus of the lamp dimensions ($\cos \beta = aV_0/V_n$). Then:

$$I = \sqrt{3}\,\frac{aV_0}{\omega L}. \quad \ldots\ldots\ldots \quad (4)$$

A phase difference of 60° also appears between the mains voltage and the mains current, which is here the same as the lamp current; this can be compensated by connecting a capacitor with capacity C across the mains (*fig. 4*). *Fig. 5* shows the vector diagram of the effective values of the currents and voltages so obtained. The capacitor must pass such a current that the combined current through capacitor and lamp is in phase with the mains. From fig. 5 it can be seen that, if $\beta = 60°$, the following equation must hold:

$$V_n \omega C = \frac{\sqrt{3}}{2} I, \quad \ldots\ldots \quad (5)$$

or, according to (4):

$$\omega C = \frac{3}{4\omega L}. \quad \ldots\ldots\ldots \quad (6)$$

The value of C can thus be calculated, either from the lamp current or from the inductance of the choke.

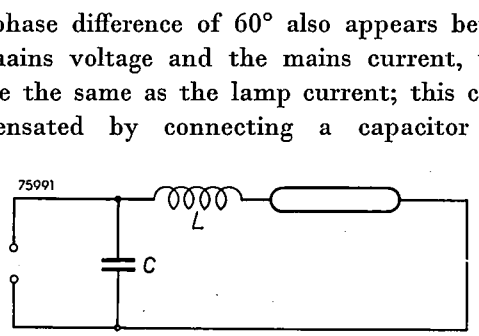


Fig. 4. Circuit for the compensation of phase displacement between mains voltage and lamp current. Compensation is achieved by suitable choice of the capacitance C.

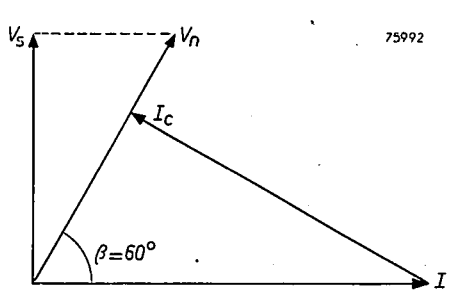


Fig. 5. Vector diagram corresponding to the circuit of fig. 4. I_c is the effective value of the current through the capacitor. I_c and I combined must be in phase with the mains voltage V_n.

The third harmonic

The equivalent circuit of the lamp also contains an e.m.f. of angular frequency 3ω, with a peak value of $4V_0/3\pi$, and this voltage generator also sends currents through the lamp and the circuit, but now with angular frequency 3ω. These currents are superimposed on the current due to the e.m.f. of angular frequency ω. There is a phase relationship which depends, however, on the ballast used; this phase relationship between the currents is determined by the already mentioned phase relationship between the voltages and the inductances and capacitances in the circuit.

Let us consider for a moment the simple case when no third harmonic is present in the mains voltage, an ideal self-inductance L again being taken as stabilizing element for the lamp. The effective value of the current component with angular frequency 3ω is then:

$$I_3 = {}^1/_2 \sqrt{2} \cdot \frac{4}{3\pi} V_0/3\omega L = \frac{a}{9}\frac{V_0}{\omega L}, \quad \ldots \quad (7)$$

in which the distortion factor a is again introduced.

Comparison with (4) gives for the ratio of third harmonic to fundamental current, for this case:

$$I_3/I = \frac{1}{9\sqrt{3}} \approx 0.06. \quad \ldots\ldots \quad (8)$$

Thus the lamp current contains about 6% third harmonic. The method employed, i.e., the treatment of the lamp current as being almost entirely due to the fundamental term, thus appears to be admissible.

Influence of the lamp circuit on the third harmonic

The extent to which a gas discharge lamp produces a third harmonic in the mains current, i.e. causes distortion, evidently depends considerably

on the circuit employed. A few examples of this dependence are given here.

A capacitive ballast is often employed in place of an inductive ballast. The former comprises a choke of inductance L, in series with a capacitor of such a capacity C that its reactance at the frequency ω is equal to $-2j\omega L$, thus $C = 1/2\ \omega^2 L$. The absolute value of the total impedance to the fundamental current is then again ωL, but the impedance is now capacitive. The advantage of such a circuit is quite clear. In a circuit containing a number of lamps, some of which are connected capacitively and others inductively, it is possible, without using correction capacitors, to reduce the phase difference between mains voltage and fundamental current to zero.

The numerical value of the impedance of the series-coupled C and L to the third harmonic is equal to:

$$|Z_3| = 3\omega L - \frac{1}{3\omega C} = 7\omega L/3$$

and has an inductive character. The third harmonic current in this case amounts to:

$$I_3 = {}^1/_2\sqrt{2} \cdot \frac{4}{3\pi} \cdot \frac{V_0}{7\omega L/3} = \frac{1}{7\sqrt{3}} I,$$

i.e. to about 8% of the fundamental current.

A special case of the above-mentioned combination of capacitive and inductive elements is the so-called "tulamp" circuit (*fig. 6*). Here also, due to the value of C chosen, the fundamental mains current is in phase with the mains voltage. Since the fundamental currents in the two circuit branches differ in phase by $2\pi/3$, the third harmonic in the

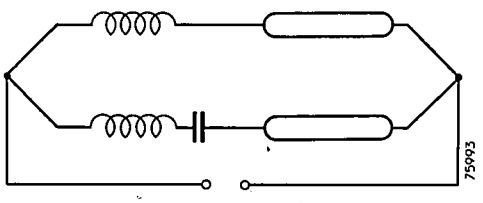


Fig. 6. "Tulamp" circuit. The value of the inductance is the same in each branch, and the capacity is so chosen that the impedance of the capacitive branch is equal and opposite to that of the inductive branch. The total current through the combined circuit is thus brought into phase with the mains voltage.

two branches are in phase. No reduction in the third harmonic component of the mains current is therefore obtained by the use of this circuit.

It is, however, possible, by use of a different type of circuit, to eliminate the third harmonic in the mains current, even when only one lamp is present. If, in the circuit of *fig. 7*, we so choose the capacitor and choke that $3\omega C = 1/3\omega L$, then no current component of frequency 3ω appears in the mains. The mains are thus, for this frequency, short-circuited. The third harmonic flows through the lamp, however, undiminished.

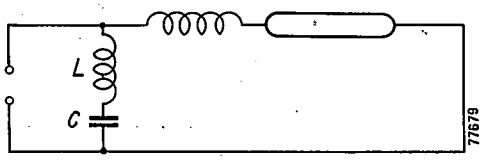


Fig. 7. Circuit with capacitor and choke chosen so that $3\omega C = {}^1/_3\omega L$. With this circuit the mains are short-circuited for currents of frequency 3ω, but the full third harmonic passes through the lamp.

A final example of the influence exerted by the circuit on the extent of the distortion in the current, is provided by the star-connected circuit of *fig. 8*. The three fundamental currents are

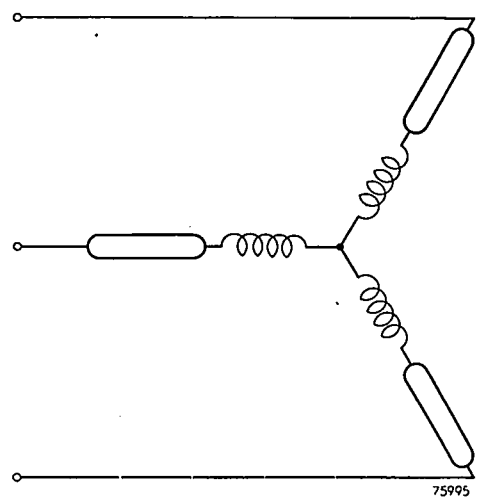


Fig. 8. Star-circuit for three "TL" lamps with their associated ballasts.

separated $2\pi/3$ in phase, and the third harmonics will be in phase if a neutral conductor is present. In that case a three-fold third harmonic passes through the neutral. In the absence of a neutral, however, the third harmonics can have no finite values; no current of frequency 3ω can then pass through the lamps.

If the mains voltage itself contains a third harmonic — and that is usually the case on week-days — all sorts of conditions arise. It is therefore never certain that the third harmonic, such as is obtained under ideal conditions with a given circuit, will be produced in practice even with the same circuit.

Distortion produced outside the lamp

The important question in connection with distortion, i.e. the amount of third harmonic which a lamp produces in its circuits, is still only partly answered. In order to be able to estimate the increase in third harmonic in the mains caused by a given installation of gas discharge lamps and ballasts, we must take into consideration also the distortion caused by the ballasts — i.e. the chokes — themselves. This arises owing to the fact that the iron in chokes always exhibits a curved *B-H* characteristic — the magnetic induction B saturates when the field-strength H, and hence the current through the choke, reaches a certain valuc. The deviation from the ideal linear characteristic is greatest when the current through lamp and choke is at peak. The magnetic induction, therefore, does not vary in a pure sinusoidal manner with time, even if the current itself were purely sinusoidal. The third harmonic i'_3 produced in the choke, is proportional to the derivative of the third harmonic in the Fourier expansion of the variation of the magnetic induction with time. Since this derivative is of course zero when B itself is at maximum, i'_3 will be exactly zero when the current i through lamp and choke is at maximum. As we have seen, the third harmonic i_3 produced by the lamp is then just at its maximum value. The total third harmonic current is then, to a close approximation, equal to $\sqrt{i_3^2 + i'^2_3}$.

We have already seen that i_3 normally amounts to about 6-8% of the fundamental current i. There is no advantage in so constructing the choke (by choosing B_{max} so low, which means making the choke very large), that i'_3 is much less than 7% of i. If i'_3 is, for example, made *equal* to i_3, the total third harmonic is still only 10% of the fundamental current. This percentage is, in fact, found in the usual installations.

Improved equivalent circuit

We will consider somewhat more closely, in conclusion, the case of an ideal choke employed as ballast. *Fig. 9* shows the current component i_1 of frequency ω, and the third harmonic i_3 passing through the lamp, together with their phase relationships. If we sum i_1 and i_3 to give i, then a current waveform is obtained similar to that shown in fig. 2, particularly in that the maximum lies to the right of the centre.

If we consider the diagram more closely, we sec that the current $i = i_1 + i_3$ passes earlier through zero than does i_1, at the point at which i_1 (ignoring sign) is approximately equal to the peak value of i_3. If we represent the phase displacement angle correspondong to this *point* by ψ, then:

$$\sin \psi \approx \psi \approx \frac{I_3}{I} = \frac{1}{9\sqrt{3}} \text{ radians.}$$

We can now recalculate the current through the lamp, bearing this phase displacement in mind, and thus obtain a closer approximation. For this

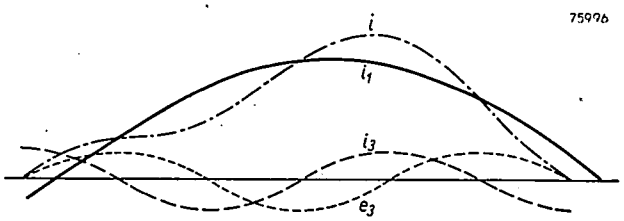


Fig. 9. Combination of the fundamental component i_1 and the third harmonic i_3 to give i, shows that the fundamental current i_1 (which was taken as the lamp current in the first approximation) passes later through zero than i (which serves as a close approximation to the true current: cf. fig. 2). The alternating voltage e_3, which is supplied by the generator with frequency 3ω in the equivalent circuit, passes through zero at practically the same moment as does the total current i. By use of an inductive ballast, i_3 is displaced in phase by approximately 90° with reference to e_3.

purpose, the vector $\dot{I}$ in fig. 3 must be replaced by I_1, which makes an angle $\psi = 1/9\sqrt{3}$ with V_1, so that $V_s = I_1\, \omega L$ is no longer perpendicular to V_1 (*fig. 10*). Since i_1 passes later than i through zero,

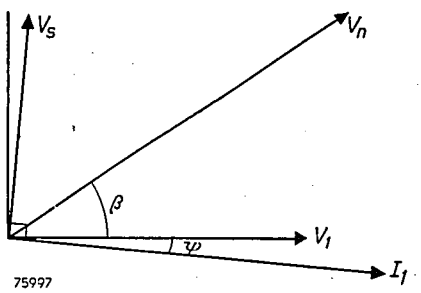


Fig. 10. Vector diagram of the effective values of the currents and voltages in a gas discharge lamp, when the influence of the third harmonic is taken into consideration.

I_1 is slightly behind V_1 in phase. It is therefore more accurate, in this second approximation, not to take a pure resistance $\varrho = aV_0/I$ (formula (3)) as the equivalent element for the lamp, but to place in series with the resistance a small self-inductance λ such that $\omega\lambda/\varrho = 1/9\sqrt{3}$, or:

$$\lambda = 0.06 \frac{V_0}{\omega I}.$$

Such a series combination of resistance and inductance can, in fact, serve in many cases as a suitable equivalent circuit for a gas discharge lamp, as has been found in practice. It is, however, always only an expedient, in view of the fact that both ϱ and λ depend on the current which has to be passed through the lamp.

Summary. After a short introduction dealing with the properties of gas discharge lamps in general, the characteristic of these lamps is further discussed. If such a lamp is connected to an alternating current supply, the lamp voltage changes, to a first approximation, discontinuously from a positive to a negative value, and vice versa; the lamp voltage is thus a square waveform which can be represented by a *signum* function. As a consequence, higher frequency components, i.e. harmonics, appear in the lamp current. The equivalent circuit of the lamp cannot, for various reasons, including the non-linearity of the characteristic, be composed only of resistances and other linear circuit elements; it can only be represented by a series connection of A.C. generators with frequencies ω (mains frequency), 3ω, 5ω etc. Employing this equivalent circuit, the fundamental component and the third harmonic of the lamp current are calculated to a first approximation. It is found that a gas discharge lamp generates a third harmonic of the order of 6-8% of the fundamental component. The amount of this passing into the mains depends further on the circuit in which the lamp is employed. In conclusion, the current through the lamp is analyzed to a second approximation, from which it is seen that a "substitute" for the equivalent circuit, consisting of a series combination of a resistor and a small self-inductance, can be useful in practice.

A PRE-AMPLIFIER FOR USE WITH ELECTRONIC VOLTMETERS AND OSCILLOSCOPES

by F. G. PEUSCHER and J. van HOLTHOON. 621.394.645:621.317.7

Although the amplifier which is the subject of this article embodies no new principles or revolutionary features, a description would seem to be justified in view of the great utility of this unit, and its value as an adjunct to other electronic measuring equipment.

Triodes and tetrodes were used in valve (or electronic) voltmeters as much as thirty years ago. Instruments of this kind have a number of advantages over other types of A.C. voltmeters. The principal advantages are their high input impedance — resulting in only a small load on the measured voltage — and, in conjunction with amplifiers, their high sensitivity. Especially in the last ten years, the development of electronic voltmeters has made great strides, mainly in the direction of still higher impedances and sensitivities. Of the innumerable types of instrument to which this development has led, the Philips voltmeter type GM 6016 is an outstanding example, being suitable for frequencies of 1 kc/s to 30 Mc/s, and giving full scale deflection on as little as 3 mV.

The advantages mentioned are also relevant to cathode-ray oscilloscopes. These are also used in conjunction with high-impedance amplifiers to give them a high sensitivity, and these amplifiers too, require to be designed for a wide frequency band.

Although sensitivity may already be high, instances occur in practice where even greater sensitivity is required. To meet such needs, an amplifier has been designed [1]) having an input impedance of 5 MΩ and which, connected between a voltage source and a voltmeter, increases sensitivity by a factor of 100. This amplifier (type GM 4574) is depicted in *fig. 1*, and forms the subject of the present article.

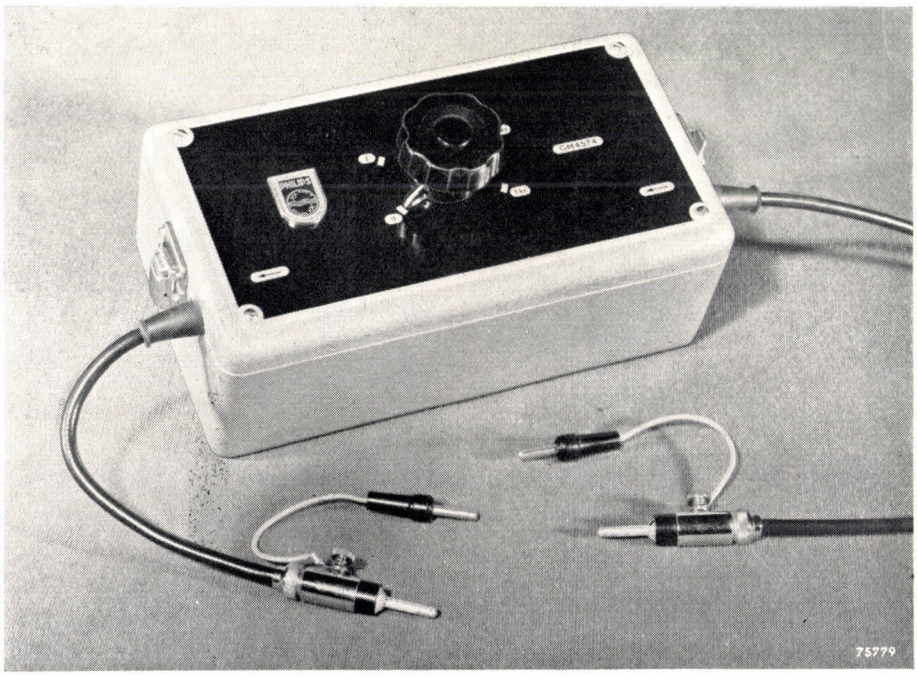


Fig. 1. Amplifier type GM 4574, giving an amplification of 100×. Base 22 cm × 11.5 cm.

[1]) The design was also the work of E. E. Carpentier, who has since left the services of the Company.

General description

In the design of this amplifier, the principal object was to produce a small unit that would not occupy more space than strictly necessary on a bench on which room has also to be found for equipment under test and numerous measuring instruments. For this reason, sub-miniature valves of the kind originally designed for hearing-aids [2]) were decided upon, these being operated from dry batteries. The need for a mains transformer, rectifier valve and smoothing equipment is thus dispensed with. Only a small amount of power is consumed by the filament circuits of these sub-miniature valves, and therefore only small batteries are required.

There is another advantage in the use of batteries: the voltage surges invariably occurring in mains voltages, as well as ripple voltages, are completely avoided. This is very important from the point of view of the purpose for which the amplifier is intended, since the input voltages concerned are often very small, viz. of the order of 0.1 mV.

At the same time, batteries are not without their disadvantages; they tend to produce substances which would set up corrosion or endanger the insulation. In the present amplifier this difficulty has been overcome by housing the batteries in insulated boxes (*fig. 2*) which entirely prevent contamination of the amplifier, and which can be quite easily removed and rinsed in water if necessary.

In order to ensure that the amplification shall remain as constant as possible, this being of course an essential feature where measurements are concerned, considerable negative feed-back is employed. This is particularly desirable because the slope of the valves varies fairly widely as a result of the gradual drop in the battery voltages.

For an overall amplification of 100×, the gain that would be obtained *without* negative feed-back must accordingly be much more than 100×. Two stages in cascade are provided, each giving an amplification of about 63×, or 4000× in all. These stages precede a third, cathode-follower, circuit, the object of which is to keep the output impedance low (less than 5000 Ω), so that the amplification shall be independent of the (high) input impedance of the various equipment connected to the output terminals. The amplification obtained from this stage is slightly less than unity.

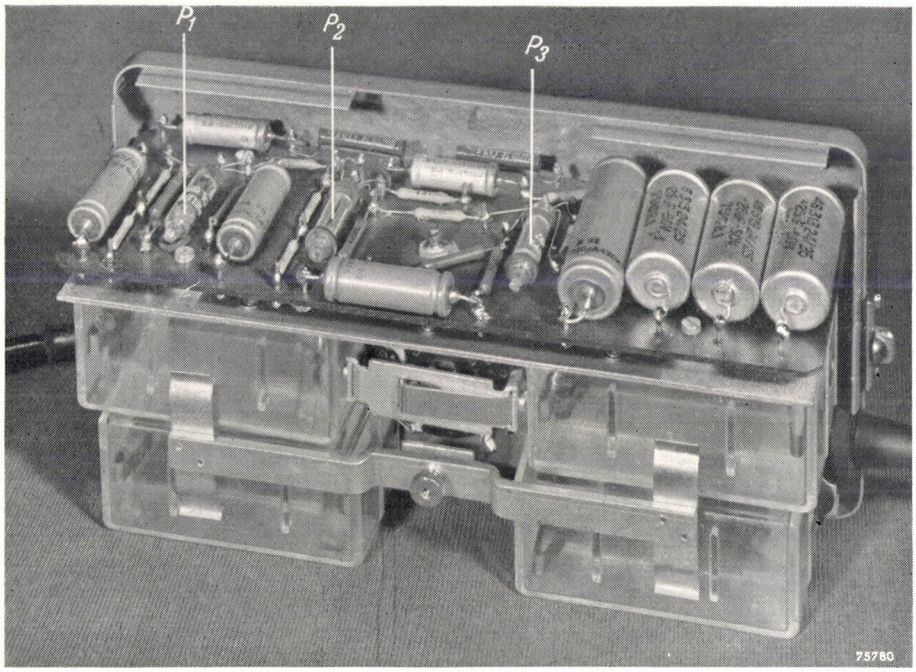


Fig. 2. Interior of the amplifier GM 4574. P_1, P_2, P_3 sub-miniature valves type DL 67. The four boxes made of transparent insulating material are the battery boxes (three for the L.T. batteries (1.5 V), and one for 2 anode batteries of 22.5 V). These boxes ensure effective insulation between the batteries individually and with respect to earth. They also protect the components from contaminating substances from the batteries, and can easily be removed for cleaning.

[2]) Philips tech Rev. **15**, 37-48, 1953 (No. 2).

The circuit

The schematic circuit is shown in *fig. 3*. The three valves are pentodes, type DL 67. As the cathodes are not interconnected in this circuit, each filament (current 13 mA) is fed from its own battery (1.5 V). The battery boxes mentioned above guarantee adequate insulation between the batteries themselves and with respect to earth.

The DL 67 is designed to give sufficient anode current even on low supply voltages (max. 45 V), this being ensured by incorporating in it a screen grid of closely spaced wires. The mutual conductance of the screen grid is accordingly very high, which means that variations in screen voltage have a pronounced effect on the working point of the valve. In order to prevent the working point from drifting too much as a result of diminishing voltage from the anode battery, the screen grids of the first and second valves are fed through high resistances (the third valve is connected as a triode). When the battery voltage drops, the screen current also drops slightly and the voltage loss in the feed resistor is reduced, thus partly compensating the fall in the battery voltage. Consequently, the screen-grid current, and hence also the anode current, are much less dependent on the battery voltage than if the screens were connected direct to a tapping on the battery. Differences in the mutual conductance of the screen between one valve and another are thus also smoothed out.

The screens are decoupled by capacitors C_4 and C_5, which are electrolytic capacitors of high capacitance (25 μF), each having in parallel with it a paper capacitor of much lower capacitance (not shown in fig. 3). Electrolytic capacitors have the advantage of a very small volume per μF, especially at the low working voltages occurring in this amplifier. The object of the paper capacitor is to ensure effective decoupling also at high frequencies, at which electrolytic capacitors have a fairly high impedance (their series resistance increases with frequency). A third electrolytic capacitor (C_6) is connected across the anode battery, to avoid undesirable coupling across the internal resistance of the battery, which can become fairly high as the battery gets older.

Capacitive coupling is employed between the input and first valve (C_1-R_1), between the first and second valves (C_2-R_2) and at the output (C_3-R_3); the second valve is D.C.-coupled to the third. The cathode connection is in each case the negative side of the L.T. battery. Grid bias for the first and second valves is established with respect to the negative side of the filament, across the high-value leak resistors (R_1, R_2); this bias is increased slightly for the first valve by the voltage drop across the impedance Z_1 which is part of the negative feed-back circuit.

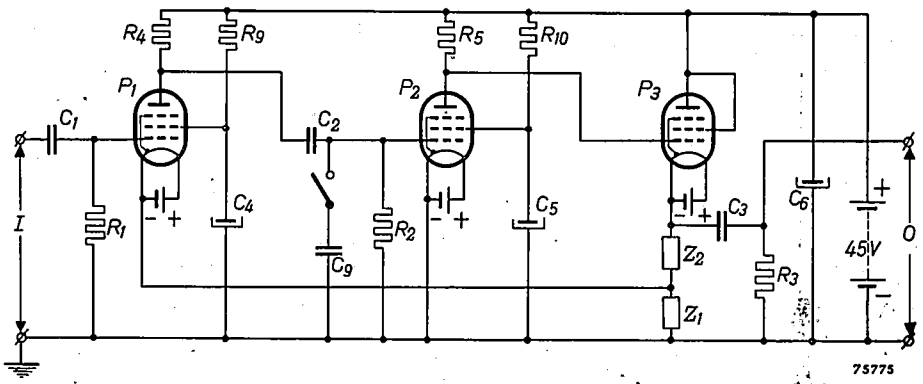


Fig. 3. Circuit diagram of the amplifier. *I* input. *O* output. P_1, P_2, P_3 pentodes type DL 67 (P_3 connected as a triode). C_1, C_2, C_3 coupling capacitors. C_4. C_5, C_6 decouplers. C_9 capacitor for suppressing the noise voltage while reducing the band-width. R_1, R_2 grid leaks. R_3 resistor across output terminals. R_9, R_{10} screen grid resistors. Z_1-Z_2 negative feed-back network.

The slope of the valves at the working point is about 0.1 mA/V; the input impedance is roughly 5 MΩ.

Negative feed-back

The negative feed-back circuit is shown in fig. 3. It consists of two impedances Z_1 and Z_2; a feed-back of $\beta = Z_1 (Z_1 + Z_2)$ is produced, this representing that part of the output voltage which is returned to the input. The ratio in which the amplification A (without negative feed-back) is thereby reduced, is $1:(1 + \beta A)$. In the present case A is about 4000 and the required amplification is $A' = 100$, so that $\beta = Z_1 (Z_1 + Z_2)$ must be roughly 1/100. Variations

in the amplification are thereby reduced in the ratio of $1:(1 + \beta A) = 1/40$, which is sufficient to ensure that the amplification (except at the very highest frequencies concerned) is not reduced by more than 10% when all the battery voltages have dropped to one half their rated values. Non-linear distortion is also attenuated in a ratio of 1/40, and this is important when the amplifier is employed with an oscilloscope.

Fig. 4 gives further details of the negative feed-back circuit. At frequencies in the region of 1000 c/s the capacitances C_p and C_7 may be neglected, and C_8 can be regarded as a short circuit. Z_1 then consists only of the resistor R_6, and Z_2 of resistors R_7 and R_8 in parallel. The required amount of amplification ($100 \pm 3\%$) is obtained by using the appropriate value for R_6.

In fig. 4, C_p represents a stray capacitance, this being mainly the capacitance to earth of the L.T. battery for the first valve. At high frequencies the effect of C_p is not negligible, but this can be counteracted by means of a small trimmer C_7, since the ratio $Z_1/(Z_1 + Z_2)$ is independent of frequency when C_7 satisfies the condition

$$C_7 \frac{R_7 R_8}{R_7 + R_8} = C_p R_6.$$

To ensure that voltages of square waveform will be amplified without distortion, C_7 is very carefully adjusted, using a square-wave voltage of about 5000 c/s. The form of the output voltage of the amplifier is simultaneously checked with an oscilloscope which immediately reveals any discrepancy in the adjustment [3]) (see *fig. 5*).

network is so designed that the feed-back is reduced with decreasing frequency. When the frequency is reduced, the capacitor C_8 (fig. 4) increases the impedance of the branch R_8-C_8 and hence also of Z_2, so that the ratio of $Z_1/(Z_1 + Z_2)$ decreases.

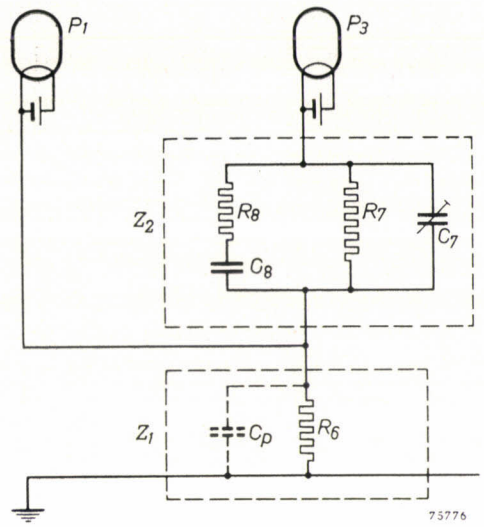
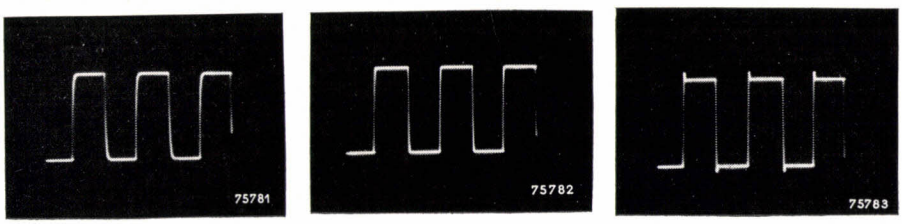


Fig. 4. Details of the negative feed-back network Z_1-Z_2. C_p is a stray capacitance in parallel with R_6; the effects of this capacitance are counteracted by means of a trimmer C_7. The fixed capacitor C_8 in series with R_8 improves the response curve at low frequencies.

This compensation is not perfect, however; to ensure that there will not be too little compensation in any given frequency range, some over-compensation must be permitted in another range. This is the reason why the curve (*1* in *fig. 6*) rises slightly at frequencies below 100 c/s, to reach a maximum



a *b* *c*

Fig. 5. Effect of the trimmer C_7 (fig. 4). The input voltage is of square waveform, 5000 c/s. Oscillograms of the output voltage: *a*) value of C_7 too low, *b*) correct, *c*) too high. Waveforms displayed on oscilloscope type GM 5653.

Owing to the finite time constant of the three RC couplings (C_1-R_1, C_2-R_2 and C_3-R_3), the gain is smaller in the lower frequencies than in the middle of the range; to compensate this, the feed-back at 2 c/s, which on average over a number of these amplifiers is roughly 12% above the nominal value of the amplification. Consequently, a perfectly square wave of 12.5 c/s is slightly distorted, as will be seen from *fig. 7a*; at 25 c/s this distortion is almost imperceptible, and at 50 c/s it cannot be

[3]) See J. Haantjes, Judging an amplifier by means of the transient characteristic, Philips tech. Rev. **6**, 193-201, 1941.

detected at all (fig. 7*b* and *c*). Without the compensation, even a 50-cycle square-wave voltage would be badly distorted, and the compensation therefore ensures a useful increase in the range of frequencies for which the amplifier can be employed.

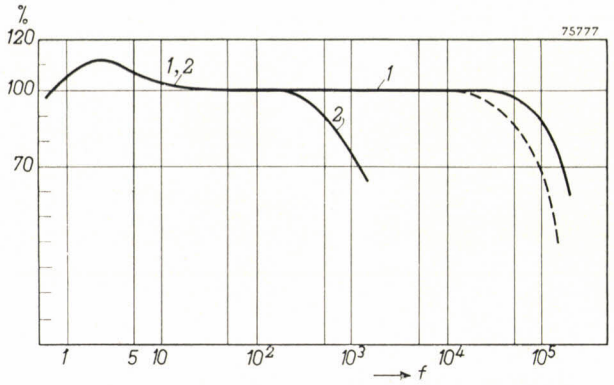


Fig. 6. Amplification (relative) as a function of the frequency, plotted for a load of 1 MΩ in parallel with 20 pF: the curve *1* is obtained without the capacitor C_9 in circuit, *2* is obtained with C_9 connected (fig. 3). 100% corresponds to amplification $A' = 100 \pm 3\%$, with nominal battery voltages, and with 10% less than this value for battery voltages of half the rated values. In the latter case the curve is shown by the broken line.

With a response curve of this shape (*1* in fig. 6), obtained by means of the compensation mentioned, the unit is suitable for amplifiying sinusoidal frequencies of from 1 to 150,000 c/s; with a square wave voltage the frequency range is 10 to 10,000 c/s.

Even when all the battery voltages have dropped to half their rated values, the response curve retains the form shown, with the exception of the extreme right-hand end, i.e., at the higher frequencies, where the drop becomes slightly steeper (broken line in fig. 6).

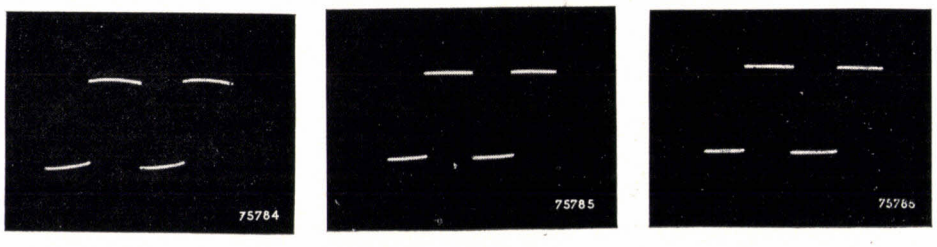
frequencies; the high-frequency end of the range can then be safely suppressed, thus eliminating the noise attributable to the higher frequencies. This is effected very simply by placing a capacitor in circuit (C_9 in fig. 3). The noise voltage, referred to the input, is less than 10 μV without the capacitor C_9, but is less than 5 μV with this capacitor in circuit. In this case the response curve of the amplifier is as shown by line *2* in fig. 6, this being suitable for the amplification of sinusoidal voltages at frequencies of 1 to 1000 c/s, and square-wave or pulse voltages of from 10 to 50 c/s.

In order to avoid possible trouble due to microphony when very low voltages are to be amplified, the amplifier chassis is flexible mounted in the case.

"Forming" of the electrolytic capacitors

As already mentioned, the screen grids of the valves are fed through high-value resistors with decoupling by means of electrolytic capacitors (C_4 and C_5, fig. 3). This arrangement necessitates certain precautions in view of the leak current of the capacitors, but which are fully justified by the advantage of electrolytic capacitors mentioned above (small volume per μF).

The dielectric in an electrolytic capacitor is a thin film of aluminium oxide on the aluminium anode [4]), this being produced by a "forming" process which consists in passing a direct current through the capacitor in the correct direction; oxygen is then evolved at the anode, and this oxidizes the aluminium. The higher the D.C. voltage across the capacitor, the thicker the layer of oxide and, hence, the



a *b* *c*

Fig. 7. Oscillograms of the output voltage as obtained from a square waveform voltage, *a*) of 12.5 c/s, *b*) of 25 c/s, *c*) of 50 c/s. At 12.5 c/s there is slight distortion (due to overcompensation by C_8 (fig. 4)); at 25 c/s this has almost disappeared, and at 50 c/s it is imperceptible. Taken with oscilloscope type GM 5653.

Where amplification of very small voltages within a limited range of frequencies is required, it is useful to be able to reduce the background (noise) at the expense of the bandwidth. Such conditions will often occur in investigations of mechanical phenomena, which are concerned with very low

higher the resistance which it offers and the lower the capacitance.

During such time that the capacitors are not in use, the oxide layer is partially broken down, so that,

[4]) W. Ch. van Geel and A. Claassen, Electrolytic condensers, Philips tech. Rev. **2**, 65-71, 1937.

when a voltage is applied across these capacitors after a long period of idleness, the leak current is at first quite high. A new layer is quickly formed, however, of thickness depending on the applied voltage.

It follows, therefore, that there may be some difficulty when an amplifier designed on the lines of the circuit shown in fig. 3 is switched on. A leak current is then able to flow through C_4 and C_5 which is of the same order as the screen-grid current (normally 10 μA), and the required screen voltage is thus by no means reached. Moreover, with the potential of only a few volts then occurring across the capacitor, there is little opportunity for the leak current to decrease within a short time owing to re-formation of the capacitor, and thus to establish the required working conditions.

This difficulty is overcome by including two intermediate positions in the switch, which is in any case needed for the filament circuits and for the capacitor C_9. The switch accordingly has five positions, which we shall now mention in turn.

In the first position ("0"), the three filament circuits are open and the anode battery is switched off, this being necessary to prevent discharge due to the leak currents in the three electrolytic capacitors when the amplifier is not in use.

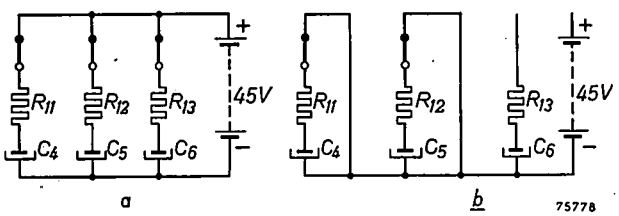


Fig. 8. *a*) In position I of the switch, the electrolytic capacitors C_4, C_5 and C_6 are "formed" at 45 V, through resistors R_{11}, R_{12} and R_{13}.
b) In position II, C_4 and C_5 are discharged through R_{11} and R_{12}.

In the second position (I) the circuit depicted in *fig. 8a* is established in order to form the electrolytic capacitors at 45 V through current-limiting resistors (C_6 is included in this to save the battery from the high initial leak current that would flow through this capacitor if placed in circuit at once). In a short time (a few seconds if the amplifier has not been used for some weeks, or about 1 minute if idle for longer periods), the leak currents are so small that the capacitors are charged almost to the voltage of the anode battery.

If the switch were now set to the normal working conditions (fig. 3), an initial voltage of 45 V would be applied to the screen grids of the first and second valves, by reason of the charge on C_4 and C_5, and this is much higher than the normal potential, which is only 10 V. The resultant heavy screen current would be detrimental to the valves, and the switch is therefore so arranged that in the third position (II), the capacitors C_4 and C_5 are discharged through resistors (fig. 8*b*).

These capacitors are connected to the screen grids only in the fourth position ("100 kc/s"), with C_6 directly across the anode battery. At the same time all the filament circuits are closed, giving the arrangement as shown in fig. 3.

The amplifier then functions in accordance with the response curve *1* in fig. 6. Once the capacitors C_4 and C_5 have been re-formed at 45 V with the switch in position I, the leak current at the 10 V now applied to them is quite low. In the long run, however, at this low working voltage, the layer of aluminium oxide becomes thinner and the leak current accordingly higher, and it is therefore advisable when the amplifier is continuously in use, to turn the switch to position I for a few moments, from time to time, to re-form C_4 and C_5 at 45 V. With continued use, a few minutes every fortnight is sufficient for this purpose.

In the fifth position of the switch ("1 kc/s"), C_9 is in circuit (see fig. 3), the response curve is in accordance with curve *2* in fig. 6 and the noise level is reduced.

Summary. An amplifier (type GM 4574) is described, designed for increasing the sensitivity of electronic voltmeters and oscilloscopes. Sub-miniature valves operated from dry batteries ensure a light, compact unit. The amplifier comprises two initial stages each giving amplification of about 63×, and a low impedance cathode-follower output stage (less than 5000 Ω). The use of considerable negative feed-back reduces the overall amplification in the straight part of the characteristic to 100 ± 3%, but renders the amplification insensitive to variations in the valve characteristics and battery voltages (amplification is reduced by at most 10% by a drop in battery voltage to half the rated value), and reduces non-linear distortion. A further object of the negative feed-back circuit is to correct the response curve. The amplification as a function of the frequency is illustrated; it is roughly 100× throughout the greater part of the range of from 1 to 150,000 c/s, with a maximum of 112× at 2 c/s and a minimum of 75× at 150,000 c/s. Noise level can be reduced by narrowing the frequency range to 1-1000 c/s. For the amplification of square-wave and pulse voltages the wide frequency band is 10-10.000 c/s and the narrow band 10-50 c/s. The noise voltage, referred to the input, is less than 10 μV in the wide band and less than 5 μV in the smaller one.

THE TROPOSPHERE AS A MEDIUM FOR THE PROPAGATION OF RADIO WAVES - II

by H. BREMMER.

621.396.11:551.510.52

*In the first part of this article [13]) six separate mechanisms were distinguished by which radio communication may be effected by microwaves *). Five of them were discussed in the first article, and the present article is wholly devoted to the remaining one, namely scattering. Such a detailed treatment is justified, not because scattering is important in itself for radio communication — it is too uncertain — but because it is largely responsible for interference in the microwave region from distant transmitters. It is therefore extremely important for such practical problems as allocating wavelengths among existing or future transmitters, and determining the minimum separation of two transmitters of the same wavelength if mutual interference is to be avoided.*

Scattering by inhomogeneities in the troposphere

So far a number of idealized cases have been distinguished whereby microwaves can be propagated through the troposphere. In each of them the refractive index depends only upon the height above the earth's surface. This was discussed in detail in the first part of this article [13]) (hereafter referred to as I). It was stated that under normal conditions, a field beyond the horizon to the transmitter was produced mainly by diffraction. However, fields have repeatedly been observed beyond the horizon of the transmitter that are much too strong to be explained by diffraction. The way in which these fields vary suggests that they are caused by inhomogeneities in the troposphere. There are always chance irregularities produced by solids, such as aeroplanes, birds, insects etc, but these cannot explain the systematically fluctuating fields, that can interfere with television reception, for example, up to several hundred kilometres from a microwave transmitter.

Fields like this may be generally attributed to scattering, i.e. to local variations in the refractive index due to small-scale inhomogeneities in the troposphere. Their scattering action on light is to be seen in the twinkling of stars, which seems to originate mainly in the lowest layer of the troposphere. Similar scattering effects are experienced by radio waves. According to a theory of Booker and Gordon [14]), it seems that the order of magnitude of the fields due to scattering agrees with the (rather scanty) meteorological data on the inhomogeneities. These data are more especially concerned with fluctuations in the wind speed, the temperature and the humidity in the lowest layer [15]).

For scattering of radio waves, only the fluctuations δn in the refractive index are important, and δn, according to formula (1) in I, results from separate fluctuations δp_a of the pressure of the dry air, δp_w of the water vapour pressure and δT of the temperature. For a typical situation with $T = 300$ °K, $p_a = 1000$ millibars and $p_w = 20$ millibars, we find:

$$\delta n = (0.26\,\delta p_a + 4.2\,\delta p_w - 1.44\,\delta T)\,10^{-6}. \quad . \quad (7)$$

Temperature and pressure change so rapidly that heat-exchange with the surroundings can be neglected. On the average, therefore, relationships hold good between the fluctuations δp_a and δT, and also between δp_w and δT, that agree with the relationship between pressure and temperature for adiabatic changes. With the help of these, (7) can be simplified to

$$\delta n \approx -2\,\delta T \times 10^{-6}. \quad . \quad . \quad . \quad . \quad . \quad (8)$$

Thus readings of temperature fluctuations give, in themselves, an idea of what fluctuations to expect in n. Gerhardt and Gordon [16]) have made observations in Texas of the variation of temperature in a small arbitrary region of the atmosphere under normal conditions. They observed temperature fluctuations of the order of 1 °C, with time constants (depending on weather conditions) between a few minutes and roughly a second. According to (8), these measurements correspond to fluctuations in n of the order of 10^{-6}, which is several hundred

*) Note that "microwave" is again used in an extended sense, to embrace all waves of $\lambda < 10$ metres.

[13]) This Review, November 1953, p. 148. Note that on p. 151 of this article (I), eight lines below fig. 3, a printing error occurs: for $\lambda = 5$ cm, read $\lambda = 5$ m.

[14]) H. G. Booker and W. E. Gordon, Proc. Inst. Rad. Engrs. **38**, 401-412, 1950.

[15]) A review is to be found in e.g. O. G. Sutton, Atmospheric turbulence, Methuen, London 1949, Chapter 2.

[16]) J. R. Gerhardt and W. E. Gordon, J. Meteor. **5**, 197-203, 1948.

times smaller than the normal value of $n-1$ (about 0.0003).

More recently Birnbaum [17]) has measured the fluctuations in n directly from the difference between the resonance frequencies of two cavity resonators, one sealed and the other open to the atmosphere. In this way he measured the maximum distance for which a correlation still existed between the fluctuations at two separate points of measurement. This distance varied, depending on the type of weather, between values of the order of 10 metres and 10 centimetres. It fixes the dimensions of what we can term separate "eddies", and the relation between these dimensions and the wavelength is all-important for the scattering phenomena. We shall come back to this point, but we shall first consider the physical conditions that determine the sizes of eddies.

Inhomogeneities caused by turbulence

The eddies are in turbulent motion with respect to each other. The eddies themselves may be of any size, from very great (e.g. atmospheric depressions) to very small, of the order of centimetres. It is the small ones that particularly interest us [18]).

In an extended medium like the atmosphere, turbulent movements can apparently be stable, provided energy is continually brought to them from outside (in the present case, radiation from the sun). When a statistical equilibrium is attained, this energy is finally dissipated as heat, through friction due to the turbulence. The dissipation is greatest for the smallest eddies, for these are responsible for the largest velocity gradients. A rapid transport of energy is therefore conducive to the formation of *small* eddies, for only then can enough friction develop to dissipate all the energy. As to the influence of viscosity, friction opposes the formation of very small eddies, since it slows down the consequent strong turbulent movements. It follows that the size l_η of the smallest eddies that play a significant rôle in the energy exchange process decreases when the rate of arrival of energy becomes greater, and increases with increasing viscosity. This agrees with the following formula, derived by Taylor [19]), which relates the "eddy size" l_η (defined by him more precisely on a statistical basis) to the coefficient of viscosity η, the energy ε absorbed per second per unit volume, and the mean kinetic energy $\frac{1}{2}\overline{v^2}$ per unit mass of turbulent gas:

$$l_\eta = \sqrt{5\frac{\eta}{\varepsilon}\overline{v^2}}.$$

The phenomenological theory leading to this result assumes an extended medium, large compared with the average size of an eddy, and further, that the turbulence is isotropic and that the effect of compressibility of the medium may be neglected (the latter is of importance only for air speeds comparable with the speed of sound). However, this theory says nothing about the statistical size-distribution of the eddies, and to do this, further physical assumptions have to be made. Using only simple dimensional arguments, Kolmogoroff arrived at a distribution for which the probability of obtaining a small eddy of size l (within which eddy the turbulence speed must be about the same everywhere) is proportional to $l^{5/3}$, but he neglected the influence of friction. After allowing for the friction, Heisenberg found that the probability for very small eddies decreased far more rapidly, namely in proportion to l^7, and that the rapid decrease begins with eddy sizes of the order of $l_{\text{crit}} = (l_\eta)^{3/2}/L^{1/2}$, where L is the maximum dimension in the medium. Friction thus imposes a limit not only on the size l_η of eddies that take a significant part in energy-exchange ("macro-eddies"), but also on the smallest eddy size l_{crit} that has an appreciable probability of existing ("micro-eddies"). The numerical value of l_{crit} is very important for scattering phenomena, as we shall see later.

Before going on to the calculation of fields due to scattering, we quote a verse attributed to L.F. Richardson, which is a variant on a well known jingle [20]). It pithily sums up what happens to the energy taken up by a turbulent medium:

> Big whirls have little whirls
> That feed on their velocity;
> Little whirls have lesser whirls,
> And so on to viscosity.

Determination of the scattering field due to a single volume element

We assume that the refractive index varies slowly enough with time for it to be regarded as constant during one period of the transmitter frequency, and further that a perceptible correlation between the values of δn at two points exists only for points

[17]) G. Birnbaum, Phys. Rev. **83**, 110-111, 1951.

[18]) A very important review of turbulence phenomena is to be found in an article by S. Chandrasekhar, Astrophys. J. **110**, 329-339, 1949.

[19]) G. I. Taylor, Proc. Roy. Soc. A **151**, 421-444, 1935 (Equation 50).

[20]) Big fleas have little fleas
Upon their backs to bite 'em;
Little fleas have lesser fleas,
And so — ad infinitum.

within a certain distance of each other — of the order of magnitude of l_{crit}.

The contributions to the field originating from scattering of the primary beam by points P within a volume element dV not smaller than a micro-eddy, may be assumed to be coherent. In contrast to this, the separate net contributions from these volume elements are to be regarded as mutually incoherent. (It should be noted that as rapid fluctuations are involved, only statistical averages of these contributions have any meaning.) As a result of this incoherence, the amplitudes of the contributions dE to the field H should not be summed but their squares, that is, their intensities.

The first step of the calculation therefore requires the intensity of the scatter field caused by a single volume element dV. The calculation is greatly simplified by assuming that the dimensions of dV are small compared with the distances TP and PR from the volume element to the transmitter and to the receiver, respectively (*fig. 8*). Just as

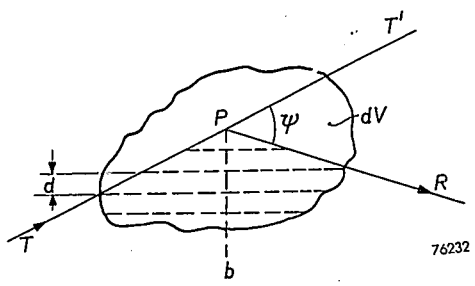


Fig. 8. The volume element dV represents a scattering element in the troposphere. TP incident ray, PR scattered ray. Angle $T'PR = \psi$ is the scattering angle. The dashed lines indicate a few planes at right angles to the bisector b of the angle TPR. The mutual separation d of these planes satisfies Bragg's law (eq. 9).

with X-ray scattering in crystals, the scattered intensity as a function of direction (RP) depends only on the distribution in space of δn which produces the scattering, where δn is, of course, a function of (x, y, z). More precisely, the intensity at R depends on only one component of the (here continuous) Fourier spectrum of δn within dV with respect to the spatial co-ordinates x, y and z. (The value of this component will be given later.) This means that a fictitious δn distribution with only this one Fourier component would produce the same scattering as the real δn distribution (in so far as this approximation is a good one). This component now corresponds to a distribution for which equal values of δn occur in planes (indicated by dotted lines in fig. 8) at right angles to the bisector b of the angle TPR: their mutual separation d is given by

$$d = \lambda/(2 \sin \tfrac{1}{2} \psi). \quad . \quad . \quad . \quad . \quad . \quad (9)$$

The scattering angle ψ is here the supplement of the angle TPR.

We recognize (9) as Bragg's law. It means that the Fourier component that completely determines the scattering corresponds to a certain array of planes with periodic values of δn, and these surfaces cause, by interference, a maximum intensity in precisely the direction PR. The analogy with Bragg's law is only possible because the fluctuations can be regarded as changing relatively slowly with time, such that the distribution of δn during a number of consecutive periods of the transmitter frequency can be thought of as fixed. The scattering is then as in a solid. An important difference from scattering by crystals is, however, that there is no orderly cell structure, so that the intensity of the scattered waves changes continuously with the direction of observation. In this respect it is more like scattering in an amorphous solid.

It follows that the intensity $d(E^2)$ of the contribution to the field from a single elemental volume dV is in the first place proportional to the intensity of the relevant Fourier component. The magnitude of $d(E^2)$ therefore depends on the location of the transmitter and the receiver, since the position of the Fourier component in the spectrum changes with these locations.

The theory also tells us that $d(E^2)$ also depends upon the physical factor $\overline{(\delta n/n)^2}$, which is the mean-square of the relative deviation of the refractive index. This makes it possible to draw a comparison between scattering of radio and of sound waves. For radio waves, according to (8), $\delta n/n$ is equal to $-2 \times 10^{-6}\, \delta T$, whilst for sound waves, for which n is proportional to $T^{-\frac{1}{2}}$, $\delta n/n$ is equal to $-\frac{1}{2}\delta T/T$. Under like circumstances, that is to say, the same condition of the troposphere and the same wavelength (it so happens that sound waves have lengths varying from a few centimetres to a few metres also), sound wave scattering produces a field intensity per unit volume greater by a factor $10^6/4T$, which is of the order of $1000\times$.

The total scattering-field at a receiver

The incoherent contributions of the separate volume elements lying between the transmitter and the receiver have to be summed by integrating $d(E^2)$ over the whole scattering space. We make the mathematics less difficult by confining ourselves to the case of a transmitter T beamed on a receiver

R (*fig. 9*). We shall assume that both have polar diagrams with half-value angles [21]) of $\Theta_{\frac{1}{2},h}$ in the horizontal plane, each bisected by the vertical plane through T and R, and a half-value angle $\Theta_{\frac{1}{2},v}$ in the vertical plane with respect to the local horizon. It is assumed that $\Theta_{\frac{1}{2},h}$ and $\Theta_{\frac{1}{2},v}$ are both small angles. As far as reception is concerned, only

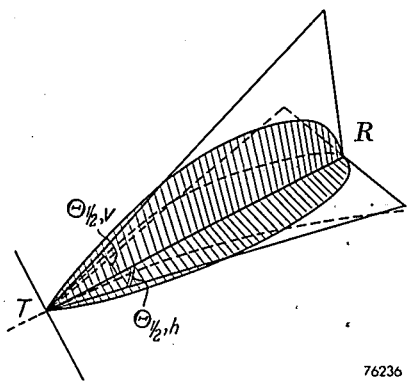


Fig. 9. Transmitter T with beam directed at the receiver. The beam has a polar diagram in the horizontal plane with half-value angle $\Theta_{\frac{1}{2},h}$ which is bisected by the vertical plane through transmitter and receiver, and a polar diagram in the vertical plane with half-value angle $\Theta_{\frac{1}{2},v}$ with respect to the horizontal planes. The angles $\Theta_{\frac{1}{2},h}$ and $\Theta_{\frac{1}{2},v}$ have been exaggerated in the diagram for clearness. The receiver is assumed to have a similar polar diagram.

scattering occurring within the volume defined by the intersection of the transmitter and receiver half-value angles is of importance. In *fig. 10* the

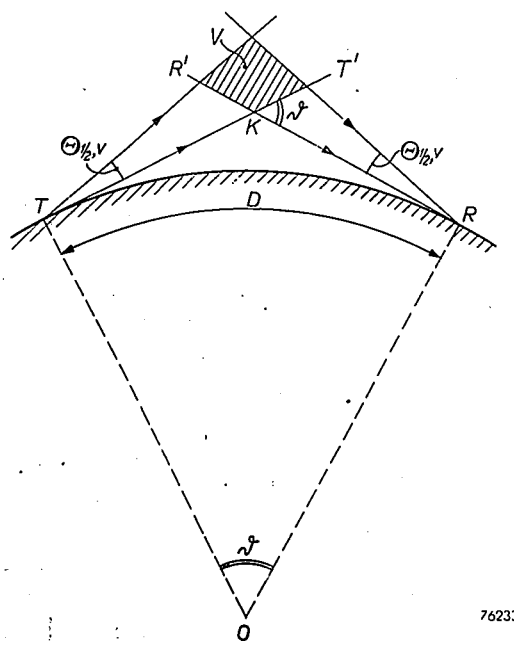


Fig. 10. From the point of view of reception, for beams like those in fig. 9, the scattering is important only when it takes place within a volume V; its cross section in the plane of the paper is shaded. Angle $T'KR$ = angle $TOR = \vartheta$.

[21]) The half-value angle is the angle between directions for which the intensity is half that along TR (fig. 9).

cross-section of this volume in the plane through T and R — the plane of the diagram — is shown shaded.

Booker and Gordon [14]) have treated many cases, of which we shall only consider those for great distances, for then the scattering fields are relatively the most important. The scattering volume is then confined to the neighbourhood of the plane bisecting TR, and it is small compared with the distance $\frac{1}{2}D = \frac{1}{2}a\vartheta$ to transmitter and receiver (D = arc TR, ϑ = angle TOR, a = radius of the earth), provided that

$$\vartheta \gg \Theta_{\frac{1}{2},v} \text{ and } \Theta_{\frac{1}{2},h} \ll \tfrac{1}{2}\pi.$$

All the scattering angles are then approximately the same, and we can take them equal to $T'KR$, (fig. 10), corresponding to the rays TK and KR, which are horizontal at the transmitter and receiver respectively. As is clear from the figure, the angle ψ (of fig. 8) is equal to the angle ϑ subtended by transmitter and receiver at the earth's centre. Since the field intensity depends on D, the total scattering field depends on the angle ψ.

This raises the question: for what values of ψ can appreciable scattering be expected? We can answer this qualitatively, provided that δn changes only very slightly over distances of the order of l_{crit}. When this is so, the components of the volume Fourier spectrum of δn must of course be negligible for periods less than l. Using this in eq. (9), when only Fourier components with periods greater than l contribute to the scattering, the corresponding angles ψ must satisfy

$$\frac{\lambda}{2\sin\frac{1}{2}\psi} > l, \text{ or } \sin\tfrac{1}{2}\psi < \frac{\lambda}{2l}. \quad . \; . \quad (10)$$

The condition is easiest to see from its two limiting cases:

1) $\lambda \gg l$. The scattering is appreciable in all directions satisfying (10). This is the well-known Rayleigh scattering, which is more or less isotropic.
2) $\lambda \ll l$. Scattering of importance occurs only through very small angles ψ; the scattering therefore takes place mainly in the forward direction.

For great distances with a large angle ϑ, the most marked contribution to the field at the receiver is found in the first case ($\lambda \gg l$), where there is no limit on the scattering angle. The importance of the parameter l now becomes evident; this is apparently of the order of l_{crit}, that is, the smallest eddy size that is at all probable (micro-eddies, see the end of the section on turbulence).

Evaluation of the scattered field

We have just seen that the total scattered field depends, in particular, upon one Fourier component of the distribution in space of δn in a coherently scattering volume δV. It appears from a fuller investigation that this dependence can be explained by statistical properties of the space distribution of δn. All we need to know is the following so-called correlation function

$$C(a,b,c) = \frac{\overline{\delta n(x,y,z) \times \delta n(x+a,\ y+b,\ z+c)}}{\overline{\{\delta n\ (x,y,z)\}^2}}.$$

In this definition the bars denote means over the volume elements, which must not be smaller than a micro-eddy. This function gives the correlation between the fluctuations of n at two points with cartesian co-ordinates differing by a, b and c respectively, so that for isotropic turbulence they depend only on the distance $(a^2 + b^2 + c^2)^{\frac{1}{2}}$. At very large distances there is no correlation and the correlation function decreases to zero. The function is also normalized; its maximum value is at zero distance and is there unity. Apart from these limiting values, the function decreases monotonically with increasing distance. As examples of mathematically tractable cases we mention two isotropic ones

$$C(a,b,c) = \mathrm{e}^{-(a^2+b^2+c^2)^{\frac{1}{2}}/l} \quad . \quad (11a)$$

and

$$C(a,b,c) = \mathrm{e}^{-(a^2+b^2+c^2)/l^2} \quad . \quad . \quad . \quad (11b)$$

In both of them the correlation begins to decrease sharply when the distance $(a^2 + b^2 + c^2)^{\frac{1}{2}}$ becomes of the order of l, so that l can be interpreted as a measure of the average micro-eddy size. The field strengths E corresponding to these correlation functions at great distances from the transmitter can now be found. For a transmitter radiating a power of 1 kW, these field strengths are respectively

$$E = \frac{4.2\times10^8 \times l_{\mathrm{cm}}^{\frac{3}{2}} \{\overline{(\delta n)^2}\Theta_{\frac{1}{2},\nu}\}^{\frac{1}{2}}}{\lambda^2_{\mathrm{cm}}\, D_{\mathrm{km}} \left\{1 + \frac{16\pi^2 l^2}{\lambda^2}\sin^2\left(\frac{D}{2a}\right)\right\}} \ \mathrm{mV/m}, \quad (12a)$$

$$E = \frac{2.0\times10^8 \times l_{\mathrm{cm}}^{\frac{3}{2}} \{\overline{(\delta n)^2}\Theta_{\frac{1}{2},\nu}\}^{\frac{1}{2}}}{\lambda^2_{\mathrm{cm}}} \times \frac{\mathrm{e}^{-\frac{2\pi^2 l^2}{\lambda^2}\sin^2\left(\frac{D}{2a}\right)}}{D_{\mathrm{km}}} \ \mathrm{mV/m}. \quad . \quad . \quad . \quad (12b)$$

(E is the r.m.s. field strength in millivolts per metre. Units of length are given as suffixes, except in dimensionless ratios of lengths.) These formulae clearly demonstrate the influence of the form of the correlation function. If $\lambda \ll l$, the first case corresponds to a decreasing field approximately proportional to D^{-3}, and the second case to a field that decreases roughly in proportion to $D^{-1}\exp(-\gamma D^2)$, where γ is a positive constant.

Interpretation of observations in terms of scattering

Although so much depends on the precise form of the correlation function C, which is not known, field strengths far beyond the horizon in general appear to be satisfactorily explained by the Booker-Gordon theory [14]). This theory starts from the correlation function (11a) and leads to a field like that in (12a) that decreases not exponentially, as does a diffraction field, but more slowly.

The theoretical results can also be compared with observations relating to the distance up to which television reception from one transmitter can be interfered with by another working at the same wavelength. As with the analysis of ionospheric propagation data, from averages of a large number of observations, field strength - distance curves can be plotted, which give the field strengths exceeded during e.g. 50%, 20% or 10% of the time. For any one distance, the greater the field strength, the smaller the percentage and the smaller the influence of the diffraction field. The latter varies as the effective radius of the earth gradually changes, but these changes are certainly not large enough to explain the rather large field strengths that are observed only during, say, 1% of the time. It appears that these field strengths are to be attributed entirely to sporadic, sometimes very strong, scattering phenomena.

In this connection the curves obtained by the Comité Consultatif International des Radiocommunications [22]) may be mentioned. These are for fields from a half-wave dipole, exceeded during 1% and during 10% of the time. For example, for distances D between 100 and 500 km, and frequencies between 50 and 200 Mc/s, it appears that the field strength E exceeded during 1% of the time can be fairly accurately represented by

$$E = 85\,000\ D^{-3}\ \mathrm{mV/m}$$

(with D in km, and again for 1 kW radiation). The factor D^{-3} is consistent with (12a) provided $l \gg \lambda$, for then the term 1 in the denominator can be neglected, and also $\sin D/2a$ can be replaced by $D/2a$, so that a factor D^3 appears in the denominator. However, this does not confirm the fundamental assumptions underlying (12a), for a similar decrease of field strength can be obtained, for example, if the turbulence is not isotropic.

[22]) C.C.I.R. Documents of the 6th plenary assembly, Geneva 1951, page 55.

This is an important point. E. G. Richardson [23]) has made direct measurements of air turbulence in the lowest layers, and these show not only that the eddy size increases with height, but also that its isotropic character, in so far as it exists close to the earth's surface, is lost. Moreover, from measurements in aircraft (N. S. Wong), it now seems likely that the refractive index varies far less horizontally than vertically. This is also confirmed by Herbstreit and Norton, who studied the correlation between the field strengths from a transmitter, measured at two receivers placed close to each other. The correlations were obtained for different directions of the line joining the receivers, and comparing them it appears that the correlation extends over a greater distance horizontally than vertically.

These observations led Feinstein to reconsider the simple model of a refractive index depending only on the height, and to assume in addition that the vertical gradient of the refractive index fluctuates around a constant mean. The theory thus assumes a strongly anisotropic turbulence. With suitable assumptions concerning the ratio of the mean vertical dimension of the eddies to the wavelength, this theory, too, can yield a field strength proportional to a negative power of the distance D, as does the theory based on isotropic turbulence. Field strength measurements at large distances are therefore still insufficient to decide between isotropy and anisotropy. More comprehensive observations on scattered fields, combined with meteorological data, are required to extend our knowledge of turbulence.

Fading explained by scattering

The scattering mechanism can also explain fading, which occurs especially at great distances. In fact, scattering apart, fading is always liable to occur whenever field strengths of the same order of magnitude are contributed by two or more different modes of propagation (e.g. direct paths, reflections at transition layers, diffraction, etc). These contributions depend on nearly independent physical conditions, and this can make their phase differences change very quickly, so that the total field strength fluctuates violently. However, for great distances, where scattering effects dominate all other contributions to the field, fading can be explained successfully only by scattering, according to the mechanism first proposed by Ratcliffe [24]) for ionospheric fading.

The eddies in the troposphere can be regarded as stationary during one period of the high-frequency radiation, but in fact they do move slightly. Even when their movement is so slow that it does not influence the average field, the scattered field due to the eddies will be subjected to a Doppler effect, and even this can cause small variations of field strength about the mean value. If v_b is the component of the velocity in the direction of the bisector of the incident and scattered rays (direction b in fig. 8), and c is the velocity of light, it is found that for a scattering angle ψ, a Doppler displacement $\Delta\nu$ with respect to the frequency ν of the carrier wave exists, such that

$$\frac{\Delta\nu}{\nu} = \frac{2v_b}{c} \sin \tfrac{1}{2}\psi.$$

The field is thus modulated with a frequency $\Delta\nu$, and fading develops with a time constant

$$T = \frac{1}{\Delta\nu} = \frac{\lambda}{2v_b \sin \frac{1}{2}\psi}.$$

For large distances ($\vartheta \gg \Theta_{\frac{1}{2},v}$), ψ can be put equal to the angle $\vartheta = D/a$ between transmitter and receiver. Fading then occurs with a time constant dependent on the tropospheric parameter v_b, and on the value of the fixed parameter λ/D. The tempo of the fading fluctuations thus becomes quicker with increasing distances. An analysis of the fading phenomena could therefore provide information on the air velocities characterising the turbulence.

Recapitulation

In conclusion, this section recapitulates the many different factors that control the propagation of short radio waves in the troposphere. Their distinguishing characteristics are outlined.

Communication between a transmitter and receiver comes about in one of the following ways, or through a combination of them:

a) *Direct propagation:* the total field is due to two rays, both being refracted according to Snel's law: one is the direct ray and the other undergoes a reflection at the earth's surface. *Occurrence:* whenever the receiver is above the horizon of the transmitter. *Reception characteristic:* the field is approximately proportional to D^{-1} for all wavelengths.

b) *Propagation via transition layers:* the field originates along a ray reflected (usually partially) at a layer in the troposphere where the refractive index varies abnormally strongly with height. *Occurrence:* particularly when a temperature in-

[23]) E. G. Richardson, Proc. Roy. Soc. A **203**, 149-164, 1950.
[24]) J. A. Ratcliffe, Nature **162**, 9-11, 1948.

version or a meteorological front gives rise to the required transition layer. *Reception characteristic:* over a limited region, the field decreases more slowly than corresponds to a factor D^{-1}, and increases somewhat with the wavelength.

c) *Propagation as in a waveguide:* the propagation is brought about by multiple reflection of many rays between the earth's surface and the top of the "waveguide". This upper boundary is a region of indefinite thickness, at which continuous refraction — resulting in reflection — takes place. *Occurrence:* only under exceptional conditions, when with increasing height, a rapid decrease in humidity is accompanied by a rapid rise in temperature. *Reception characteristic:* for very short wavelengths, a field is observable below the horizon that is much stronger than the normal field brought about by diffraction, see (e), and decreases with increasing wavelengths.

d) *Gradient reflections:* the field is produced by partial reflections at the infinite number of infinitely thin layers of constant refractive index which constitute the continuous variation of refractive index with height in the troposphere. *Occurrence:* always present, but not noticeable except when they contribute to the field below the horizon due to diffraction. *Reception characteristic:* the decrease in the diffraction field just under the horizon is less pronounced than usual.

e) *Diffraction:* the field is produced by deflections of the rays towards a region beyond the horizon unattainable by geometrical-optical rays. *Occurrence:* always present, but contributing a significant field only at moderate distances beyond the horizon. *Reception characteristic:* the field decreases exponentially with distance, without being subject to pronounced fluctuations in time.

f) *Scattering:* the field is caused by scattering at small inhomogeneities that are present throughout the troposphere. *Occurrence:* always present, but only noticeable at great distances, where the other modes of propagation no longer contribute significantly to the field strength. *Reception characteristic:* the field at great distances decreases more slowly than exponentially, and is also very sensitive to weather conditions.

Summary. This second part of a survey of radio propagation in the troposphere deals with scattering. Inhomogeneities (eddies) are discussed, which occur in all sizes and are subject to turbulent motion. The smallest eddies that are at all likely to occur (micro-eddies) are especially important for scattering. In order to evaluate the scattered field, the field due to the scattering of one volume element is first considered. An interpretation then appears to be possible in accordance with Bragg's well known law. The total scattered field at a receiver is then found by integration over all active volume elements. The mathematical difficulties are minimized by assuming that the transmitter and receiver are beamed on one another. If a mathematically tractable function is now chosen for the volume fluctuations of the refractive index, numerical results can be obtained. Two such functions are mentioned together with the corresponding scattered fields. Observational data are next considered from the point of view of scattering, including observations of fading. The article ends with a recapitulation of the six idealized ways by which radio waves can be propagated in the troposphere, and gives their special characteristics.

PHOTOMETRY AT LOW LUMINANCE LEVELS

by W. de GROOT.

535.24

The establishment, by the Commission Internationale de l'Éclairage (Stockholm 1951), of a relative efficiency function for the dark-adapted eye (scotopic efficiency function) and the discussions held on that occasion, are once more drawing attention to the problem of photometry at low levels of illumination and corresponding low luminances. A survey of this matter is given below, which may be considered as a supplement to earlier articles on the same subject).*

Photometry

If a flat surface, for example a diffusely reflecting white screen, is exposed to radiation with wavelength(s) between 0.4 and 0.7 μ, an observer will obtain from it an impression of light — provided the power radiated on to the screen is not too low — and he will be able to form a judgment as to the brightness of the screen. This subjective "brightness impression" depends on a number of conditions. The principal ones are:

1) the power of the radiation emitted by the screen per unit area and the spectral distribution of the radiation,
2) the brightness of adjoining parts of the field of view and the state of adaptation of the eye,
3) the part of the retina involved in the observation.

With (1) we shall deal more in detail afterwards. As an example of the influence of the factors mentioned under (2) the appearance of the moon at night and in the daytime may be considered. If we look at the moon against a clear sky in daylight, we see it as a whitish patch on a blue background. After sunset the same moon appears to us as a glaring white disk with black surroundings. Nevertheless, the radiation which the moon sends to our eye is the same in both cases; only the brightness of the surrounding sky has diminished [1]).

An example of (3) is the well-known phenomenon of the disappearance of a small and feeble light at which we stare in the dark. It is caused by the fact that, in the dark, the central part of the retina (the fovea) is much less sensitive to light than the surrounding part (the parafovea) and the rest of the retina (the periphery). In photometry, the observer makes himself independent of this variability of the brightness impression, by comparing only fields of approximately equal brightness. In instruments constructed for the purpose (photometers), the observer sees two adjacent fields of view, one illuminated by the unknown source and the other — used for comparison — by a source of adjustable and known intensity.

In view of what has been mentioned under 2) and 3), care has to be taken that both fields of view are in the same condition as regards their surroundings and the part of the retina involved in the observation. The two fields of view may, for example, be the two halves of a circular field (*fig. 1*). The

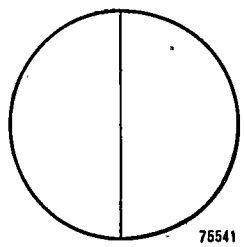


Fig. 1. Photometer field (angular diameter 2° in photopic region, 6°-10° in scotopic region).

surroundings may be either dark or they may have a certain brightness, for instance, about equal to that of the circular field. The viewing angle of the photometric field will be chosen according to circumstances. With high brightnesses, at the level of illumination which makes colour vision possible, an angle of 2° is usually chosen, so that only the fovea and its immediate neighbourhood take part in the observation. With low brightnesses, having regard to the low sensitivity of the fovea, a larger angle (e.g. 6° to 10°) is chosen, so that the influence of the fovea may be disregarded and the observation is effected mainly by the parafovea.

Influence of intensity and spectral distribution on the judgement of brightness

The influence of the conditions mentioned under (1), above, will now be considered, beginning with the case of high brightnesses (luminance > 3 cd/m^2).

*) See P. J. Bouma, Philips tech. Rev. **1**, 102-106, 142-146, 1936; G. Heller, Philips tech. Rev. **1**, 120-125, 1936; **5**, 1-5, 1940; W. de Groot, Philips tech. Rev. **10**, 150-153, 1948.

[1]) The radiation which we receive from the moon in the daytime is actually slightly greater, since the light of the sky itself is superimposed on that of the moon proper.

High brightness levels

At high brightness levels, the observer obtains impressions of colour and brightness simultaneously. It is difficult to separate the colour impression from the brightness impression. When comparing the two fields in a photometer, one illuminated by monochromatic green light, the other by monochromatic red light, it is difficult to state whether the brightness of the two patches is the same, although after some practice, judgement in this respect is improved.

In such cases, a step-by-step method is often used. The green colour is first compared to another colour, somewhat nearer (say yellow-green); this is compared to a colour which is still more yellowish, and so on until red is reached. This is convenient because the adjustment to equal brightness is much easier if the colour difference is small.

If, in this way, the whole visible spectrum is run through, it will be seen that the ratio of powers P_λ, to be radiated per unit area and solid angle in the direction of the eye, to ensure equal impressions of brightness at each wavelength, are different for each pair of wavelengths. For a certain wavelength (λ_m) in the neighbourhood of 555 mμ (5550 Å) the power is a minimum. The ratio $P_{\lambda m}/P_\lambda = v(\lambda)$ is called the relative efficiency of the observer at the wavelength λ. For $\lambda = \lambda_m$, $v(\lambda) = 1$, for other wavelengths $v(\lambda)<1$.

If the experiment is repeated with the same observer at a still higher brightness level, the same value of λ_m and the same function $v(\lambda)$ are found. Thus in the region of high brightness levels, $v(\lambda)$ is independent of the level of illumination. If, on the other hand, the experiment is made by different observers, different functions $v(\lambda)$ and different values of λ_m will generally be found. The CIE (or Commission Internationale de l'Éclairage) therefore standardised, in 1924, an efficiency function which is considered to be valid for a "normal observer". We shall denote this function by $V_p(\lambda)$, because the region of high brightnesses is also called the *photopic* region [2]). The corresponding value of λ_m we shall denote by λ_p.

Unless the contrary is indicated, we shall suppose in the following that the observer is a "normal observer".

Suppose that the radiation to be investigated is not monochromatic but is provided by a source with spectral distribution $P'(\lambda)$. This means that the power [3]) (per m^2 and per unit solid angle) radiated in the direction of the eye in the interval λ, $\lambda+d\lambda$ is $P'(\lambda)d\lambda$.

If the brightness impression in the photometer is equal to that of a monochromatic power P_m at wavelength λ_m, experiment shows that

$$P_m = \int P'(\lambda)\ V_p(\lambda)\ d\lambda. \quad . \quad . \quad . \quad . \quad (1)$$

The expression on the right-hand side of equation (1), multiplied by a suitable factor, may also serve as a measure of the brightness. The usual unit is the candela per m^2 (cd/m^2) or nit. If the brightness is defined in this way (and preferably measured in this unit) this quantity is given the name luminance (L). The unit cd/m^2 is so chosen that a perfect radiator ("black body") at the freezing temperature of platinum ($\approx$2042 °K) has a luminance of 6×10^5 cd/m^2.

From the known spectral distribution of the perfect radiator it is possible to calculate the value of the constant (K_p) occurring in the formula

$$L = K_p \int P'(\lambda)\, V(\lambda)\, d\lambda \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The value [4]) of K_p is about 650 lm/W.

Low brightness levels

If the experiments just described are repeated with luminances less than 3 cd/m^2 the wavelength λ_m is found to be displaced to shorter values. This displacement depends on the part of the retina involved in the observation [5]). For the fovea the displacement is small, but at decreasing intensity the ability of the fovea to detect light soon ceases. As regards the periphery, already at high luminances the wavelength λ_m does not coincide with the value λ_p characteristic of the fovea. At decreasing intensity, λ_m shifts to still lower values. If the

---

[2]) From the Greek photos = light, and opsis = vision, in contrast with scotopic, from the Greek skotos = darkness.

[3]) In the literature, the symbol E(λ) is often used. We prefer the letter P because the quantity under consideration is a power and not an energy, and provide it with a dash to keep the reader aware of the fact that $P'(\lambda)$ has not the dimension of a power but of a power per unit of (wave) length, and thus has the character of a derivative with respect to λ.

[4]) If 3.74×10^{-16} W.m^2 and 14385 μ. °K are taken as the values of the constants $c_1 = 2\pi hc^2$ (radiation over a solid angle of 2π) and $c_2 = ch/k$ respectively, in Planck's radiation formula, and if the freezing temperature of platinum is based on optical pyrometry, taking 1336 °K as the melting temperature of gold, one finds K_p = 683 lm/W (see Philips tech. Rev. **10**, 150-153, 1948). This value is higher than the value found experimentally (about 630 lm/W). This points to the fact that the real melting temperature of gold is probably higher than the internationally accepted value (between 1340 and 1342 instead of 1336 °K; see F. Henning, Temperaturmessung, Barth, Leipzig 1951, pp. 140 and 256, W. de Groot, Physica **16**, 419, 1950).

[5]) See W. D. Wright, Researches on normal and defective colour vision. Kimpton, London 1946, pp. 73-87.

experiment is arranged in such a way that at high luminances the fovea alone takes part in the observation and at lower brightnesses the parafovea is allowed to take part in a prescribed way (e.g. by a gradual increase of the visual angle), it is permissible to say that the value of λ_m gradually shifts from the photopic value λ_p to lower values. At the lowest perceptible brightnesses, this shift comes to a standstill. The wavelength λ_m has then assumed the value 507mμ, which we shall denote by λ_s, because the region of lowest perceptible brightnesses is also called the *scotopic* region [2]).

With a special experimental arrangement, it is possible to define unique values of λ_m, and of an efficiency function $V(\lambda_m,\lambda)$, corresponding to each value of P_m. Conversely, a unique value of P_m may be found corresponding to each value of λ_m, between λ_s and λ_p. The form of the function $V(\lambda_m,\lambda)$ in which λ_m occurs as a parameter, depends on λ_m. If a field with spectral distribution $P'(\lambda)$ appears equally bright as a field radiating monochromatic light of wavelength λ_m with a power P_m, experiment shows [6]) that

$$P_m = \int P'(\lambda)\, V(\lambda_m,\lambda)\, d\lambda.$$

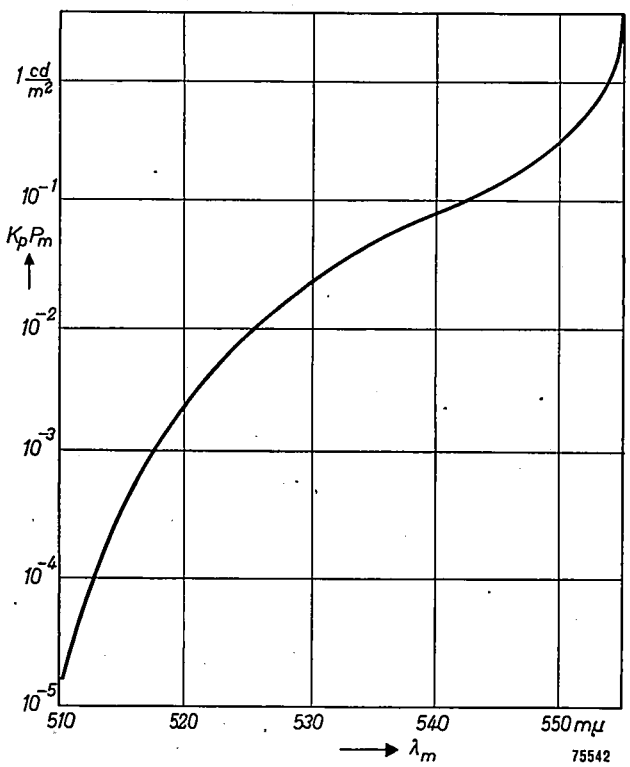


Fig. 2. P_m as a function of λ_m (Weaver). The quantity $K_p P_m$ (expressed in cd/m²) is plotted on a logarithmic scale.

[6]) P. J. Bouma, Proc. Kon. Acad. Wet. Amsterdam, **38**, 33-35, 1935, P. J. Bouma, Physical aspects of colour, Ch. II (Philips technical Library, 1947). See also W. de Groot, Appl. sci. Res. **B2**, 131-148, 1951, J. Vorobeitchik, Bull. Soc. Belge. Electr. **68**, (1952), No. 2.

The connection between λ_m and P_m was first investigated in 1890 by König [7]).

Up till now the functions $P_m(\lambda_m)$ and $V(\lambda_m,\lambda)$ have not been standardised, as agreement has not

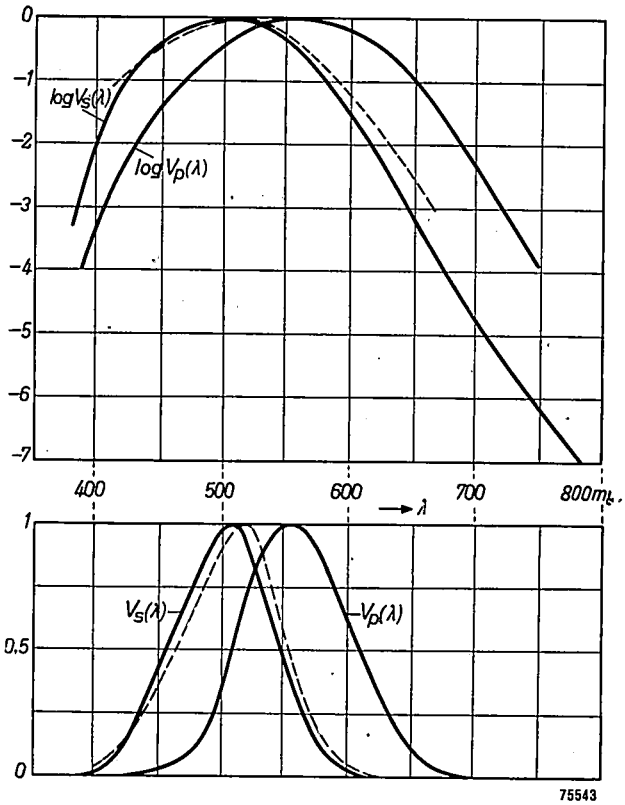


Fig. 3. The functions $V_p(\lambda)$ (CIE 1924) and $V_s(\lambda)$ (CIE 1951) plotted against λ. In the upper graph the ordinate is logarithmic, and in the lower graph, linear. The broken line corresponds to Weaver's $V_s(\lambda)$-function.

yet been reached on the way in which the experiment should be arranged. In 1949 Weaver [8]) published a series of tables for the functions $V(\lambda_m,\lambda)$, which may serve as a basis for calculations. In 1951, at the CIE conference at Stockholm, a function $V_s(\lambda)$ was standardised [9]), valid for young eyes, in the region of very low brightnesses. In *fig. 2* the connection between P_m and λ_m according to Weaver is shown, and in *fig. 3* the functions $V_p(\lambda)$ and $V_s(\lambda)$, as given by the CIE.

Photometry at low brightnesses

What are the consequences of all this for photometry at low brightness levels? Suppose that in a photometer, two fields of view are observed,

[7]) See P. J. Bouma, Philips tech. Rev. **1**, 142-146, 1936. P. G. Nutting, Bull. Bur. Stand. **7**, 236-238, 1911.

[8]) K. S. Weaver, J. Opt. Soc. Amer. **39**, 278-291, 1949.

[9]) Proc. CIE (Stockholm 1951), part III, pp. 33-39. For the sake of clarity in the present article, the symbols K_m, K_m', V_λ, V_λ', as used by the CIE, are replaced by K_p, K_s, $V_p(\lambda)$ and $V_s(\lambda)$.

radiating light of wavelengths λ_1 and λ_2 respectively, with corresponding powers P_1 and P_2 such that the observations are made in the photopic region. When luminances are equal, we have

$$K_p P_1 V_p(\lambda_1) = K_p P_2 V_p(\lambda_2).$$

If we now decrease both powers by multiplying them by the same factor (<1) — this may be done, for example, by inserting a neutral filter — so that we are no more in the photopic region, both fields will, according to definition, still have the same luminance, but to the eye they will no longer appear equally bright. Conversely, two fields appearing equally bright will no longer have the same luminance if we are not in the photopic region.

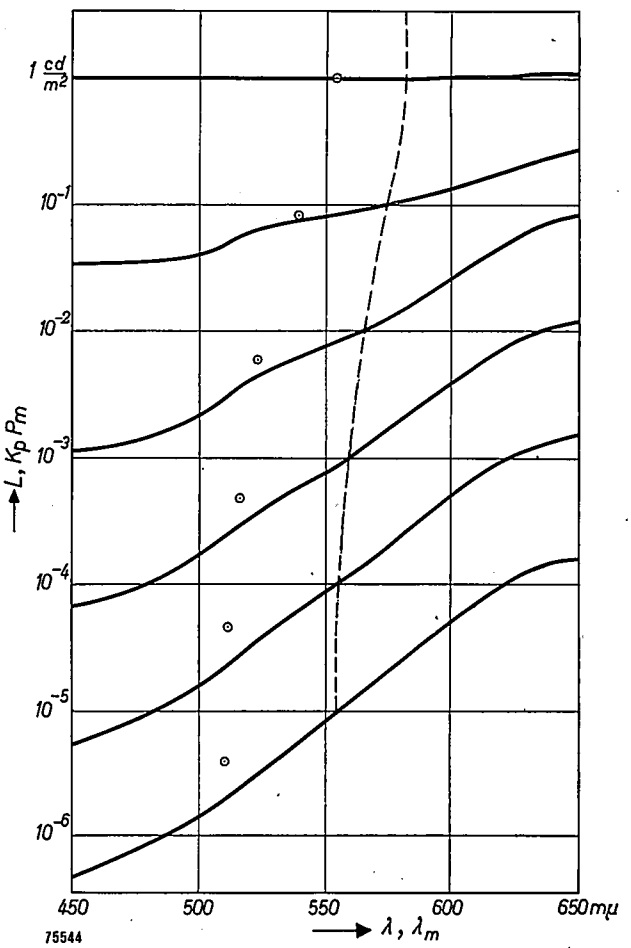


Fig. 4. Lines of equal "subjective brightness" for monochromatic radiations, according to Bouma. Ordinates (logarithmic scale) are luminances L in cd/m². In the same diagram, values of $K_p P_m$ (see fig. 3) are plotted against λ_m. The broken line joins the wavelengths for which the luminance is equal to that given by a source of $T_c = 2042$ °K, appearing equally bright.

In *fig. 4*, which is due to Bouma, the luminances L of monochromatic fields appearing equally bright, are given on a logarithmic scale. Corresponding values are joined by a curve (full lines), each curve representing a certain brightness level. The values of $K_p P_m$ and λ_m corresponding to each curve are also indicated.

The function plotted against λ is

$$\log [K_p P_\lambda V_p(\lambda)] = \log [K_p P_m V_p(\lambda)/V(\lambda_m,\lambda)].$$

In the photopic region $\lambda_m = \lambda_p$ and $V(\lambda_m,\lambda) \equiv V_p(\lambda)$. Therefore, in this region, the value is constant and equals $\log [K_p P_m]$. In the region of low brightnesses, $\lambda_m < \lambda_p$ and $V_p(\lambda_m) < 1$, whereas $V(\lambda_m,\lambda_m) \equiv 1$. Therefore, the point with coordinates λ_m and $\log (K_p P_m)$ is above the corresponding curve.

The points of any one curve in fig. 4 correspond, as already indicated, to fields appearing equally bright. The curves are plotted in order of increasing apparent brightness.

We now put the question: is it possible to define a "measure" for this apparent brightness? To each curve corresponds a value of P_m. The higher the value, the higher the curve will be in the diagram. It is natural to use P_m as a measure of apparent brightness. We may also use $K_p P_m$, i.e. the luminance of the field radiating monochromatic light of wavelength λ_m with power P_m, which appears equally bright. But it is evident that an arbitrary monotonic function $f(P_m)$ (i.e. a function which always increases with increasing P_m) also satisfies the requirements [10]).

Bouma, following König, chose the luminance of a field appearing equally bright and radiating monochromatic light of wavelength $\lambda = 535$ mμ as a measure of what he called "subjective brightness". This definition is unambiguous, but it is not well suited for photometric practice, because a source radiating this wavelength is not readily available.

It is, however, not necessary to restrict oneself to monochromatic sources, because the choice of the function $f(P_m)$ is arbitrary within wide limits. A definition which has found practical application uses as a light source an incandescent lamp of colour temperature 2360 °K. This, of course, is the type of lamp used in most photometers. The subjective brightness is now made equal to the luminance of a field appearing equally bright and radiating light with a relative spectral distribution equal to that of a black body of the said temperature. The advantage of this definition is that now "subjective brightnesses" can be measured using an ordinary photometer.

The colour temperature of 2360 °K is rather arbitrary, and another choice might be considered.

---

[10]) P. J. Bouma, Physica 8, 413-424, 1941.

At the CIE conference 1951 this question was discussed and agreement was attained on the desirability — at least from a theoretical point of view — of using the same light source, for the measurement of the "subjective brightness" as used to fix the unit of light intensity (candela) at high luminances, viz. a perfect radiator with a temperature of about 2042 °K.

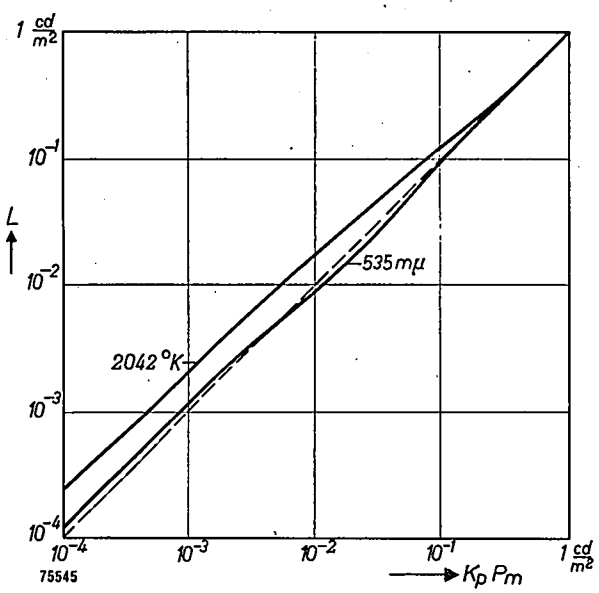


Fig. 5. Luminance of photometer fields, illuminated by two different sources ($\lambda = 535$ mμ, $T_c = 2042$ °K), plotted against K_pP_m (both logarithmic). The broken line would correspond to K_pP_m equal to L. The scotopic region is the region where $K_pP_m < 10^{-4}$ cd/m², and the photopic region where $K_pP_m > 3$ cd/m².

In *fig. 5* the luminances of two radiations (monochromatic, $\lambda = 535$ mμ, and black body with $T = 2042$ °K) are plotted against K_pP_m (both on a logarithmic scale). This figure therefore shows the function $f(P_m)$, both for the definition of König-Bouma and for that of the CIE. It is seen that at high brightnesses the two functions are identical, and at the lowest brightnesses (in the scotopic region) K_pP_m and the "subjective brightness" are again proportional to each other. Their ratio [11]), however, depends on the choice of $f(P_m)$. The latter fact simplifies the case in the scotopic region with which we are often concerned. In this region the "subjective brightness" is represented by the expression

$$K_s \int P'(\lambda)\, V_s(\lambda)\, d\lambda,$$

which formula is analogous to formula (2) for the luminance, with the difference that $V_p(\lambda)$ is replaced by $V_s(\lambda)$ and the factor K_p by K_s. The value of K_s depends on the definition of $f(P_m)$, and therefore on the choice of the standard source. For the CIE source ($T_c = 2042$ °K), taking the CIE values for $V_s(\lambda)$, $K_s = 1746$ lm/W.

The question arises, whether a light source exists for which, in the entire region, the subjective brightness is simply proportional to P_m, in which case $f(P_m)$ would be represented by a straight line in fig. 5. Such a light source does not exist, as was pointed out by Bouma. If a source with colour temperature $T_c = 7500$ °K were chosen [12]), K_s would be equal to K_p.

Another way of arriving at the same result [13]) would be to choose a monochromatic source, for which $V_s(\lambda) = V_p(\lambda)$, i.e. a source with wavelength $\lambda = 528$ mμ. In both cases, however, $f(P_m)$ is not proportional to P_m in the intermediate region. The deviation from the straight line is greater for the curve corresponding to $\lambda = 528$ mμ than for that corresponding to $T_c = 7500$ °K.

The proposal of the CIE

The choice of the CIE reduces photometry at low brightness levels to a simple question (at least in theory). The comparison field in the photometer is to be illuminated with light of the candela-source, which is to be attenuated by a neutral attenuator such as a rotating sectored disc, a neutral filter or a diaphragm. The attenuation is to be regulated in such a way that the comparison field, illuminated by the standard source, appears equally bright as the field to be measured. The result is obtained in units cd/m², as in the case of high luminances.

A difficulty arises because the sensitivity curve of the parafovea differs from that of the fovea, as explained above. Hence, in reality the question is not so simple, owing to the great difference between the spectral distribution of most light sources and that of the source of colour temperature $T_c = 2042°$ K. Even if the colour difference is no hindrance in itself (as is the case with parafoveal or peripheral observation in the intermediate and in the scotopic region), measures have to be taken to ensure that the fovea is not involved in the observation [14]).

In practice, therefore, a photometer lamp with a higher colour temperature (2360 °K or more) is chosen. The measured luminance will then have to be corrected, in order to reduce it to the desired luminance of a field of 2042 °K appearing equally bright. In the scotopic region this may be done simply by multiplying by a constant factor.

One question remains: in which unit is the "subjective brightness" to be expressed? If the

[11]) See also W. D. Wright, Photometry and the eye, Hatton, London 1949.

[12]) W. de Groot, see note [6]).

[13]) P. J. Bouma, see note [10]).

[14]) A case in which these conditions were not observed is described in Proc. CIE 1951 (III), Gen. Paper D: A. M. Kruithof and W. de Groot, Photometry at very low brilliance levels.

recommendations of the CIE are followed, a result expressed in cd/m² is obtained. It is then possible, however, that for a certain lamp, a different number of lm/W is obtained depending on whether the illumination belongs to the scotopic or the photopic region (as may be the case with the illumination of roads).

In general, the scotopic radiation equivalent

$$k_s = K_s \frac{\int P'(\lambda)\, V_s(\lambda)\, d\lambda}{\int P'(\lambda)\, d\lambda} \text{ lm/W (s)}, \quad . \quad . \quad . \quad (3)$$

differs appreciably from the photopic equivalent

$$k_p = K_p \frac{\int P'(\lambda)\, V_p(\lambda)\, d\lambda}{\int P'(\lambda)\, d\lambda} \text{ lm/W (p)}. \quad . \quad . \quad . \quad (4)$$

Omission of the addition "scotopic" (s) or "photopic" (p) to the symbol lm/W might give rise to misleading statements.

Table I. Ratio of radiation equivalent (lm/w) in the scotopic region (k_s) to that in the photopic region (k_p), for various light sources [15]).

Monochromatic light		Perfect radiators	
λ(mμ)	k_s/k_p	T_F(°K)	k_s/k_p
450	21 (30)	2042	**1.00**
500	7 (7.7)	2360	1.2
550	1.3 (1.2)	2850	1.4
600	0.2 (0.13)	5500	2.15
650	0.07 (0.02)	7500	2.4
sodium	0.3 (0.23)	sunlight	2.0
mercury	1.1	daylight	2.3

[15]) The figures are calculated from Weaver's data. The figures in brackets in the first column are from the CIE data. The composition of the mercury light source is: λ 405 mμ (12%), λ 408 mμ (2%), λ 436 mμ (22%), λ 546 mμ (29%), λ 578 mμ (35%).

The above-mentioned expressions (3) and (4) have the same value only if $P'(\lambda)$ represents the spectral distribution of a source with $T_c = 2042$ °K (the calculation of K_s from K_p was based on this equality).

In Table I the ratio k_s/k_p is given for a number of light sources. It is found that for light sources with a continuous spectrum (incandescent lamps) the ratio is not greater than 2 or 2.5. It may reasonably be supposed that the same holds for fluorescent lamps. It is remarkable, however, that for monochromatic light, especially for wavelengths near the ends of the visible spectrum, the values of k_s and k_p may differ by a factor of more than 10.

Summary. In photometry the observer makes himself independent of the variance of brightness impression by always comparing two patches of light which produce about equal brightness impressions. In the region of high brightnesses an efficiency function of the human eye has been standardised (CIE 1924). The definition of luminance is based on this function. The unit of luminance is the cd/m² or "nit". At low luminances (< 3cd/m²) equality of luminance does not correspond to equality of brightness impression. Following Bouma, "subjective brightness" is defined as the luminance of the comparison field, illuminated by a special source, which in the photometer appears equally bright with the field to be measured. For this source, Bouma chose monochromatic light (λ = 535 mμ). Another often recommended choice is a source of colour temperature T_c = 2360 °K. In 1951 the CIE recommended a source with T_c = 2042 °K (black body at the freezing temperature of platimun, on which the definition of the candela is also based). At the same time, an efficiency curve for the region of lowest brightnesses (scotopic region) was standardised. It is now possible to express the "subjective brightness" in cd/m², and to use the units lumen, lux etc. also in this region. In order to avoid misunderstanding, however, it is advisable to mark these units with the suffix: scotopic or photopic. In scotopic photometry use can be made of other standard sources (e.g. T_c = 2360 °K, T_c = 7500 °K), but in this case the measured luminances must be multiplied by a correction factor in order to reduce them to the luminance of an equally bright field, illuminated by the candela source.

BOOK REVIEW

Artificial light and photography, by G. D. Rieck and L. H. Verbeek, pp. 347, 180 illustrations, 52 plates (4 in colour), 10 tables. — Philips Technical Library — This book can be ordered through your technical bookseller.

On opening this book, the first impression is one of mathematical formulae set out against a background of exquisite photographic reproductions. This is only superficially an incongruous combination: in fact, it serves to demonstrate the two-fold nature of photography, being an art and at the same time a science, or better, a technique based on scientific principles.

In the handling of this technique and its varied apparatus (e.g. cameras, developing methods, enlarging and projection equipment and their associated light sources) the serious photographer will want to know something of the fundamentals of the subject and the factors which influence his results. He will find the present book to be a rich source of information. The authors have succeeded in producing a comprehensive and often enthralling picture of the scientific background of their subject. The book is suitable both for the professional photographer with a considerable photographic background and for the intelligent and technically interested amateur. Some of the more specialized matters are dealt with in detail in small print and a number of literature references are given. On the other hand, readers who seek practical information can omit the theoretical parts, which are chiefly concentrated in chapters II and III (Light and lighting, and sensitive materials and their reaction to light, together about 100 pages), and enjoy chapters V, VI and VII (exposure using filament and gas-discharge lamps, flash exposures and the use of artificial light for projection, reproduction and processing — together 120 pages). Many practical hints, useful to both professional and amateur, are given. In between these sections of the book, Chapter IV (80 pages) gives an extensive description of the light sources themselves. Finally, Chapter VIII briefly discusses some special applications of artificial light, e.g. for infra-red and ultra-violet photography, microphotography etc.

The contents of the book are well balanced. For example, considerable attention (more than usual in books of this nature) has been given to the construction and characteristics of the incandescent lamp and the flashbulb, in view of their practical importance. Some sections of the book deserve separate mention, because of their original presentation of certain practical problems: the discussion on the combination of sunlight and flashlight (p. 239), the "shadowing" of objects (p. 37), and the colour rendering of monochrome and colour materials (p. 66) are especially noteworthy. As a handbook its value is greatly enhanced by the tables of practical data and by the two extensive indexes (authors and subjects).

Distributed througout the book are about fifty large reproductions (some in colour) of pictures taken with artificial light. Small-scale reproductions of these pictures are printed at the end of the book together with their technical data, in particular the lighting methods used. This "concise course in artistic photography" may well induce the serious amateur to aim at emulating these results in his own work.

S. GRADSTEIN.

VOL. 15 No. 7, pp. 189-220 JANUARY 1954

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
THE PHILIPS INDUSTRIES

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, NETHERLANDS

AN AUTOMATIC RECORDING POTENTIOMETER FOR INDUSTRIAL USE

by H. J. ROOSDORP.

621.317.733.083.4.078:621.317.39

In electrical equipment for the measurement and control of industrial processes, high sensitivity and accuracy must be combined with robust construction and reliability. The following article describes an instrument operating on the null principle and automatically balanced, which goes a long way towards meeting these requirements.

The great progress in electronics, which has made such a valuable contribution to the development of telecommunications, has also resulted in great improvements in the field of electrical measurement. Much thought has also been devoted to the problem of applying these perfected measuring techniques to the measurement of non-electrical quantities. In order that such measurements may be effected with the aid of electrical equipment, it is necessary to convert the quantity to be measured to a corresponding electrical quantity. Devices that will do this are known in general as transducers, and examples of these are given in the table.

From this table it will be seen that most of the electrical quantities produced by such transducers are D.C. voltage or resistance values. This is not altogether fortuitous, for these quantities can be easily and accurately measured: the transducer employed for any particular purpose is chosen to give the most suitable electrical quantity for measurement.

In this article an instrument will be described which is capable of measuring very accurately direct voltages, resistances or even impedances, and which, because of its robust construction, is suitable (in conjunction with some form of transducer) for the measurement, recording or automatic control of many non-electrical quantities in industry [1]).

Quantity to be measured	Transducer	Electrical quantity
Temperature	Thermocouple	D.C. voltage
	Resistance thermometer	Resistance
	Total radiation pyrometer	D.C. voltage or resistance
Pressure Force (weight)	Strain gauge	Resistance
Displacement	Differential transformer	Self inductance
Rate of flow of gas or liquid	Diaphragm with pressure difference pick up	Resistance or self inductance
Humidity	Dew-point pickup (resistance thermometer)	Resistance
	Wet and dry bulb, with resistance thermometer	Resistance
	Conductivity pickup	Resistance
PH	Glass-calomel electrode	D.C. voltage
Luminous intensity	Photo-electric cell	D.C. voltage
	Barrier-layer cell	Resistance
Ionizing radiations	Ionization chamber	Direct current
	Geiger counter	Direct current

[1]) A recording instrument somewhat similar to that described here was designed in the Philips Laboratories during the war.

Measuring methods

Voltages, currents and impedances can be measured in accordance with two different principles, viz. the deflection method and the null or comparison method. Some instruments using the first method are, e.g., electrostatic voltmeters for voltages, moving-coil and moving-iron instruments for currents, and the cross-field or cross-coil type of instrument for impedances.

The great advantage of the deflection system is that it gives a *direct* reading, which bears a known relationship to the quantity to be measured. It suffers from the drawback, however, that the measuring instrument itself determines the attainable accuracy, which may not always be adequate. Another disadvantage is that a certain amount of power is required to produce the needle deflection, so that, if the transducer is capable of delivering only a very small amount of power, an extremely sensitive instrument, or alternatively an amplifier, has to be used. In the latter instance, the measuring accuracy is also dependent on the stability of the amplification.

With the null method, the unknown quantity is compared with another (known) quantity of the same kind. This can be done by means of a bridge circuit or a potentiometer. Only a detecting instrument is then required and the accuracy of the measurement does not then depend on that of the instrument but only on the comparison elements (e.g. standard resistances). Also, since the instrument is employed only as a zero indicator, it can be made very much more sensitive than those of the direct-reading type. (Moreover, the sensitivity can in this case be increased by using an amplifier, the amplification of which need not be known, or constant).

The null method thus provides at once greater sensitivity and accuracy than the direct-reading system. The only disadvantage of the null method is that an adjustment has always to be made before the value of the unknown quantity can be obtained from the known comparison quantity.

Automatic balance

The comparison or null method is normally employed when precise measurements are required. The disadvantage of having to adjust the balance manually for each and every measurement can be eliminated by using a servo motor to effect the adjustment automatically [2]). With such an "automatic bridge" the quantity to be measured is thus read direct, so that here we have in effect a direct-reading instrument. Another great advantage of the servo motor is that it makes a much larger force available than the conventional direct-reading system. The null method with automatic balance therefore provides better facilities for connecting the measuring system to a recording mechanism to provide a continuous record.

The following numerical example will serve to illustrate this.

In a moving-coil meter the driving torque is:

$$M = B\,i\,l\,d\,w, \quad \ldots\ldots \quad (1)$$

where:

M = torque in N·m *),
B = magnetic induction in the air-gap, in Wb/m^2 (1 Wb/m^2 = 10^4 gauss),
i = current in Amps flowing in the coil,
l = length in m, of that part of the coil within the magnetic field,
d = diameter of coil in m,
w = number of turns.

If the resistance of the coil is, say, 400 Ω, the current i, for the measurement of a voltage of 10 mV (e.g. from a thermocouple), is 25 μA. Now, if we further assume that $B = 0.3$ Wb/m^2, $l = 3$ cm, $d = 2$ cm and $w = 1000$, the torque will be 4.5×10^{-4} N.cm = 46 mg.cm. This torque must counterbalance that of the spiral springs. When the current to be measured is varied, the difference between the torque applied to the moving coil by the magnetic field and that of the spiral springs is available for the variation in the deflection of the needle.

If the current varies, say 1%, a torque of 0.46 mg.cm is available for the deflection. The force applied to the end of a needle 10 cm in length is then 46 μg, which is just enough to initiate an indication, but quite inadequate for direct recording purposes.

The amount of force delivered can be increased by means of an amplifier, but the amplification that can be employed is limited by the maximum dissipation of the moving coil. Assuming this to be 1 W, and again taking the resistance of the coil to be 400 Ω, the maximum current that the coil will carry is 50 mA. The maximum possible torque is then 2000 times greater than in the example given above; a force equal to 92 mg will now be available at the extremity of a needle 10 cm long. This is just sufficient to drive an inking stylus, but the greatest care would have to be taken in the design and operation of such an instrument.

[2]) An early example of a measuring circuit embodying this principle was described in this Review by R. Vermeulen, Philips tech. Rev. 4, 354-363, 1939.

*) N = Newton, the unit of force in the M.K.S. system (= 10^5 dynes).

Using the null method with a servo-driven balancing mechanism as described below, a force of some hundreds of grammes is available. This force is supplied by an induction motor, to the control winding of which a power of about 4 W is applied by an amplifier. (To obtain so much force at the pointer of a moving-coil direct-reading instrument, 10^{10} W would have to be applied to give full-scale deflection!)

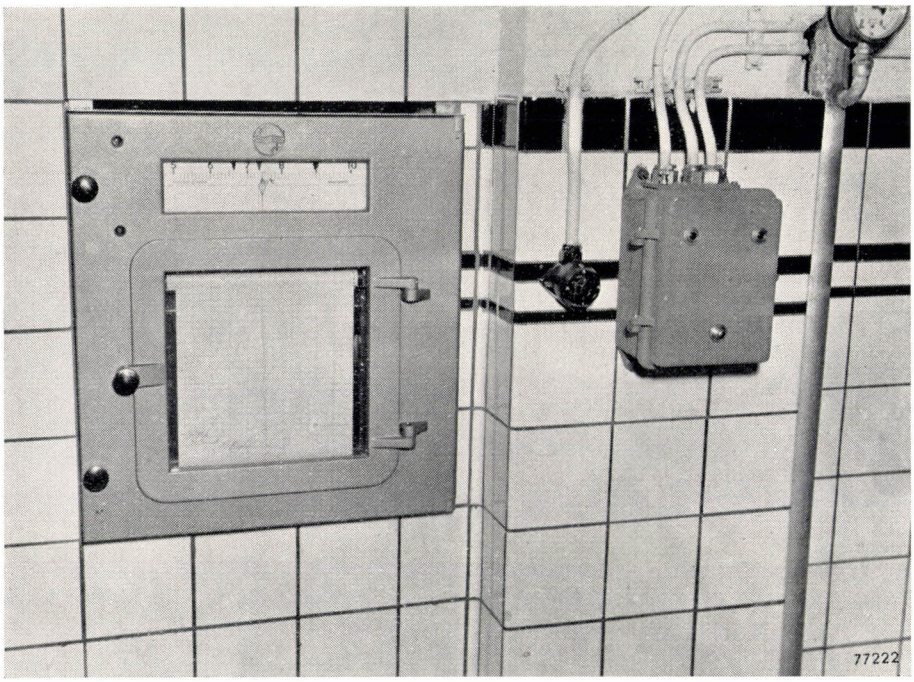


Fig. 1. The automatic measuring and recording bridge (type PR 2000) used for measuring and recording the temperature for the pasteurisation of milk.

A measuring circuit operated on the null principle and incorporating automatic balancing can be quite easily designed for a setting time very much shorter than that of a sensitive direct-reading instrument without an amplifier. Equipment can be made for industrial purposes in which the time necessary for traversing the whole scale is less than 1 sec.

For industrial applications, the null method with automatic balancing accordingly has many advantages, viz. high accuracy, considerable power and high speed. The equipment about to be described based on this system, will measure voltages and impedances and, if necessary, record them [3]).

Fig. 1 shows the unit employed for measuring and recording the temperature for the pasteurization of milk.

Principle of the automatic bridge

Figures 2, 3, 4 and 5 depict a number of circuits suitable for the measurement of direct voltages, and resistances, by the null method with automatic balance.

In *fig. 2* the voltage E_x from a thermocouple is compared with a voltage E_0 from a potentiometer connected to a source E_n whose voltage is precisely known. The difference E_i (the so-called error voltage) between E_x and E_0 is applied to an amplifier A, and the amplified voltage is used to drive a motor M which is coupled mechanically to the contact of the slidewire R_n. This contact is always moved by the motor in the direction that will reduce the difference between E_x and E_0. Provided that the amplification by A is sufficient, the motor operates

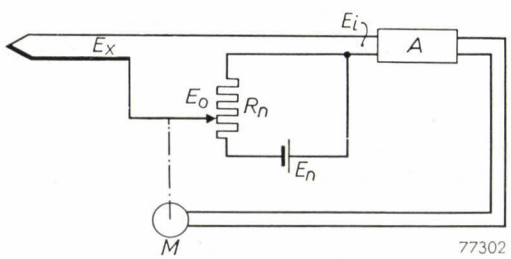


Fig. 2. Basic potentiometer circuit for measuring the voltage E_x of a thermocouple by the null method with automatic balance. E_n source of reference voltage. R_n potentiometer slidewire. A amplifier. M motor. The motor sets the contact of the slidewire to the point where the difference E_i between E_x and E_0 is practically zero.

[3]) Series PR 1000 for indicating, and series PR 2000 for recording instruments. The use of the automatic potentiometer/bridge for the control of industrial processes will be dealt with in a subsequent article.

at very low values of E_i, so that, practically speaking, when the motor stops, E_0 is equal to E_x. Accordingly, E_x can be read from a scale mounted on R_n.

In many cases it may be desirable to have the adjustment corresponding to zero voltage at a point other than the beginning of the scale; e.g. in measurements over a range of say from 50 to 150 V instead of 0—100 V, or where the range of measurement covers both positive and negative voltages (e.g. from —10 to + 100 mV). To achieve this, the potentiometer is made into a bridge circuit (*fig. 3*), which has the additional advantage that variable resistors (e.g. temperature-dependent resistors) can be included in one or more of the bridge arms, in order to render the circuit independent of certain disturbing effects. In this way, for example, compensation can be provided for variation in the e.m.f. of a thermocouple due to changes in temperature at the cold junction.

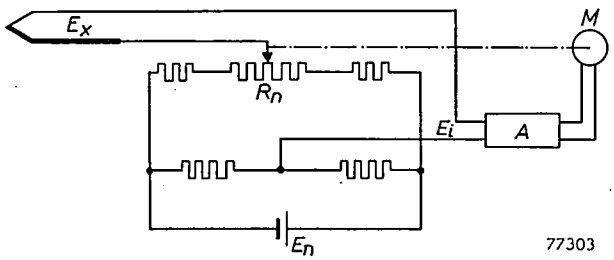


Fig. 3. Potentiometer circuit for measuring the voltage E_x of a thermocouple. With this arrangement, zero voltage may be made to correspond to any point along the length of the slidewire. The balancing voltage is obtained from a bridge circuit.

Fig. 4 shows a circuit suitable for the measurement of the resistance R_x of a resistance-thermometer. In the simplest form of circuit, this resistance is included in one of the arms of the bridge. Here again the sliding contact of the potentiometer R_n is moved by the motor M in such a way as to balance the bridge so that the error voltage E_i delivered to the amplifier A will be practically zero.

In *fig. 5* we illustrate a circuit that is widely employed in conjunction with strain gauges. These gauges, represented in the figure by the resistances R_s, are wired up in the form of a bridge, so as to balance out the effect of any resistance variations in the connections. Measurement is effected with the aid of a second bridge circuit in which the measuring slidewire, controlled by the motor M is included. Both bridges are fed from a single transformer; the motor sets the slider of R_n to the point where the voltage between points a and b is the same as that between c and d: at this point the adjustment is complete and the motor stops since the error voltage E_i applied to the amplifier is then effectively zero.

With sufficient amplification, the accuracy of measurement is determined by the tolerance within which the resistance values or the potential E_n

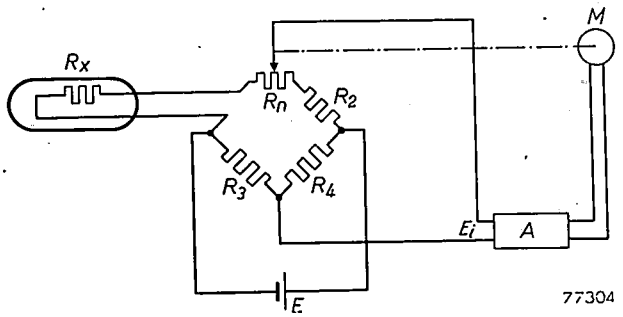


Fig. 4. Simple bridge circuit for measuring the resistance R_x of a resistance thermometer.

can be set. For industrial use this type of instrument is usually calibrated to an accuracy within $\frac{1}{4}$%.

The sensitivity, that is the lowest value of E_i at which the motor will revolve, is determined by the amplification from A. Very high amplification can be combined with robust construction in electronic equipment, to give an effective sensitivity of 1 μV.

The speed at which the adjustment is effected is governed by the design of the servo system. With a high-speed motor the time can be made very short, viz. less than 1 sec, but there is then some risk of the balancing mechanism maintaining an oscillation about the point of balance. To prevent this, the mechanism must be damped. The amount of damping required can be computed in the following manner.

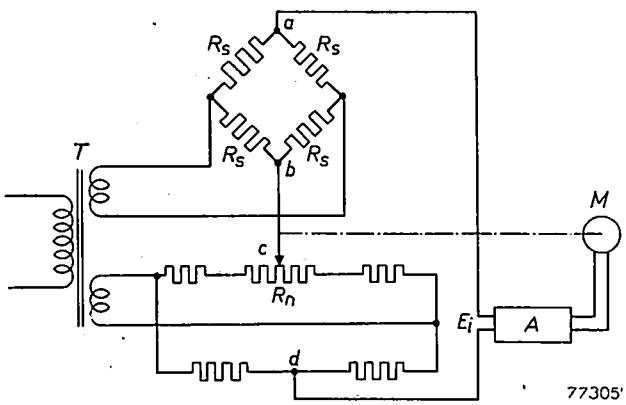


Fig. 5. Circuit for strain gauge measurements. The gauges, represented by the resistances R_s, are connected as a bridge circuit. The slidewire R_n is included in a second bridge circuit, and the two bridges are fed from a single transformer T. The motor sets the slider on R_n to the point where the difference E_i between the voltage across a-b and the voltage across c-d is practically zero.

In *fig. 6* the point on the slidewire R_n corresponding to zero voltage applied to the amplifier A, is denoted by O. Let us employ the following notation:

x = displacement of the slider from the point O;
vx = voltage applied to the amplifier A, corresponding to displacement x;
a = amplification provided by A;
$kavx$ = force applied by the motor M to the moving parts when a voltage avx is applied to it;
m = mass of the moving parts.

With no damping applied to the slider, the equation of motion is

$$m\frac{d^2x}{dt^2}+kavx=0. \quad . \quad . \quad . \quad . \quad (2)$$

The values of x satisfying this equation may be represented as a function of time by an undamped oscillation of angular frequency given by

$$\omega=\sqrt{\frac{kav}{m}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

We shall now suppose that a damping force is present which is proportional to the velocity of the

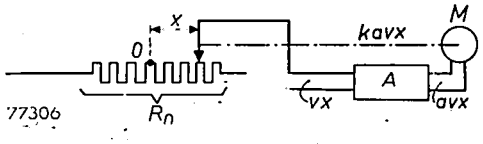


Fig. 6. When the contact of the slidewire R_n is located at a distance x from the point of balance O, a voltage vx is applied to the amplifier A. The amplifier then delivers a voltage avx which causes the motor to exert a force $kavx$ on the slider.

moving parts. For a velocity dx/dt, let this damping force be pdx/dt; the equation of motion is then:

$$m\frac{d^2x}{dt^2}+p\frac{dx}{dt}+kavx=0. \quad . \quad . \quad . \quad (4)$$

From this it follows that x as a function of time can be represented by a damped oscillation whose amplitude is proportional to $e^{-pt/2m}$. As with most measuring instruments, what we require is aperiodic motion, so that the value of p will have to be high enough to introduce critical damping. This occurs when:

$$p^2=4\,mkav. \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

Mechanically, it is extremely difficult to obtain a reproducible value of p. Electrically, however, this is quite a simple matter; in effect, the motor M is coupled to a tachometer generator G, which delivers a voltage that is proportional to the speed. For a speed of dx/dt, let this voltage be represented by gdx/dt. This voltage from the generator G, together with that from the potentiometer R_n, is supplied to the input of the amplifier (*fig. 7*). Without the mechanical damping, the equation of motion is now

$$m\frac{d^2x}{dt^2}+ka\left(vx+g\frac{dx}{dt}\right)=0. \quad . \quad . \quad (6)$$

The solution of this equation, giving x as a function of time, may be represented by a damped oscillation of amplitude proportional to $e^{-kagt/2m}$.

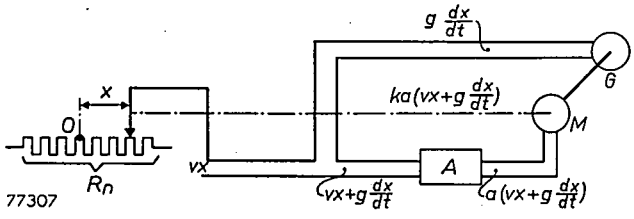


Fig. 7. A tachometer-generator G, coupled directly to the motor M, delivers a voltage that is proportional to the speed. This voltage is added to the input of the amplifier, producing an electrical damping of the balancing mechanism.

Critical damping occurs when

$$g^2=\frac{4\,mv}{ka}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

The damping can accordingly be adjusted to the required value by varying the output from the generator (i.e. by varying g).

To ensure high sensitivity, it is essential that the motor, even for the smallest value x_{min} of the displacement x, delivers a certain minimum power K_{min} to drive the recording mechanism, viz.

$$kavx_{min}=K_{min}, \quad . \quad . \quad . \quad . \quad . \quad (8)$$

and the required amplification is, therefore,

$$a=\frac{K_{min}}{kvx_{min}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

From (7) and (9) it follows that the value of g at which critical damping occurs is given by

$$g^2=\frac{4\,mv^2x_{min}}{K_{min}}. \quad . \quad . \quad . \quad . \quad (10)$$

We see from (9) and (10) that the desired values of a and g are dependent on v, that is, on the measuring range. If measurements are to be made in any other range, both g and a have to be re-adjusted, but this can be obviated by splitting the amplifier into

two parts, A_1 and A_2, in the manner shown in *fig. 8*. The equation of motion is now obtained by replacing v in (6) by va_1, and a by a_2. According to (8) the following equation must be satisfied if the recording mechanism is to operate at a small value x_{min} of the displacement x:

$$k a_1 a_2 v x_{min} = K_{min},$$

or

$$a_1 = \frac{K_{min}}{k a_2 v x_{min}} \quad \ldots\ldots \quad (11)$$

Critical damping occurs when

$$g^2 = \frac{4 m v a_1}{k a_2}, \quad \ldots\ldots \quad (12)$$

from which, together with (11),

$$g^2 = \frac{4 m K_{min}}{k^2 a_2^2 x_{min}}. \quad \ldots\ldots \quad (13)$$

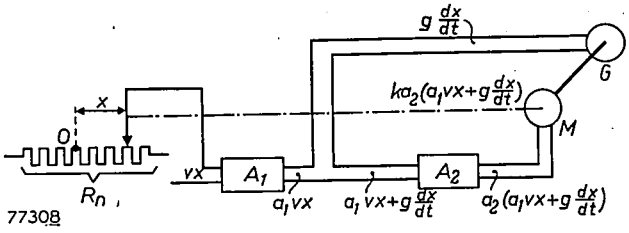


Fig. 8. With the amplifier divided in two sections, A_1 and A_2, it is not necessary, when the measuring range is changed, to modify both the voltage which the generator delivers at a given speed (factor g) and, at the same time, the amplification. It is sufficient to vary only the amplification a_1 of the amplifier section A_1.

This shows that g is now independent of v. When the measuring range is varied, g need not be readjusted; in accordance with (11) it is sufficient to adjust a_1.

Construction

The general arrangement of the automatic bridge isshown in the diagram in *fig. 9*. The resistors, which together with the slidewire R_n, form the bridge circuit proper, are grouped in the form of an interchangeable unit B corresponding to the required measuring range. The latter can then be changed simply by replacing the unit B. The slider of R_n is mounted on a cursor W which is moved along the rail C by the motor M through gearing and a friction drive D. A pointer indicates the position of the cursor on a scale S. The rotor of the tachometer generator G is coupled to the shaft of the motor M. The amplifiers and supply units, in the form of interchangeable units, are mounted in a rack E. Unit *1* supplies a voltage E_B to feed the measuring element. Unit *2* is the amplifier denoted by A_1 in fig. 8. The amplifier shown in that figure as A_2 is in two sections, *3* and *4*, of which *3* receives both the output voltage from *2* and the voltage E_G from the generator. Unit *4* is the output stage, of which the output voltage E_u is applied to the motor M. Lastly, the supply voltages for the two amplifiers are furnished by unit *5*.

Between *4* and *5*, other units can be placed in circuit when the equipment is to be used for automatic control, but these will not be discussed here [3]).

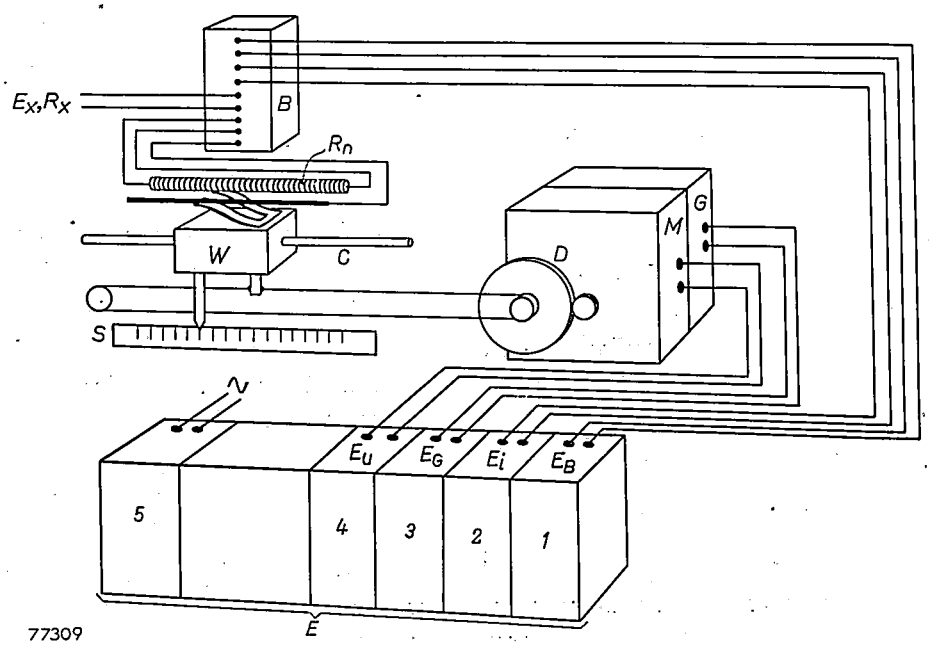


Fig. 9. Block diagram of automatic potentiometer bridge, types PR 1000 and PR 2000. The cursor W, which is moved along the rail C by the motor M through a friction drive, carries both the slidewire contact and the pointer, which moves along the scale S.

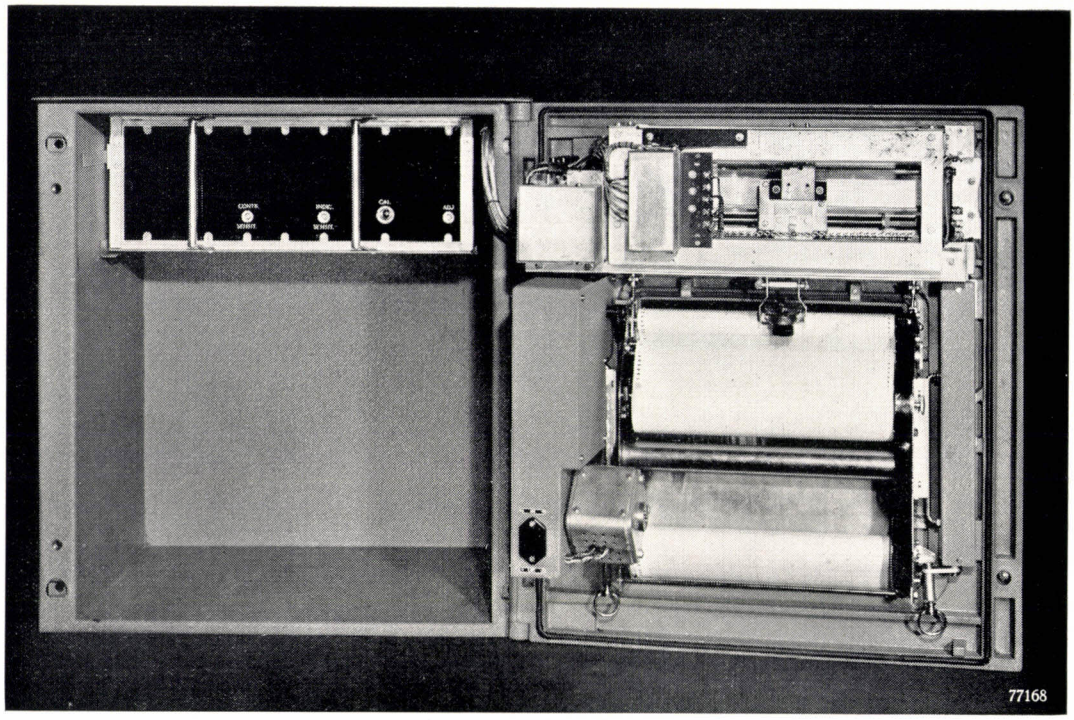


Fig. 10. Automatic recording potentiometer/bridge type PR 2000, showing recording mechanism. Top left: rack of amplifiers and supply units. Top right: in the opened cover of the instrument will be seen the slidewire, servo-motor and tachometer-generator, below which the recording mechanism is mounted.

In order to ensure the high degree of reliability demanded of equipment intended for industrial use, all the components are loaded far below their rated values; the use of components having only a limited life has been avoided.

Fig. 10 is a photograph of the instrument opened to show the interior. The front cover (at right) contains the slidewire, servo motor, tachometer-generator and recording apparatus, the electronic section (rack *E*, fig. 9) being mounted in the body of the unit (left). This rack is shown in detail in *fig. 11*, whilst *fig. 12* depicts one of the amplifiers removed from the rack.

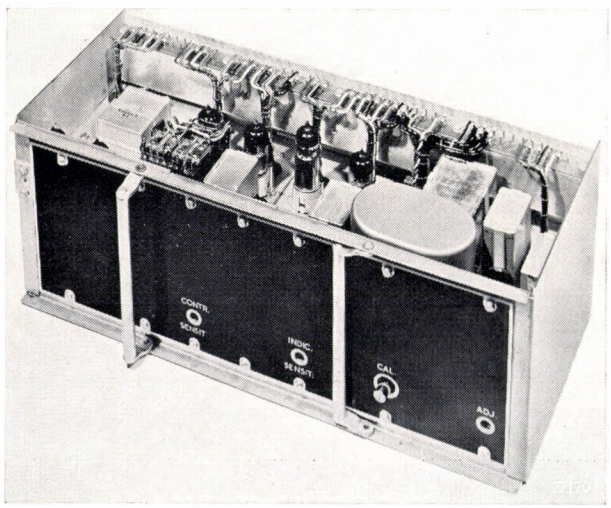


Fig. 11. Chassis with amplifier and supply units. All the leads are taken to a central connecting strip, to facilitate testing.

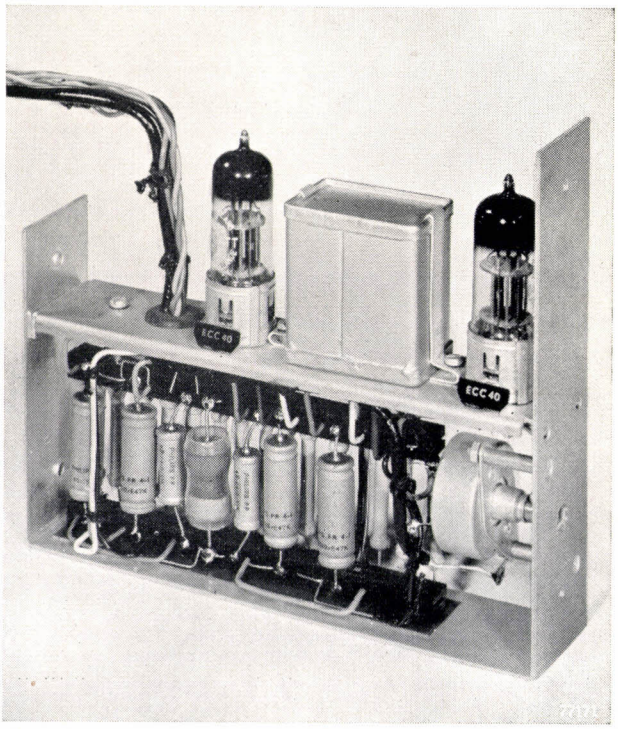


Fig. 12. One of the amplifier units. All the components are readily accessible.

To facilitate maintenance, all the main components, moreover, can be checked without any difficulty *in situ*, as all the principal leads are mounted on a central connecting strip on the rack (see fig. 11). The rack is connected to the other parts of the circuit through a line of plug sockets at the rear. All the components of each of the units are immediately accessible, once the chassis is removed from the rack (fig. 12).

The slidewire

As previously mentioned, the measuring accuracy is determined by the precision of location of the slidewire contact and the effective value of R_n, as well as the precision of the other resistors and the reference voltage. It is relatively easy to manufacture fixed resistors to a fine tolerance and with a high degree of stability, but with slidewires this is not such a simple matter. The total resistance of a slidewire can certainly be made quite constant, but it is equally necessary that its absolute value, as well as its fractional variation as a function of the position of the slider shall be carefully predetermined. (The maximum tolerance is 0.1 %.) In this way it is possible to change to a different measuring range without recalibrating the whole system, merely by changing the fixed resistors (unit B) and the scale. The slidewire R_n can then also be changed without necessitating recalibration of the scale. A fixed resistor is permanently connected across the slidewire, such that the equivalent resistance of the two is exactly 1000 Ω.

The cursor W is designed to carry the contacts of two slidewires simultaneously, thus making it possible to build up more complex circuits if required.

In order to secure a positional accuracy compatible with the measuring tolerance, the slidewire must have a high degree of linearity and must have a sufficient number of turns. In this instrument it is made with 1300 turns over a length of 260 mm, and has a linearity within ± 0.1 %. This means that the adjustment is accurate to about 0.1 mm.

Bridge supply voltage

For resistance measurements the bridge may be fed with alternating current at mains frequency. Sometimes, however, it may be necessary to employ direct current, for example, when the resistance to be measured is such a long way from the equipment that the capacitance of the leads would influence the result (these capacitances can, of course, be balanced out at the bridge, but it may be more convenient to avoid this complication by using direct current). When D.C. is used, the supply voltage is obtained from a rectifier with the necessary smoothing.

For D.C. voltage measurements, the bridge is used as a potentiometer. For this purpose the current through R_n must be of a constant and known value: it is therefore obtained from a source of voltage that is stabilized by means of a gas-discharge tube, whose operating voltage remains extremely constant throughout its life (type 85 A 1). The circuit is fed with a current of 1 mA (*fig. 13*) through a number of resistors, one of which (R_v) is variable, to enable the current to be adjusted to this value.

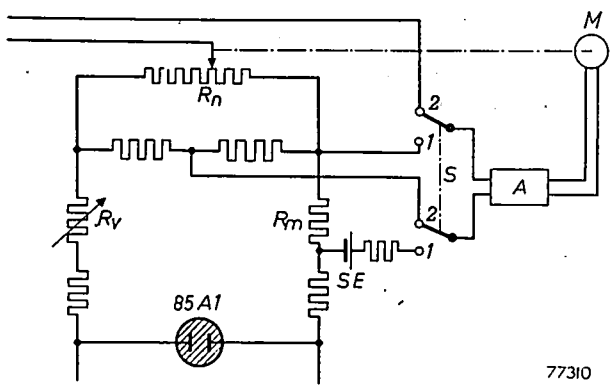


Fig. 13. Calibration of potentiometer. By means of the variable resistor R_v, the current for the potentiometer network is adjusted to 1 mA. This is verified by comparing the voltage across the resistor R_m (1018 Ω) with that of a Weston standard cell *SE*. For this operation the switch *S* is set to position *1*. With correctly adjusted current, the motor *M* does not operate.

Resistor R_m has a value of 1018 Ω, and in order to adjust the current to the exact value, the voltage across this resistor is compared with that of a Weston standard cell (1.018 V). This is effected by means of a switch *S* (fig. 13). When this switch is set to position *1*, the difference between the voltage across R_m and that of the standard cell *SE* is applied to the input of the amplifier *A*. If these voltages are equal, no voltage is applied to *A*, and the motor *M* is inoperative, but if a potential difference exists the motor revolves. This difference can be adjusted to zero by means of R_v, the indication of this being that the motor stops. The current is then at the correct value. Owing to the stability of the operating voltage of the tube 85 A 1, this calibration need not be made very frequently (e.g. once a month).

The amplifier

As electronic amplification can be effected most simply with A.C. voltages, the amplifier used is constructed as an A.C. amplifier. For the measurement of D.C. voltages, or for resistance measurements with the bridge operated on D.C., a converter is included between the measuring element circuit

and the amplifier. This converter consists of a contact vibrator, the energizing coil of which is connected to a source of A.C. voltage at mains frequency. For electrostatic measurements, a vibrating capacitor is employed. For both resistance and D.C. voltage measurements, therefore, the frequency of the voltages to be amplified is 50 c/s.

To ensure high sensitivity, the greatest possible amplification is required, but in practice this is limited by the inevitable noise and pick-up; there is of course no object in increasing the gain to the point where the amplified voltages of these disturbances would be sufficient to start the motor or saturate the amplifier. The spurious voltages occurring in the first valve correspond to a voltage of 5 μV on the grid of this valve. For this reason it is desirable to adjust the amplification to such a value that only when a signal voltage of about 10 μV is applied to the grid of the first valve, will the motor start to move. (This assumes that there is no pick-up superimposed on the signal voltage, which might saturate the amplifier). By using a transformer before the input valve, the minimum input voltage at which the motor will operate is reduced to about 1 μV. The input transformer also has the effect of reducing the input impedance of the amplifier (without the transformer, this would be about 1 MΩ; with transformer it is roughly 1 kΩ). As the motor positively rotates at 100 V, the voltage gain required in the amplifier in then about 10^8.

When a vibrating capacitor is used, the input impedance can be very high, say 10^{12} Ω. Now, owing to this high impedance across the cathode and grid of the first valve, the spurious voltage produced in the amplifier — designed, it should be remembered, for industrial use — is considerably greater than the value mentioned above. For this reason, the amplification is so adjusted that the motor operates only on an input voltage of at least 300 μV.

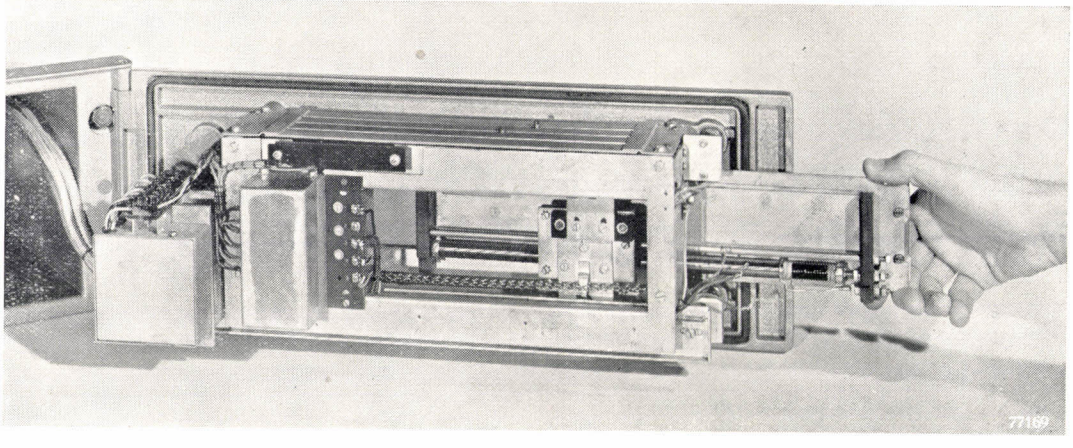


Fig. 14. The scale and slidewire are assembled as a single unit, as shown in this photograph of the non-recording version of the instrument.

The driving mechanism

For high accuracy it is essential that the position of the pointer in relation to the scale should agree exactly with that of the slidewire contact. Any backlash or flexibility in the transmission between slider and pointer would at once introduce errors. Such sources of error are avoided by mounting the pointer directly on the slider; in addition, the scale and the slidewire are rigidly mounted together (see *fig. 14*). The transmission between the motor and the cursor can thus in no way affect the accuracy of measurement.

The driving force must be at least high enough to overcome all sources of friction, and 300 grammes is adequate in this respect. To protect the cursor from being driven too forcibly against the end of of the scale by test signals beyond the actual measuring range, the potentiometer drive is of the slipping friction type.

The motor and generator

The motor is a two-phase induction motor with short-circuit rotor, as shown in diagram in *fig. 15*. One of the windings (*A*) is fed from the output amplifier, and the other (*B*), with a supply 90° different in phase, from the mains. When the point of balance of the measuring circuit is passed, the voltage delivered by the amplifier undergoes a phase shift of 180° and the motor reverses. To secure accuracy and stability in operation, a motor with a high starting torque is used.

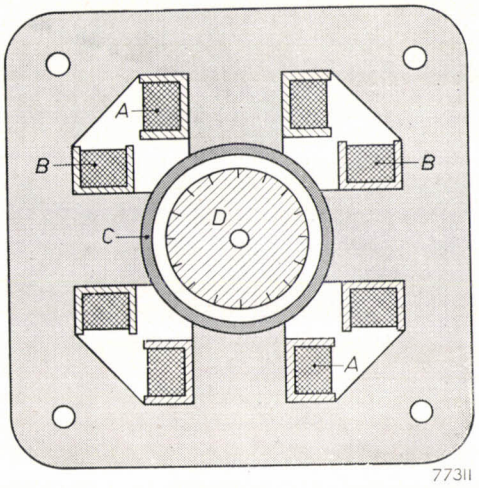


Fig. 15. Construction of the induction motor. Windings A are connected to the amplifier. Windings B are fed from the mains in such a way that the currents flowing in A and B differ in phase by 90°. An iron ring C distributes the magnetic fields set up by A and B so as to approximate to a rotating field. D is the short-circuit rotor.

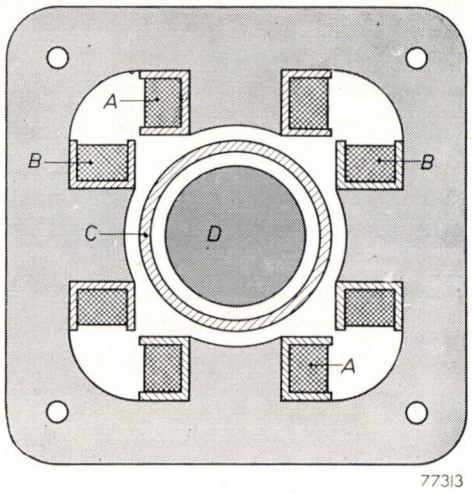


Fig. 16. Construction of the tachometer generator. Windings A are fed from the mains. When the aluminium rotor C is in motion, the magnetic field due to eddy currents in the rotor gives rise to a voltage in the windings B porportional to the speed of rotation. The output from the windings B is fed to the input of the amplifier section A_2 (see fig. 8).

Fig. 16 shows the principle of the tachometer generator. This also has two windings, one of which (A) is connected to the mains and the other (B) to the input of the amplifier section A_2 (see figs 8 and 9). The rotor consists of an aluminium cylinder (C) which revolves round a stationary iron core (D). When the rotor is brought into motion by the motor, a voltage is induced in winding B which is proportional to the speed of rotation. The frequency is the mains frequency and is therefore independent of the speed.

Reversal of the direction of rotation results in a phase shift of 180° in the voltage and, if the sign of the supply voltage has been correctly chosen, the phase relationship between the voltage delivered by the amplifier A_1 and that from the generator will be such that the latter provides a damping e.m.f. proportional to its speed. The ratio of the voltage delivered by the generator when stationary to that at full speed (approx. 2000 r.p.m.) is in the region of 1:1000. With the rotor stationary, the voltage induced in the winding B is therefore so small that it has a negligible effect on the result of measurement.

Summary. This article contains a description of an instrument for measuring D.C. voltages, resistances and impedances, of robust construction and high driving power. In conjunction with a suitable transducer, this instrument is suitable for the industrial measurement and recording of non-electrical quantities such as temperature, pressure, displacement, humidity, etc. The action of the instrument is based on the null principle which has many advantages over the direct-reading system, e.g. greater accuracy and higher sensitivity. The drawback of the null method, that an adjustment has to be made for each reading, is eliminated by using a servo-motor to effect the necessary adjustment automatically. The servo-motor, moreover, provides a considerable driving force; the setting time for full-scale deflection is about 1 sec. A tachometer generator coupled to the motor shaft provides a voltage which critically damps the motion, to prevent oscillation about the balance point (aperiodic damping).

AN APPARATUS FOR PROTECTION AGAINST DANGEROUS VOLTAGES IN WELDING EQUIPMENT

by F. H. de JONG and D. W. van RHEENEN. 621.791.735:621.316.722.3

In arc welding, the voltage requirements for good and rapid welding are, in general, in opposition to the demands of safety. A rather high no-load voltage is required which, in certain circumstances, may be dangerous to the welder. The apparatus described below gives complete protection to the welder, while in no way impairing the efficiency of the welding equipment.

Introduction

If alternating current is used for welding, this is usually furnished by a welding transformer. If direct current is used, a rectifier or a dynamo supplies the current. These diverse types of current source — here referred to under the general name of welding generators — have two electrical characteristics in common:

1) the output voltage drops considerably at increasing current;
2) the welding current at a voltage equal to the arc voltage can be adjusted between wide limits.

In *fig. 1* these typical properties are depicted for a welding transformer, by means of two characteristics, representing the output voltage V as a function of the welding current I. The one characteristic (*1*) refers to the adjustment using the highest welding current, the other (*2*) to the adjustment with the lowest welding current. Any characteristic between these two limiting curves may be obtained [1]). Other welding generators with similar voltage ranges, though with widely different current ranges, have similar characteristics.

The no-load voltage V_0 must be several times greater than the actual welding voltage (25-40 V), since otherwise the arc is difficult to start and will easily go out. On the other hand, it would be desirable to make V_0 as low as possible in order to minimize risk to the welder. The Netherlands Labour Inspectorate assumes that direct or alternating voltages up to 42 V will cause no fatal accidents in practice, but prohibits the use of higher voltages on exposed conductors. A welding generator with a no-load voltage of 42 V would, however, be of little use in practice for welding.

One is therefore compelled to work with no-load voltages with which danger to the welder is not entirely absent: as a rule, alternating voltages of about 80 V are used. Attempts have been made to reduce the danger by constructing the electrode

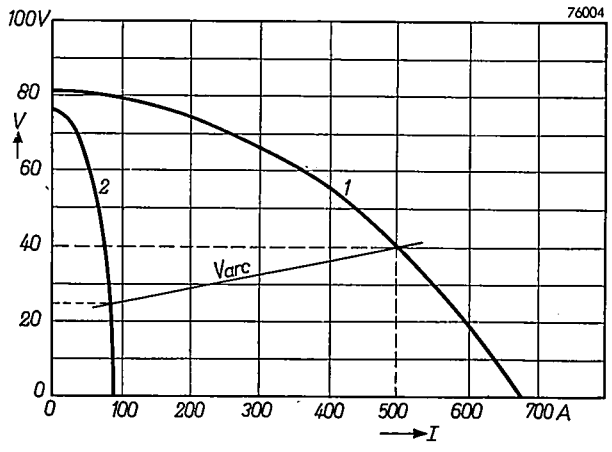


Fig. 1. Output voltage V as a function of the current I of an arc welding generator (here a Philips' welding transformer, type 3565). Curve *1* applies to the adjustment with the largest welding current, and curve *2* to that with the smallest. Continuous adjustment within these limits is possible. As the welding current is increased, the arc voltage rises from about 25 V up to about 40 V.

holder in such a way that involuntary contact with live parts is practically impossible (*fig. 2*), or by prescribing the use of insulating gloves for the welder (to which, however, there are some practical objections).

Such measures are by no means entirely satisfactory. Means have therefore been sought whereby the welding electrode and the electrode holder are isolated from dangerous voltages when the welding arc is not actually burning, in such a way that the welding operation itself is unaffected. The protective apparatus described below is distinguished from other devices on the market by the fact that it may be fixed to any welding generator whatsoever without causing any complications. It will presently be evident that in this appliance use is made of

[1]) The way in which these characteristics are obtained in the Philips welding transformers and rectifiers has been described by H. A. W. Klinkhamer, Philips tech. Rev. **1**, 338-345, 1936.

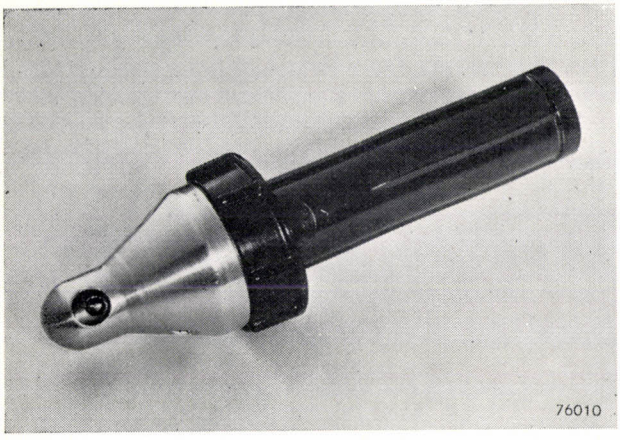


Fig. 2. Example of an electrode holder with insulated handle ("Safeweld" make).

the fact that a voltage of about 80 V is not dangerous unless one is exposed to it longer than a certain period of time. The Netherlands Labour Inspectorate has fixed this period at 0.2 sec., which includes a wide margin of safety.

Electric currents through the body are dangerous only when they cause so-called heart fibrilation. By this is meant a sudden and irregular contraction and relaxation of the heart muscles, which usually results in death. The mere feeling of a shock or irritation is not necessarily dangerous. Extensive investigations on this subject have been made by C. F. Dalziel of the University of California, Berkely. The following data are taken from one of his publications [2]).

Assuming the usual minimum value for the resistance of a current path through the body via the heart, viz. 500 ohms, it follows (from Dalziel's paper) that a voltage of 80 V, giving rise to a current of 160 mA, may be endured for one second without harmful results. The limit of only 0.2 sec, fixed by the Netherlands Labour Inspectorate, thus observes a wide margin of safety and at the same time allows some play in the actual value of the voltage.

Principle of the apparatus

The basic idea of this equipment is that, during the time when no welding takes place, the voltage V between welding electrode and the earthed piece of work is lowered to a harmless value by means of a voltage divider (R_1-R_2, *fig. 3*). As soon as welding is started, this voltage divider is automatically put out of circuit, so that then the full no-load voltage V_0 is available. The functioning will be clear from the simplified circuit diagram of fig. 3, and from *fig. 4*, which shows the voltage V and the position of various contacts as functions of time.

At the instant t_0, the welding generator (G, here a transformer), is connected with the mains via a switch (S_1). At first, V will then take the value $V_0R_2/(R_1 + R_2) = pV_0$; the ratio is so chosen that pV_0 is a harmless voltage, e.g. 40 V. At the same time, the coil of an electromagnetic switch (S_2) in series with the transformer primary is energized. After a specified, very short delay period, this switch makes contact at the moment t_1, which short-circuits R_1. V now rises to V_0. This does not yet entail any danger, however, provided that, within 0.2 seconds, V has again dropped sufficiently. This is effected by means of a relay (Re) and a delay device (T), which are joined in parallel across the welding arc. At the instant t_0 the current had already begun to flow through the coil of the relay via the delay device. After a certain interval τ,

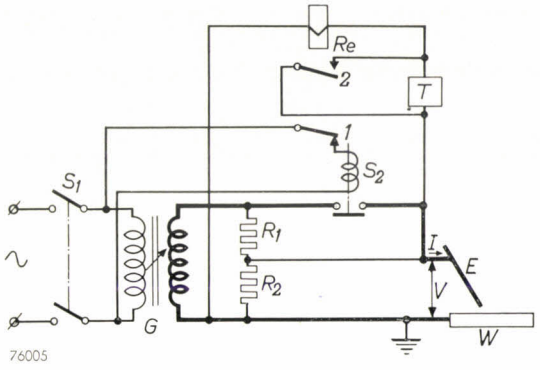


Fig. 3. Simplified diagram of the protective appliance. G welding generator, here represented as a transformer. W earthed piece of work. E welding electrode. R_1-R_2 voltage divider. S_1 mains switch. S_2 electro-magnetic switch. Re relay with break contact *1* and make contact *2*. T delay device.

at the instant t_2, the relay begins to move. In doing so it opens the contact (*1*), which is in series with the coil of the switch S_2. The latter therefore opens (at the instant t_2 if the delay of opening may be neglected), and V again decreases to pV_0. At this

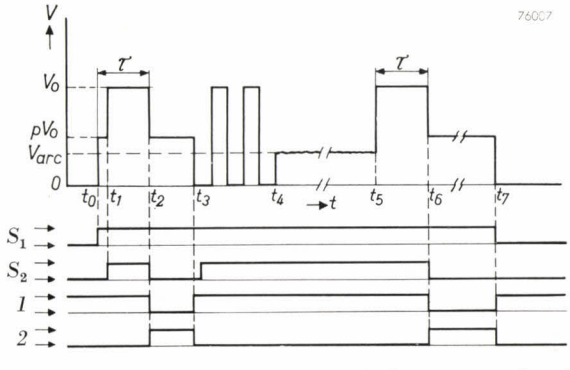


Fig. 4. Schematic representation of the r.m.s. value V of the voltage between welding rod and the work, as a function of the time t, in the circuit of fig. 3. V_0 = no-load voltage, pV_0 = voltage decreased by voltage divider. t_0 moment at which S_1 (fig. 3) closes. t_1 moment of S_2 closing contact. At t_2 after the delay time τ, relay Re is actuated. Between t_3 and t_4 an arc is drawn by tapping the welding rod on the work (arc voltage V_{arc}). At t_5 welding is finished, at t_6 the relay is released and at t_7 the mains switch is opened.

Below the graph, the positions of the switches S_1 and S_2, and of the relay contacts *1* and *2* are shown.

[2]) C. F. Dalziel, Dangerous electric currents, Trans. Am. Inst. El. Engrs. **56**, 579-585, 1946.

decreased voltage the relay remains attracted. In this position a second contact (*2*) short-circuits the delay device which is in series with the relay coil. The delay device is thus rendered inoperative, so that the relay can now work with practically zero delay. The maximum value of the delay time τ is fixed at 0.2 second, in accordance with the official safety requirements. The reason why τ may not be made indefinitely small will become apparent presently.

Welding may now be started. The welder brings the welding rod into contact with the piece of work (say at t_3), V consequently dropping to zero. As a result, the relay is released immediately and the contact *1* is again closed, so that the electromagnetic switch is energized and short-circuits R_1. The full short-circuit current of the generator can now flow through the welding electrode. On lifting the welding electrode, the arc will be started either immediately, or after few taps. From now on, V is equal to the arc voltage (25 to 40 V), i.e. not to be shorter than the 0.2 sec allowed for reasons of safety. The delay can therefore be chosen so as to meet both requirements.

As it happens, this delay has an important technical advantage in welding. If the momentary interruptions of the welding current were to result in the opening of the contacts of S_2, the voltage V would drop again to pV_0, and the welder would have some difficulty in starting the arc again, and faulty welds might result. The delay on the relay *Re*, however, prevents the protective device from reacting to such short interruptions, and the welder is therefore not hindered in any way.

It is assumed that at t_5 (fig. 4), the welder finishes his work. He lifts the electrode from the work, so that the welding arc is broken and the voltage rises to V_0. After the delay time τ the relay will be actuated again (at t_6), the contact *1* opens, the switch S_2 opens its contacts (no current at present through these contacts), and the voltage V drops to the safe value pV_0.

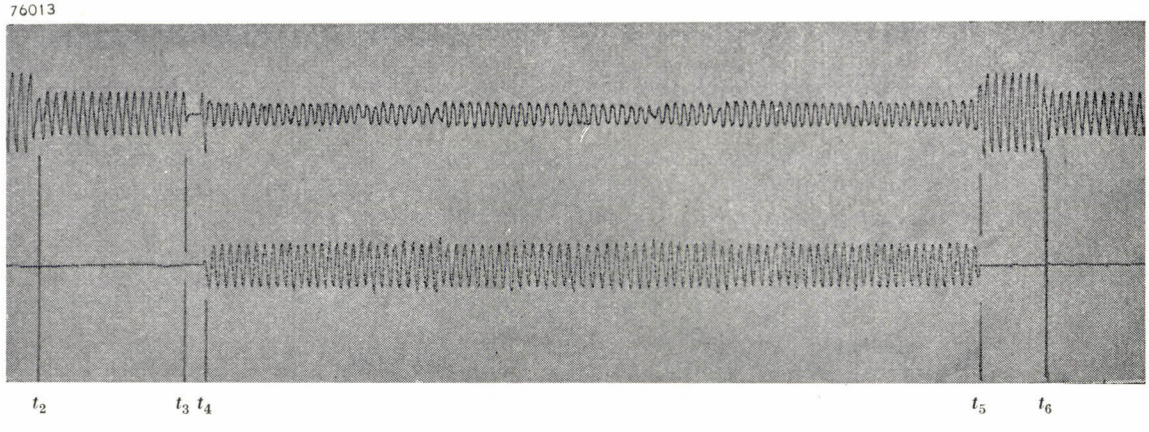


Fig. 5. Oscillograms of the voltage on the welding rod and the welding current (instantaneous values v and i). The instants $t_2 \ldots t_6$ correspond to those in fig. 4. The interval τ from t_5 till t_6, covers 8 cycles of the mains voltage (frequency 50 c/s), i.e. <0.2 sec. In about 1 cycle after tapping, the current is established.

large enough to operate the relay again. As welding proceeds, it sometimes occurs that the arc is broken momentarily. For a few cycles the welding current is then zero and the voltage is therefore equal to the no-load voltage. This should not cause the relay to open: if it did, the electromagnetic switch, because of its rapid response, would open perhaps just when the arc was struck again. The interruption of the full welding current, which may amount to hundreds of ampères, may seriously overload the contacts of the switch S_2. This is one of the reasons why a certain delay must be associated with the relay *Re*. The delay time required for this proves

The various operations just described can be found in the oscillograms of *fig. 5*, which show the instantaneous values of the welding current i and of the arc voltage v as functions of time. It can be seen that the interval τ, during which the full no-load voltage is present, lasts for only 8 cycles of the mains voltage (frequency 50 c/s), i.e. less than 0.2 sec. The oscillogram also shows that when the voltage becomes zero (welding rod in contact with the work, e.g. at the beginning of the welding period), the non-delayed relay quickly releases, and in 1 to 1.5 cycles (0.02 to 0.03 sec) the welding current will again begin to flow.

Construction

In fig. 3 the relay *Re* is for simplicity shown as an A.C. relay. A relay working on direct current, however, has certain advantages, especially if it is to be given a specific delay time. A direct current relay is therefore used in the present equipment. It is fed from a small full-wave rectifier (bridge circuit, connected as shown in *fig. 6*) which takes the place of the resistor R_2 (fig. 3).

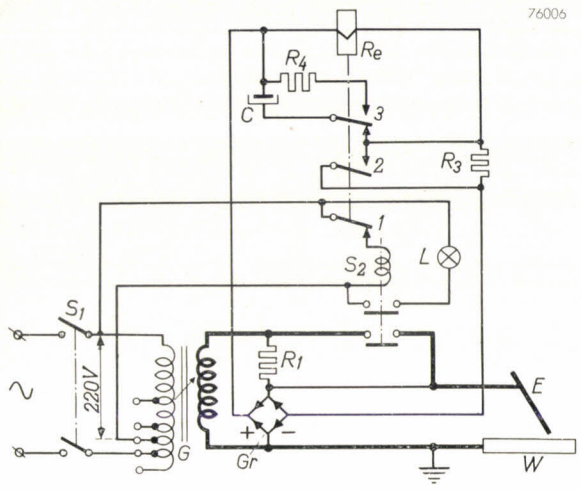


Fig. 6. Complete diagram of the protective device. *Gr* small selenium rectifier. R_3 resistor and *C* capacitor which together form the delay device (*T* in fig. 3). R_4 resistor, and *3* contact for discharging the capacitor. *L* pilot lamp. For the meaning of the other letters, see fig. 3.

A delay for the relay is obtained by means of a capacitor (*C*) in parallel with the relay coil, and a resistor (R_3) in series with this parallel connection. The delay time τ then depends on the values of *C* and R_3. The delay is removed by short-circuiting the resistor R_3 through the relay contact *2* and by simultaneously breaking the circuit containing the capacitor by means of a third relay contact (*3*). The same contact provides for the discharge of the capacitor through a resistor (R_4).

A further protective feature of this instrument is a pilot lamp which shows red as long as the electromagnetic switch (S_2) is closed, i.e. whenever a dangerous no-load voltage is present. The lamp will flash when the mains switch is first closed, and thereafter burns only while welding takes place. If some mechanical or electrical fault should cause the switch S_2 to remain closed when welding has finished, this would also be made known by the red light.

The interior of the apparatus (type PE 3100), showing some of the components, is shown in *fig. 7*.

Some further remarks should be made concerning the electromagnetic switch. This may be used for welding currents up to 500 A. For such a large current the switch is very small, due primarily to the fact that it breaks the circuit only when zero current is flowing, but also due to the manner in which the circuit is closed. A well-known objection to many electromagnetic switches is the rattle of the contacts which occurs on closing. This results in pitting and deterioration of the contacts. This is prevented on the switch used in this equipment, by causing the moving contacts (*4*, *4'*, fig. 7) to first touch the preliminary contacts (*5*, *5'*), which consist of flat limp springs so that no rattling occurs and the circuit is closed without interruptions. Only then do the moving contacts touch the main contacts (*6*, *6'*). Any rattling which now occurs is of no consequence, since the preliminary contacts remain closed. The latter, in the course of time, suffer some damage from the continual closing of the circuit, but they are easily renewed, as can be seen from fig. 7. The main contacts and

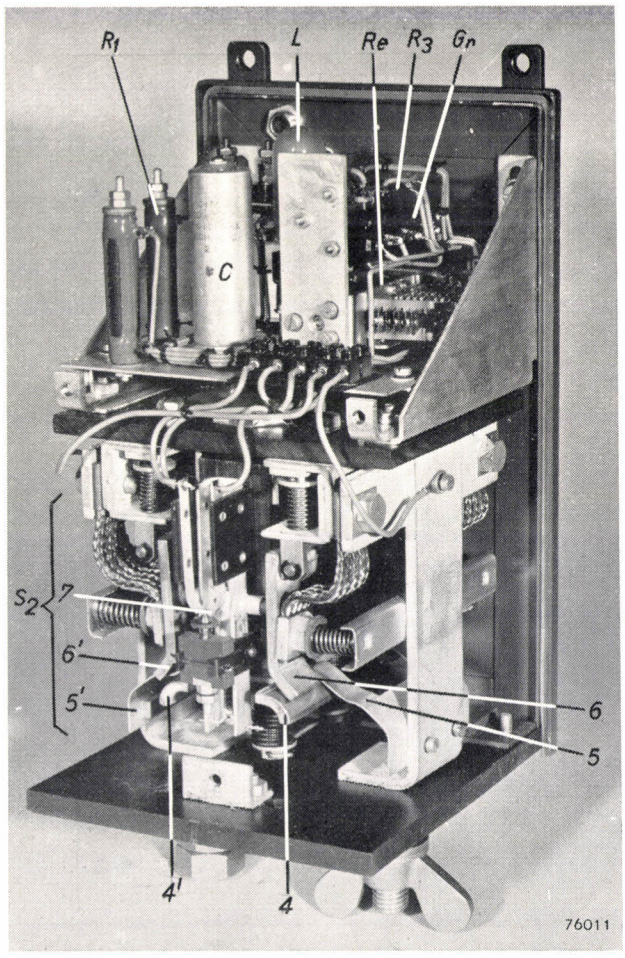


Fig. 7. The protective appliance type PE 3100 (cover removed). *4-4'* moving contacts of the electromagnetic switch S_2. *5* and *5'* preliminary contacts, *6* and *6'* main contacts, *7* auxiliary contact for the pilot lamp *L*. For the meaning of *Re*, *C*, R_1, R_3 and *Gr*, see fig. 3 and fig. 6.

the moving contacts are made of a chrome copper alloy that will stand a good deal of wear and tear. Switching tests showed that after 100 000 operations at 500 A, the latter contacts were still sound, and that the preliminary contacts showed some signs of burning, but not sufficient to merit their replacement.

Application

From the above description it will be clear that this protective appliance may be used with any type of welding generator without involving any major modifications to the latter. In general, with other protective devices, this is not the case; usually the no-load voltage is decreased by means of a voltage divider attached to the *primary* side of the welding transformer or rectifier, consisting of the primary coil of the transformer and a series choke coil, which is automatically short-circuited during welding. This type of voltage divider always has a rather low impedance. The protective appliance described here, with its high-impedance voltage divider attached to the secondary side, has a number of advantages, viz:

a) The large choke coil necessary for most protective appliances, with tappings for different mains voltages, is replaced in this apparatus by a small resistor (R_1, fig. 7).

b) Since the primary winding of the welding generator is not concerned in the protective device, no changes are necessary in the connections of power factor condensers.

c) Again, since the primary winding of the welding generator is not concerned, the device may be used for multiple welding with a common transformer, as often used in shipbuilding. The arrangement in this case is shown in the circuit of *fig. 8*; it can be seen that in this case a protective appliance connected to the primary side would not be feasible.

d) Owing to the high impedance of the voltage divider (R_1-R_2, fig. 3), even a somewhat deficient contact between the welding rod and the work will be sufficient to lower the voltage V such that the relay opens and the switch S_2 closes. With a voltage divider of low impedance, however, a good contact would be necessary: a rusty piece of metal would have to be previously cleaned in order to ensure that S_2 closed and rendered the voltage divider inoperative.

To connect the Philips protective appliance to a welding generator, only the following connections are necessary:

1) Connecting the contacts of the electromagnetic switch S_2, in series with the welding electrode. This involves only external connections, as can be seen from *fig. 9*, which shows the appliance mounted on a welding transformer.

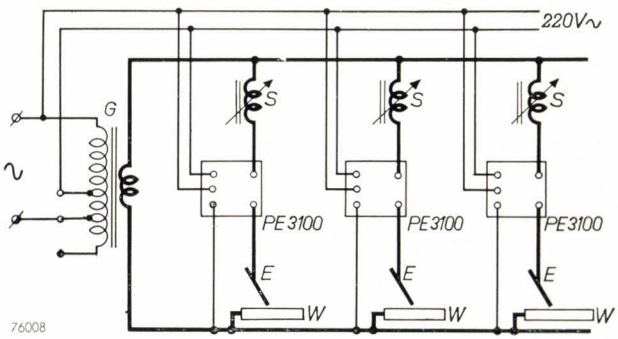


Fig. 8. Connection for multiple welding with common transformer (G). The dropping of the V-I curve (of fig. 1) is in this case only to a slight extent caused by the transformer; it is mainly obtained by means of choke coils (S), of which one has been fitted to each welding circuit (work W, welding rod E). The self-induction of the choke coils is adjustable, for regulating the welding current. The protective applicane PE 3100 is suitable for such an installation, as shown in the diagram.

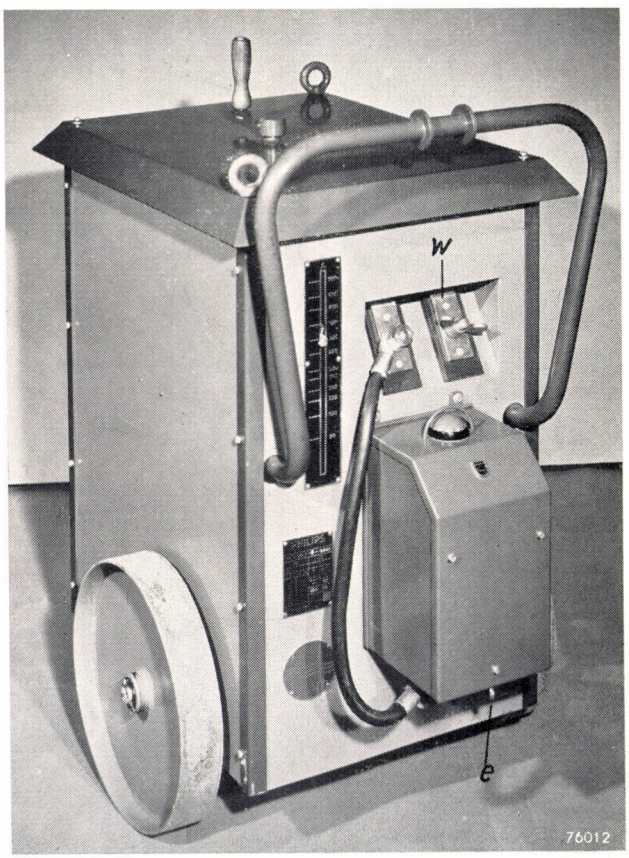


Fig. 9. The protective appliance PE 3100 mounted on a welding transformer. The work has to be connected to the terminal w, the welding rod to the terminal e.

The welding current may be adjusted by means of the handle on top of the transformer and can be read on the scale by the side of the terminals.

2) Making a connection with the earthy generator terminal (w), with which the work is connected (*fig. 10*).
3) Connecting the unit with a supply of 220 V alternating voltage. Even if the mains voltage has another value, the required voltage can usually be obtained from tappings on the primary side of the welding transformer, the rectifier transformer, or the motor driving the dynamo.

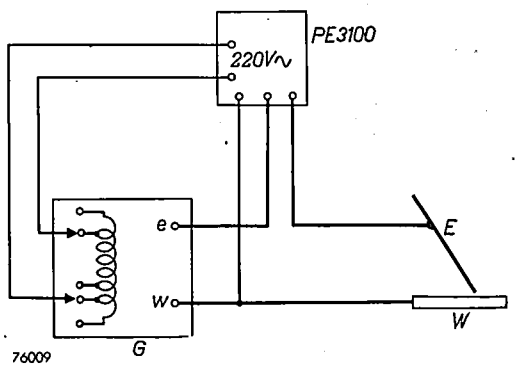


Fig. 10. Connection of the protective appliance PE 3100 to a welding generator. The only connections necessary are:
1. Two leads, one from the live transformer terminal to the apparatus and one from the apparatus to the welding rod.
2. A connection from the earthy terminal (w) of the transformer, with which the work (W) is connected.
3. Supply voltage of 220 V A.C. (from the mains side of the welding generator G).

Summary. The no-load voltage of most welding generators has a value of about 80 V. In certain circumstances this may be dangerous to the welder, but it cannot be decreased without adversely affecting the welding efficiency.

A protective appliance is described which eliminates the danger without affecting the welding efficiency. It is based on the principle of a voltage divider, which lowers the voltage between welding rod and the earthed work to a non-dangerous value while no welding is taking place, but which cuts out automatically as soon as welding is started, so that the full no-load voltage becomes available. This operation is done by an electromagnetic switch fitted with a delay time (0.2 sec, in accordance with the requirement of the Netherlands Labour Inspectorate). The delay time ensures that, while the switch will always cut out when no current is flowing, the welder will not be hindered by short interruptions of the welding current. The voltage divider itself is connected to the secondary winding of the welding generator and has a high impedance. This has a number of advantages against the more usual primary-connected voltage divider, which necessarily has a rather low impedance. Thus the appliance is easily mounted on any welding generator (transformer, rectifier, dynamo), without any changes in the connection of power factor condensors. The high impedance of the voltage divider ensures that the appliance works well even when the contact between the welding electrode and the work is not particularly good, whilst its connection on the secondary side permits the appliance to be used for "multiple welding". The appliance described (PE 3100) is suitable for welding currents up to 500 A.

A SIGNAL GENERATOR FOR THE TESTING OF TELEVISION RECEIVERS

by B. W. van INGEN SCHENAU. 621.396.615.17:621.397.62.08

The increasing number of TV receivers has made it necessary to develop methods for the rapid rectification of faults. Particularly useful for this purpose would be an easily portable generator, capable of producing all the signals necessary for this form of trouble-shooting; with the aid of such a generator tests could be made outside the normal transmission period of TV transmitters.

The generator described below produces the various signals required for the rapid tracing of a large number of faults occurring in receivers, both in the video and audio channels.

The various television standards

Considered as a whole, a television receiver is a fairly complicated piece of apparatus, and it is therefore hardly surprising that faults develop from time to time. A radio mechanic, called in to deal with a fault, will in most cases be able to do little as long as no normal TV signal is available. Since in most countries the TV transmitters are on the air for only a few hours per day, there is a great demand for a generator capable of supplying all the desired signals. Such a generator should be portable to facilitate use on the spot. The generator about to be described *), type GM 2887 (*fig. 1*), satisfies this requirement: the weight is approximately 7 kg ($15^1/_2$ lb), and the dimensions are $335 \times 220 \times 165$ mm ($13'' \times 8^1/_2'' \times 6^1/_2''$).

*) The development of this generator was initiated by Mr. H. C. Wardenaar.

The signals produced should obviously conform to the TV standard of the country in which the generator is to be used. Different standards prevail in the various countries. *Table I* shows the main characteristics for so far as our purpose is concerned. It comprises (a) the "International" standard, suggested some time ago by the Comité Consultatif International des Radiocommunications (C.C.I.R.) at Geneva, and adopted by a large number of countries [1]), (b) the British, (c) the American [2]) and (d) the French standard.

The differences between these standards are such that a single version of this signal generator would

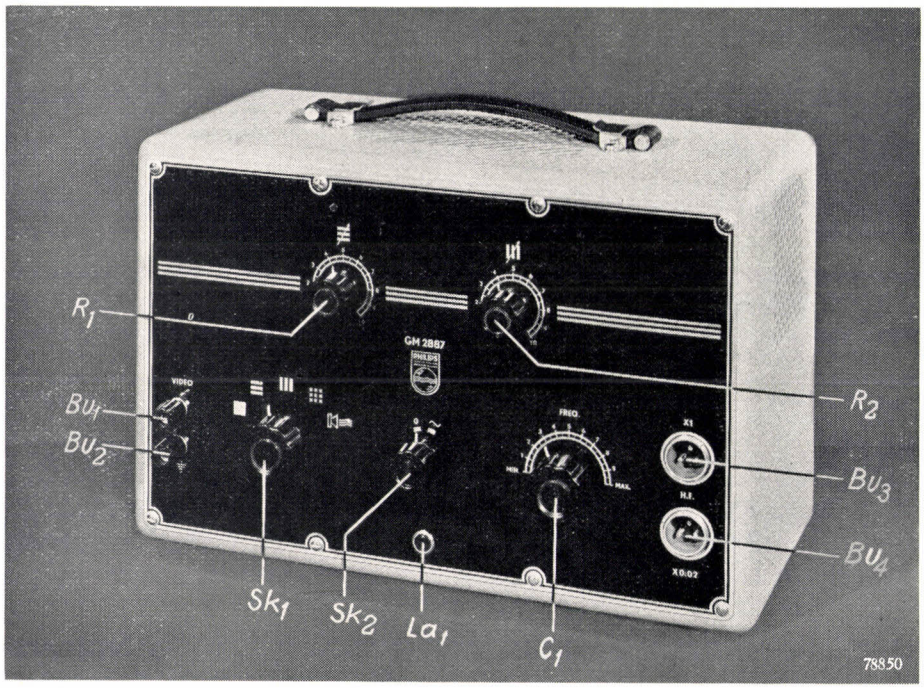


Fig. 1. Signal generator type GM 2887. Various video signals and an audio signal can be selected with the switch Sk_1 and the knobs R_1 and R_2. The carrier-wave frequency can be varied with knob C_1. The video signal appears at terminals Bu_1-Bu_2, the modulated H.F. signal at sockets Bu_3 and Bu_4. Sk_2 mains switch, La_1 pilot lamp.

[1]) These countries are: Argentina, Austria, Denmark, Finland, Italy, the Netherlands, Norway, Spain, Sweden, Switzerland, Turkey, West-Germany, and Yugoslavia.

[2]) In use in the U.S., Canada and various countries of Latin America. The American as well as the International standard are used in Brazil.

not be suitable for all four systems. For this reason the generator GM 2887 is produced in four models, designated by the letters A, B, C and D. The main data are given in *table II*.

Table I. Data on the television standards of various countries. f_s = frequency of the sound carrier-wave. f_v = frequency of the video carrier-wave. f_r = frame frequency. T_l = time interval between the beginning of one line and the beginning of the next. A.M. = amplitude modulation. F.M. = frequency modulation.

	International standard [1] (C.C.I.R.)	British standard	American standard	French standard
Picture modulation *)	negative	positive	negative	positive
Low. freq. band (Mc/s)	41- 68	41-68	49- 88	—
High freq. band (Mc/s)	174-223	—	174-216	174-216
f_s-f_v (Mc/s)	+5.5	−3.5	+4.5	−11.15
f_r (c/s)	50	50	60	50
Duration of frame-sync. pulse	$3T_l$	$4T_l$	$3T_l$	$0{,}4T_l$
Duration of frame blanking	$(19\text{-}31)T_l$	$14T_l$	$(13\text{-}20)T_l$	$41T_l$
Number of lines	625	405	525	819
Duration of line-sync. pulse	$0.09T_l$	$0.10T_l$	$0.08T_l$	$0.05T_l$
Duration of line blanking	$0.18T_l$	$0.15T_l$	$0.16T_l$	$0.16T_l$
Sound modulation.	F.M.	A.M.	F.M.	A.M.

*) See later (fig. 3).

Table II. Main characteristics of the four models of the GM 2887 signal generator.

Model	Number of lines	Line freq. c/s	Frequency band Mc/s	Picture mod.	Sound modulation	Standard
A	625	15 625	40- 80	neg.	F.M.	International
	525	15 750				American
B	405	10 125	40- 80	pos.	A.M.	British
C	625	15 625	170-230	neg.	F.M.	International
	525	15 750				American
D	819	20 475	170-230	pos.	A.M.	French

As will be shown later, with the aid of the generator GM 2887, a television receiver can be tested for the following:

1) synchronization,
2) linearity of the horizontal and vertical deflection,
3) correct connection of the deflection coils,
4) step-function response of the video amplifier,
5) sensitivity of the receiver (very roughly),
6) functioning of the video-frequency stages,
7) sound channel.

Signals produced by the generator and their use

The generator contains an *LC* oscillator, the frequency of which can be continuously varied (either between 40 and 80 Mc/s, or between 170 and 230 Mc/s) by means of knob C_1 (fig. 1), which operates a variable capacitor. The knob is not provided with a frequency scale, as this would have raised the price considerably, and in nearly all cases it is sufficient to test each channel to which the receiver can be tuned, without the necessity of knowing whether the frequency limits of these channels are correct.

The oscillator produces a voltage of $2\times(25\text{-}50)$ mV, symmetrically disposed with respect to earth [3]). The full voltage appears across the output socket Bu_3, whilst a value attenuated by a factor of approximately 1/50, appears at the socket Bu_4. The signal is conveyed from either of these sockets to the input of the receiver by means of a screened cable, supplied with the generator.

The cable has a characteristic impedance of 80 ohms, a value well adapted to many receivers. There are, however, also receivers with an input impedance of 300 ohms; for these the small box at the end of the cable (*fig. 2*) must be provided with a network giving the exact matching.

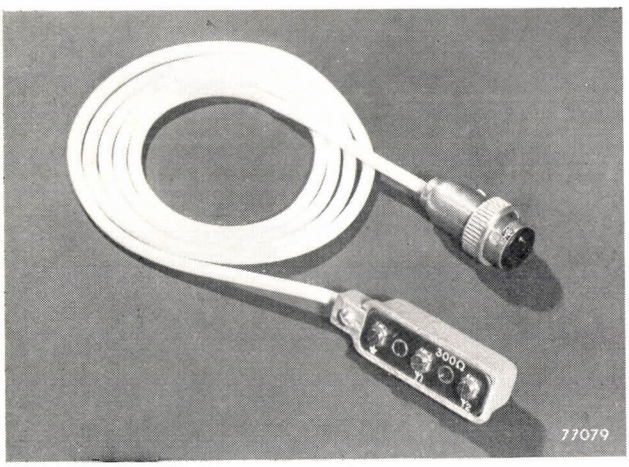


Fig. 2. Cable for connecting the generator to the TV receiver. The terminal box in the foreground contains a matching network for receivers having an input impedance of 300 ohms.

Five different signals can be modulated on the oscillator voltage; a choice of these is given by the switch Sk_1 (fig. 1). In the positions *1*, *2*, *3* and *4* of this switch — the video positions — the oscillator voltage is modulated with synchronizing signals and certain picture signals. In position *5* — the audio position — it is modulated either in frequency (models A and C) or in amplitude (models B and D) with an audio-frequency voltage, for the testing of the audio sections of the receiver.

With models A and C the video modulation is negative, whilst in models B and D it is positive.

[3]) This is not the case with model D, since the majority of French receivers have an asymmetrical input.

"Positive" and "negative" modulation are explained in *fig. 3*. The following discussion of the various output signals of the generator will be restricted to negative modulation.

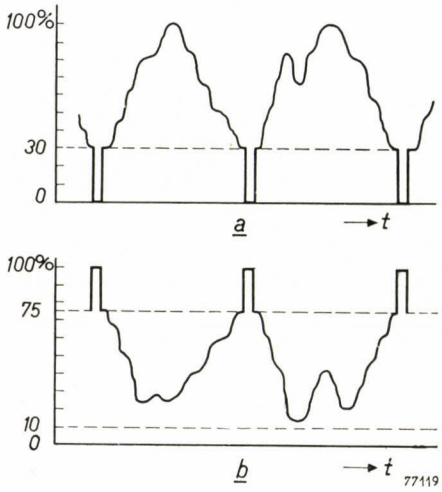


Fig. 3. *a*) Positive, *b*) negative modulation. The white level, the black level and the synchronization peaks lie respectively at 100%, 30% and 0% of the signal strength, with positive modulation, and at 10%, 75% and 100% with negative modulation.

1) Synchronization

With the switch Sk_1 at position *1*, the output voltage at Bu_3 has the shape shown in *fig. 4a*. The square peaks are the line synchronization signals. On either side of this, the signal is for a moment at the black level: the prescribed "front porch" and "back porch". These "porches" and the line synchronization pulse together form the line blanking signal (fig. 4*b*); this suppresses the beam during flyback. In the case shown in fig. 4*a*, the high frequency voltage present between the line synchronization signals has a small amplitude which corresponds to "white".

A frame synchronization signal with a frequency equal to the mains frequency is also given by the generator (not shown in fig. 4*a*); this signal too is accompanied by a front and back porch, so that vertical flyback is also suppressed.

With the signal shown in fig. 4*a* fed to the aerial terminals of a TV receiver, the horizontal and the vertical deflection should be synchronized with the generator so that the flyback becomes invisible, and the screen appears uniformly white.

2) Linearity of the deflections

If the receiver synchronization is satisfactory, the switch Sk_1 is set to position 2. In this position the generator gives a signal (*fig. 5a*) such that a series of white lines are traced but suppressed at regular intervals, so that horizontal white bars appear on the screen (*fig. 6a*). If these bars are not of equal width, then the vertical deflection is not linear. The majority of sets have a trimmer with which the linearity can be adjusted.

In position *3* of switch Sk_1 the output signal has the shape shown in fig. 5*b*. This gives rise to lines which are blacked out at regular intervals, forming a series of vertical bars on the screen (fig. 6*b*). Any differences in width between these bars show that the horizontal deflection is not linear and needs adjustment.

The horizontal and the vertical bars are combined when the switch Sk_1 is put in position *4*. Fig. 5*c* shows the shape of the signal, and fig. 6*c* is a photograph of the pattern which should appear on the screen.

3) Correct connection of the deflection coils

When a repair is made, it sometimes happens that the connections of the deflection coils are inadvertently reversed. This would give a reversed picture (top and bottom reversed, or left and right, or both). Owing to the symmetry of the bars, such reversal cannot be detected. A means of detecting this fault is given by the fact that the *number* of bars is continuously variable (from 6 to 10) by means of the control knobs R_1 (horizontal bars) and R_2 (vertical bars) (fig. 1). If the number of horizontal bars is increased, then the additional bars should appear on the screen from the bottom upwards; when the number of vertical bars is increased, the bars should appear from the right-hand side.

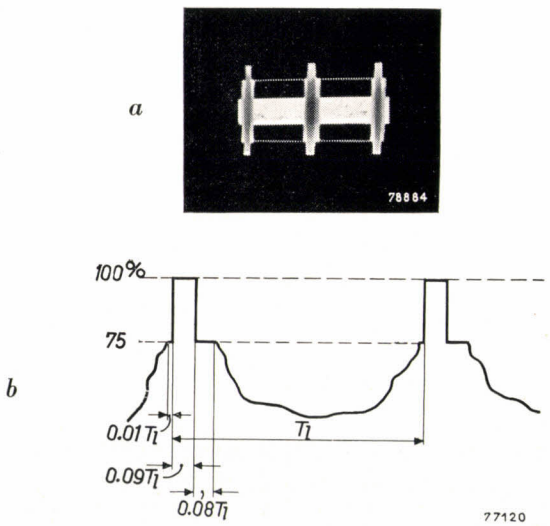


Fig. 4. *a*) Oscillogram of the high frequency output voltage. The line blanking signals can be seen, with the picture signal "white" (negative modulation) in between. *b*) Line synchronization pulses, with front porch on the left and back porch on the right (both at black level), and an arbitrary picture signal (negative modulation). The specified times apply to the International standard.

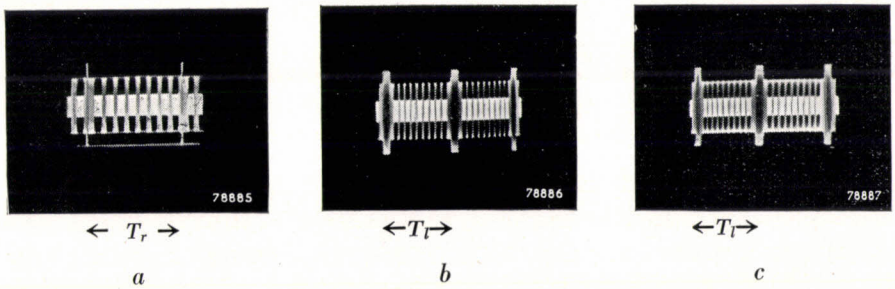


Fig. 5. Oscillograms of the high frequency output voltage (all with negative modulation). *a*) Horizontal bars, *b*) vertical bars, *c*) horizontal and vertical bars. T_r frame period, T_l line period.

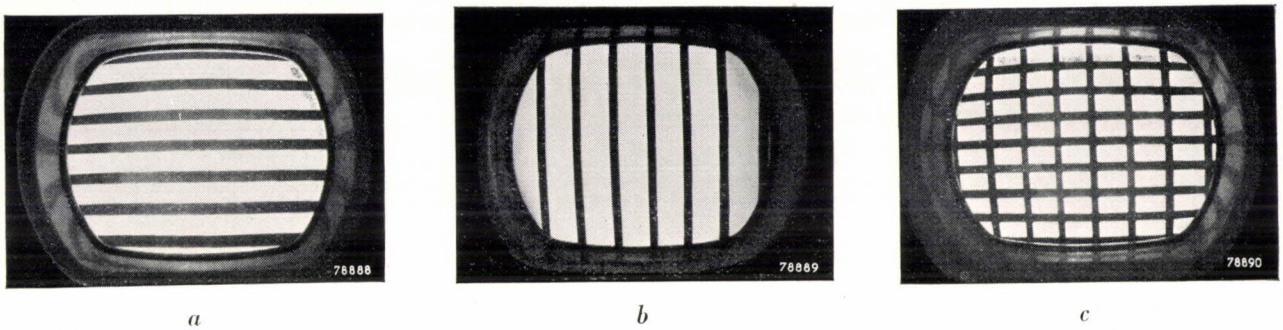


Fig. 6. Images on the picture tube, corresponding to the signals of fig. 5. *a*) Horizontal bars, *b*) vertical bars, *c*) horizontal and vertical bars.

Additional bars appearing from above or from the left are an indication that the vertical or horizontal deflecting coils have been incorrectly connected.

4) Step function response

The bars should have a uniform brightness over their entire width but deviations from this can occur, an example of which is given in *fig. 7a*. Here the receiver distorts the square pulses of the video signal into "sagging" pulses (fig. 7*b*) as a result of attenuation of the low frequency components. The cause of this may be a disconnected coupling capacitor in the receiver, so that only the very small capacity between the ends of the two wires is operative.

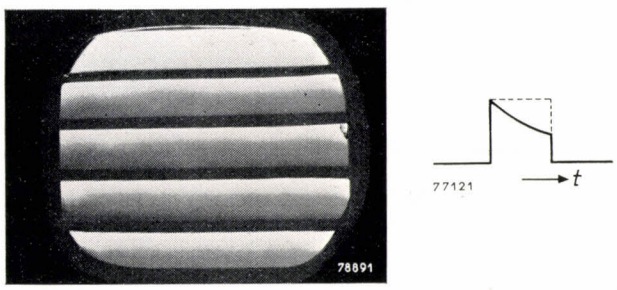


Fig. 7. *a*) Horizontal bars with non-uniform brightness caused by a square pulse (dotted in *b*) in the receiver being distorted to a "sagging" pulse (full line).

Even in the absence of a fault in the receiver, a brightness variation can exist in the vertical bars, viz., if the signal generator is badly tuned to the receiver. The frequency of the generator or of the receiver should therefore be adjusted so that the brightness is as uniform as possible before any conclusions may be drawn with regard to faults in the receiver.

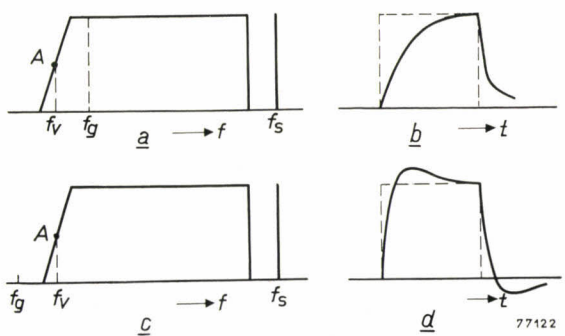


Fig. 8. *a*) Frequency characteristic of one channel of a TV receiver conforming to the CCIR (International standard). The video carrier-wave frequency f_v must be equal to the frequency corresponding to the centre A of the left flank. f_s audio carrier-wave frequency. The carrier-wave frequency f_g of the generator is here (*a*) shown higher than f_v, with the result that a square pulse in the receiver is distorted as in (*b*). *c*) As *a*), but with $f_g < f_v$, which results in a square pulse being distorted as in (*d*).

Fig. 8 shows why incorrect tuning gives rise to non-uniform brightness in the vertical bars. Fig. 8*a* shows an idealized frequency characteristic of one channel of a receiver conforming to the International standard. The video carrier-wave frequency f_v should be equal to the frequency corresponding to the centre A of the left flank of the curve; the low frequency

video components sent out by the transmitter in two side-bands then add to produce a signal as strong as that of the higher frequency video components, transmitted only in the upper side-band. If now the carrier frequency f_g of the signal-generator lies somewhat above f_t, then there will be a surplus of the lower frequency video components in the receiver, with the result that a square pulse will be distorted as drawn in fig. 8*b*, and the brightness will gradually increase from left to right. If f_g should lie somewhat below f_t (fig. 8*c*), then there will be a deficiency of the lower frequency components, and the square pulse will be distorted by "overshooting" (fig. 8*d*), with a corresponding brightness variation.

5) Sensitivity

By connecting the receiver to the socket Bu_4 (fig. 1), where the voltage is only 1/50 of the voltage at Bu_3, a rough impression of the sensitivity of the set can be attained. With this attenuated input voltage, the synchronization should be maintained and the pattern of bars should remain well visible, although the image may possibly show signs of "noise" (similar to swirling snow flakes).

6) Video signal

It may be desired to check the functioning of the output valve of the video section and the picture tube, independent of any faults which may be present in earlier stages.

For this purpose the video signal of the generator, of amplitude approx. 1 volt at the terminals Bu_1-Bu_2 (fig. 1), can be fed directly to the output valve concerned.

Use can also be made of these terminals to check the generator itself, by connecting it to an oscilloscope. In order to give a true representation of the video signal, the oscilloscope used must be capable of giving a good reproduction of pulses of duration 1/4 μ sec. (This requirement is met by the oscilloscope GM 5653.)

7) Sound

It has already been mentioned that with the switch Sk_1 in position 5, the audio channel of the receiver can be checked. The oscillator is then modulated (in frequency or in amplitude) with an audio frequency signal, viz. the voltage supplied by the generator for the horizontal bars. This is a square-wave voltage with a fundamental frequency variable from about 300 to 500 c/s. In order to make the sound less unpleasant to the ear, the square shape is rounded off slightly by a simple filter, and the frame synchronization signal (pulses of mains frequency) is kept away from this voltage.

Economizing in weight by simplification

The present television standards, however much they may differ from each other, have interlaced scanning in common: this implies that the line frequency is an odd multiple of the frame frequency [4]). Now, to derive the frame frequency from the line frequency by frequency division involves somewhat complex apparatus. Furthermore, with all TV standards, the synchronization signals, especially the frame synchronization signals, are fairly complicated in shape [5]). The generation of these signals likewise requires considerable equipment. These are all combined (together with a high frequency oscillator and a small oscilloscope) in a standard generator (type GM 2657) which has been developed for laboratory purposes. For use in the home of the set-owners and in the repair shop, however, such an apparatus is too bulky, too heavy and too expensive. In the design of the portable signal generator GM 2887, dealt with here, the aim has been to drastically cut down on dimensions, weight and price without in any way impairing its usefulness for the specific purpose for which it was intended.

This aim has been achieved largely by abandoning the demand for a fixed ratio between line frequency and frame frequency, so that the frequency dividers could be dispensed with. The result is, of course, that now the picture lines are not absolutely stationary but shift arbitrarily in the vertical direction. A slight deterioration of the picture quality results from this, although with the simple bar patterns this is hardly noticeable and in no way impedes the tracing of faults.

Dispensing with the interlacing has the effect that the line synchronization signal can be simplified since the equalizing pulses are then unnecessary and can therefore be omitted (equalizing pulses are extra line pulses which are necessary in systems with interlacing because there the frame synchronisation signal occurs alternately at the end of a line and in the middle [6])).

The number of valves has been kept down, not only by these simplifications, but by the use of double valves: the signal-generator has nine triode-pentodes ECL 80 and two double triodes ECC 81, and further, one rectifier AZ 41 and one neon tube, a total of 13 tubes. The total D.C. consumption is

[4]) See, e.g. Philips tech. Rev. **10**. 307, 1949, or W. Werner and F. Kerkhof, Television (Philips Technical Library 1952) pp 4 and 5.

[5]) For the synchronization signals conforming to the international standard, see "Standards for the International 625-line black-and-white television system", Sub-group Gerber, CCIR, Geneva, Oct. 1950, or Philips tech. Rev. **13**, 313 (fig. 1), 1952. For those conforming to the British and the (then) French standards, see Philips tech. Rev. **10**, 365 (fig. 1), 1949. See also, International Television Standards, Wireless World **48**, 296-297, Aug. 1952.

[6]) See, e.g. Philips tech. Rev. **13**, 313 (fig. 1), 1952.

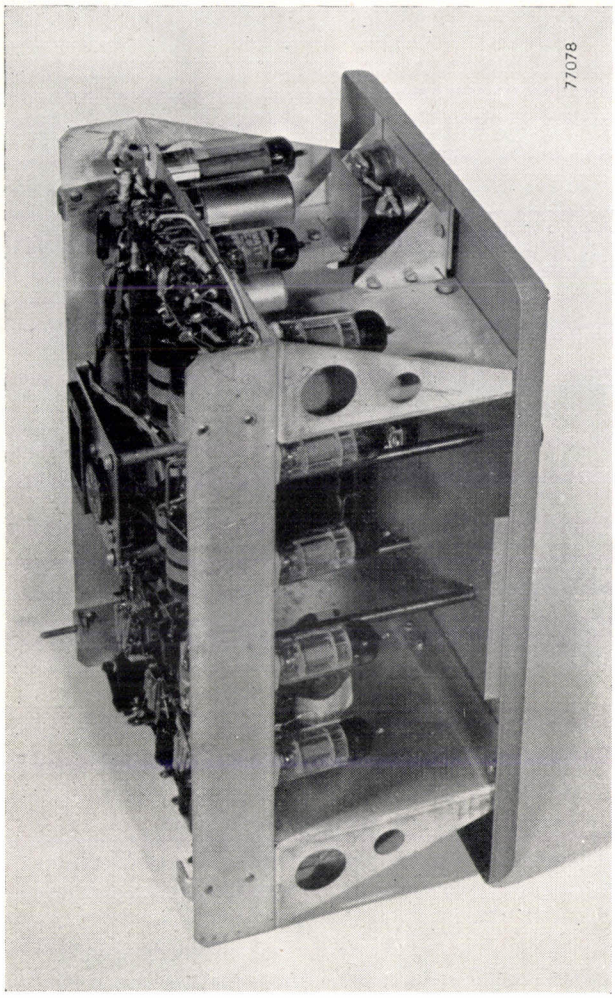


Fig. 9. The interior of the generator GM 2887. The heaviest component, the transformer, is located in the middle with the valves around it.

low (45 mA), and the total power taken from the mains is only about 40 W, so that a small and light transformer is capable of supplying the apparatus.

Fig. 9 gives a view of the interior. To facilitate carrying, the heaviest component, the transformer, is mounted in the centre. The tubes are located around it, and are thus easily accessible.

Description of circuits

Block diagram

Fig. 10 is the block diagram of the generator. On the right-hand side of the diagram is the modulator (*Mod*) in which the high frequency voltage from the oscillator (*Osc*) is modulated by the video signal originating in the unit *Vid*. The unit *Vid* is an "adding stage" in which the picture signal, supplied by *P*, and the synchronisation signals from *Sync* are added together. *P* and *Sync* are also adding stages. The signal corresponding to the horizontal bars is fed from an oscillator *HB*, and that for the vertical bars from an oscillator *VB* to the stage *P*. Both line and frame synchronisation pulses are supplied to the adding stage *Sync*. The former are produced by a unit *LS* controlled by the freely running oscillator *LO*. An oscillator *RS* produces the frame synchronization signals and is synchronized by an oscillator *RB* which produces the frame blanking signals. This in turn is controlled by pulses of mains frequency which are derived from the mains voltage via a network *N*. An oscillator *LB* supplies the line blanking signal and is synchronized by *LS*.

Using switches S_1 and S_2 a choice can be made between horizontal bars, vertical bars, or both at the same time. For testing the audio channel, S_1 is put in the top position (fig. 10); the horizontal-bar generator, on the models with frequency modulation, then feeds a reactance valve (*Re*) which modulates the oscillator *Osc* in frequency. At the same time, with S_1 in the top position, S_3 and S_4 are closed: S_3 short circuits the input of the modulator so that no video signal can occur on it (the modulator then simply operates as an amplifier), and S_4 prevents blanking signals from *RB* occurring in the audio signal supplied by *HB*; these signals, due to their frequency (50 c/s), would make the sound less pleasant to the ear.

Some of the networks mentioned will now be dealt with more fully.

Relaxation oscillators

Several of the networks dealt with are relaxation oscillators which generate a square wave signal. These are the networks *RB*, *HB*, *RS*, *LB* and *VB*.

In *HB*, the generator of the horizontal bars, a triode-pentode ECL 80 is used, operating as a multi-vibrator. The frequency can be varied from about 300 to 500 c/s, corresponding to about 6 to 10 horizontal bars.

The generator of the blanking signal, *RB*, likewise operates with a valve ECL 80 in a similar type of circuit, but the two halves of the tube are asymmetrically connected so as to give a monostable flip-flop. The leak resistor of the triode grid is held at a high positive voltage, whilst the pentode control grid has a negative grid bias (the voltage drop across a cathode resistor); consequently, the circuit is stable only when the triode is conducting. If now a negative pulse is applied to the triode grid, the triode will be temporarily cut-off, and the pentode will be conducting. A quasi-stable condition now exists, which will last for a certain period (dependent on the particular resistance and capacitance values), after which the original condition is restored. In the quasi-stable state, the anode potential of the pentode drops, because of the voltage drop in its anode resistor. Hence the screen grid of the multivibrator *HB* also drops in potential, since it is directly connected to this anode. The screen grid potential drops below the cathode potential of *HB* so that this multi-vibrator does not work while *RB* is in the quasi-stable condition.

As already stated, this latter condition is initiated each time by a negative pulse on the triode grid of *RB*. The network *N* shapes these pulses from the mains voltage, in the

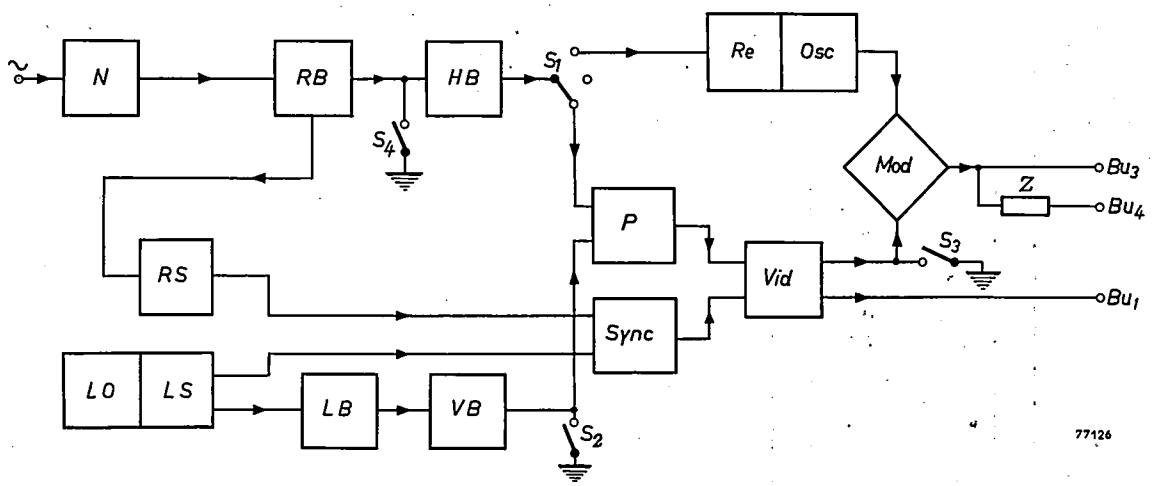


Fig. 10. Block diagram of the generator 2887. *N* generator of pulses of mains frequency. *RB* generator of frame blanking signal. *RS* generator of frame synchronization signal. *HB* generator of horizontal bars. *LO* freely running line oscillator. *LS* generator of line synchronization signal. *LB* generator of line blanking signal. *VB* generator of vertical bars. *P* adding stage for horizontal and vertical bars. *Sync* adding stage of the signals for horizontal and vertical synchronization. *Vid* adding stage for picture and synchronization signals. *Re* reactance modulator (in the models with frequency modulation only). *Osc* high frequency oscillator. *Mod* modulator. Bu_1 output terminal for the video signal. Bu_3 output for modulated high frequency voltage. Bu_4 same output, attenuated by a factor of approximately 1/50 by the attenuater *Z*. S_1, S_3, S_2 and S_4 switches.

following manner. A transformer connected to the mains has its secondary in series with a resistor and a neon tube (type Z1M). The voltage waveform across this tube is represented by the full line in *fig. 11a*. Each time the tube fires, a sudden voltage drop occurs from the ignition voltage to the running voltage. With the aid of a differentiating circuit, a pulse (fig. 11*b*) is obtained at each voltage drop, so that one positive and one negative pulse occur per cycle of the mains voltage. Each negative pulse brings the multivibrator *RB* from the stable to the quasi-stable state, and this stops the multivibrator *HB*.

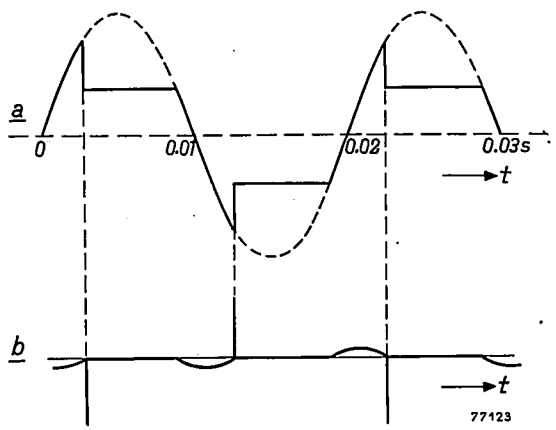


Fig. 11. *a*) Dashed curve: sinusoidal voltage having mains frequency. Full line: voltage across the neon tube in network *N*. From this latter voltage, by differentiation, a steep-shaped pulse (*b*) is produced.

The generator *RS* of the frame synchronization signal is, in the main, arranged similarly to *RB* and is controlled by *RB* — thus ultimately by the mains frequency. *RB* and *RS* thus go over simultaneously into the quasi-stable state, but with *RB* this lasts longer than with *RS*; the difference between them is the duration of the back porch.

The generators *LS*, *LB* and *VB* generate the line synchronization signal, the line blanking signal and the vertical bars signal respectively, and are analogous to *RS*, *RB* and *HB*. Of course, the frequency is much higher: *LS* and *LB* operate with the line frequency f_l, which, according to the system, is approximately 10, 15 or 20 kilocycles; *VB* operates with a frequency which is variable from $6f_l$ to $10f_l$.

Line oscillator and line synchroniser

The line frequency f_l is determined by an *L-C* oscillator *LO*, oscillating at its natural frequency. A fixed capacitance is used and the inductance is permanently adjusted, by means of a sliding core of Ferroxcube, so that f_l approaches the required value as closely as possible.

The pulses which synchronise *LS* and *LB* have to be derived from the sinusoidal voltage of this oscillator. This is done in the following manner.

Across the *L-C* circuit of *LO* there is a sinusoidal voltage approximately 200 volts in amplitude. A part of this (v_2) is fed, via a capacitor and a leak resistor, to the triode grid of the valve ECL 80, the pentode of which acts as the oscillating valve in *LO*. Since the amplitude of the alternating voltage v_2 exceeds the cut-off voltage by many times, there is a high negative grid bias, (*fig. 12a*) so that the triode passes current only at the positive peaks of the alternating voltage (fig. 12*b*). These anode current pulses drive a second *LC* circuit; the voltage v_3 developed across this circuit is shown in fig. 12*c*. This voltage is used for generating the line synchronization pulses, in such a way that these pulses continue during the frame synchronization, so that the saw-tooth generator of the receiver has no chance of getting out of step. It would be going too far to deal with the details of this circuit here.

Adding stages

There are three adding stages in the GM 2887 signal-generator: *P*, *Sync* and *Vid* (fig. 10); in *P* the picture signal is formed, and in *Sync* the synchronization signal, whilst in *Vid* these two are combined to form the video signal. The "adding" of the two signals v' and v'' is actually here the shaping of a

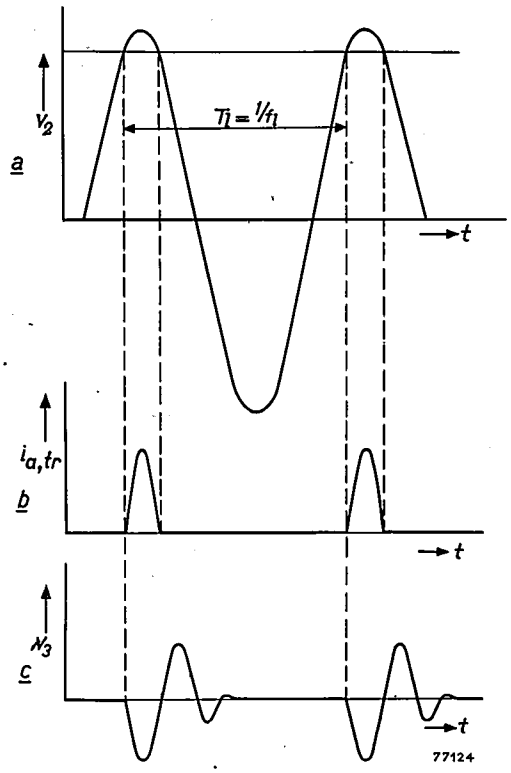


Fig. 12. *a*) Sinusoidal voltage v_2 across a part of the LC circuit of the line oscillator LO. Only the peaks of v_2 are above the cut-off voltage of the triode of LS. *b*) Pulse-shaped anode current of the triode of LS. *c*) The damped alternating voltage v_3 across the LC circuit of LS.

linear combination of them, say $av' + bv''$. As shown in *fig. 13* this can be effected with the aid of two valves having a common anode resistance. The adding stages in question function according to this principle; in each of the adding stages two valves as shown in fig. 13 are combined into one double tube, again of the type ECL 80.

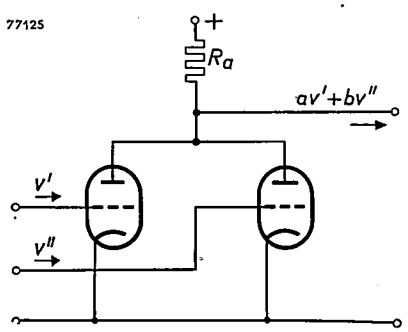


Fig. 13. Principle of an adding stage: two amplifying valves with a common anode resistance R_a. A signal $av' + bv''$, which is a linear combination of the signals v' and v'' on each of the grids, is produced at the anodes.

Oscillator, modulator and reactance modulator

The networks drawn in fig. 10 are almost identical for the four models (A, B, C and D), except for the line frequency and the frequency range of *VB*. There remain, however, three more networks: the high frequency oscillator *Osc*, the modulator *Mod* and (in the models with frequency modulation) the reactance modulator *Re*. Together these form a unit. The models with frequency modulation (A and C) will mainly be dealt with here.

One half (*II*) of a double triode ECC 81, (*fig. 14*), together with a resonant circuit L_1C_1 forms the high frequency oscillator. The other half of the tube (*I*) in the models with frequency modulation (A and C) serves as a reactance valve, i.e. it behaves as a reactance connected in parallel to the circuit L_1C_1, of magnitude dependent upon the slope of the valve. This slope varies with the audio frequency signal, so that the oscillator voltage is modulated in frequency. As already remarked, the signal of the horizontal bars, produced by the generator *HB*, functions as an audio-frequency signal.

The oscillator voltage is fed to one half (*III*) of a second double triode ECC 81. This half is arranged as a cathode follower. The cathode is thus the output electrode; it is connected with the cathode of the other half (*IV*) of this tube, which functions as a modulator. The video signal derived from the adding stage *Vid*, is applied to the grid of *IV* which for high frequency voltages is earthed across a capacitor. The output is taken off at the anode of *IV*, and in models *A*, *B* and *C*, passes via two coupled coils (L_2-L_3), so as to give a signal symmetrical with respect to earth. It is available in this form at the socket Bu_3 and, attenuated by a factor of 1/50 by a potentiometer (*Z*), at socket Bu_4.

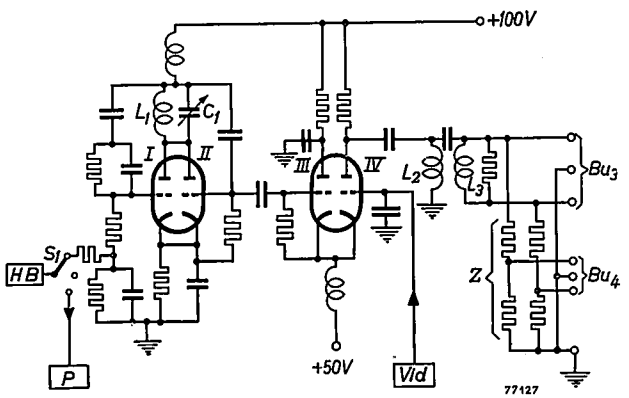


Fig. 14. Circuit of the reactance modulator, the high frequency oscillator and the modulator. Of two double triodes ECC 81, *I* serves as a reactance valve, *II* (together with the vibrator circuit L_1C_1) as oscillator, *III* as amplifier, arranged as cathode follower, and *IV* as modulator. For high frequencies the grid of *IV* is earthed across a capacitor. To this grid is applied the video signal from the adding stage *Vid* which combines the picture signal and the synchronization signal. A symmetrical output voltage is obtained at the socket Bu_3 by means of the coupled coils L_2 and L_3. This voltage is attenuated at the socket Bu_4 to about 1/50, by a potentiometer composed of resistors (*Z*). The audio-frequency signal is fed to the reactance valve via the switch S_1.

The video-signal is supplied to the modulator with such a polarity that the modulation has the required sign (the International and the American standards specify negative modulation, while the British and the French standards specify positive modulation).

The versions of the generator GM 2887 for these latter two systems — models B and D — differ from the diagram in fig. 14 in so far as the second channel uses amplitude modulation instead of F.M.

Summary. The signal generator GM 2887 described supplies signals with which various types of faults in TV receivers may be traced, independent of the operation of the local TV transmitters. The apparatus is also useful for adjustment of

the picture. In the design, the aim has been to make the generator easily portable (weight 7 kg, dimensions 335 × 220 × 165 mm, or $15^1/_2$ lb. and $13'' \times 8^1/_2'' \times 6^1/_2''$). The number of valves has been restricted to 13 (9 triode-pentodes ECL 80, 2 double triodes ECC 81, 1 rectifying tube AZ 41 and 1 neon tube Z1M) — thanks to the use of double tubes, and the omission of frequency dividers for the derivation of frame frequency from line frequency. The mains frequency is used as the frame frequency, and the line frequency is generated by a freely running LC oscillator. As a consequence, the pattern lines do not remain absolutely stationary, but this is not a serious objection in view of the simplicity of the patterns supplied by the generator (horizontal and/or vertical bars). Due to the fact that in the various countries different TV standards are in use, the generator is manufactured in four models:

Models A and C with negative picture modulation, 625 lines or 525 lines, and frequency-modulated audio-channel. Model A is for the low frequency band and C for the high frequency band of the countries with either the "International" or the American standard; B and D with positive picture modulation and amplitude-modulated audio channel. Model B is for 405 lines (British standard), Model D for 819 lines (French standard).

BOOK REVIEW

Data and circuits of radio receiver and amplifier valves (Second supplement: Valves developed between 1945 and 1950), by N. S. Markus and J. Otte, pp. 487, 505 illustrations. — Philips Technical Library, 1952 (Volume IIIa of the Electronic Tube Series) — This book may be ordered through your technical bookseller.

This volume, together with volumes II (1933-1939) and III (1940-1941), forms a compendium of the electronic valves (and some measuring instruments) developed by Philips during the years 1945-1950. Volume III, which precedes this one, has not been reviewed in this journal, as it was published shortly before the Philips Technical Review had to cease publication in 1942. The present volume deals chiefly with the small tubes produced since the war for broadcast receivers - the so-called "Rimlock" valves (E, U and D series). In addition, the book deals with four miniature D series valves for battery operation, and three "Noval" tubes which have a nine-pin base. The book also contains descriptions of 15 new measuring equipments manufactured by Philips, which supplement those given in Volumes II and III.

The great improvements in the mechanical construction of receivers which has resulted from the development of these small tubes is at once evident from a comparison of the interior of a set equipped with these tubes with an earlier set embodying similar circuits. In the introduction to this book, it is explained how the dimensions of these valves were reduced without impairing their characteristics, reliability or cost.

In all, 36 types of valve are described, together with a number of detailed circuits of complete receivers and amplifiers in which they are used. An outstanding feature of the book is the comprehensive way in which each type of tube is discussed. The reader is thus able to find everything he wants to know about a certain valve. This method of treating the subject has made the book large but it enhances its value: the more so because the text also gives the agreeable impression that it has been compiled with great care.

The scope of this review does not permit a separate enumeration of all the Philips measuring instruments described: they include two cathode-ray oscilloscopes, high and low frequency oscillators and electronic voltmeters. The most impressive apparatus is the large oscilloscope GM 5653, whose amplifier has a frequency range of 5 Mc/s, which makes it admirably suitable for television servicing. Also described is a standard oscillator, type GM 2885, driven by a crystal-controlled 1 Mc/s oscillator, from which a large number of precise derived frequencies can be obtained by frequency multiplication and by mixing with an auxiliary oscillator (0.5-1.0 Mc/s). A third apparatus worthy of mention is the variable low frequency *RC*-generator, GM 2315.

A few of the valves described are special types whose field of application lies outside that of normal broadcast reception. The most important of these are the "φ-detector" or enneode [1]) EQ 80 (for F.M.), the multi-purpose pentode EF 42 (for wide frequency band amplification) and the ubiquitous double triode ECC 40 (for push-pull circuits, television apparatus, pulse generators, electronic computors, etc.). The discussion of this last type of tube and its applications, though occupying 16 pages, could, in the writer's opinion, have been more elaborate; some literature references for this type of tube and also for the EQ 80 are desirable. This criticism, however, should in no way mitigate the favourable general impression of this very practical book. J. van SLOOTEN.

[1]) See also J. L. H. Jonker and A. J. W. M. van Overbeek, Philips tech. Rev. **11**, 1-11, 1949.

THE FOCUSING OF TELEVISION PICTURE-TUBES WITH FERROXDURE MAGNETS

by J. A. VERHOEF.

621.318.124:621.385.832:621.397.62

*In 1950-51 Philips developed Ferroxdure *), a ceramic permanent-magnet material. One of the most important applications of this material, which takes full advantage of its special properties, is in focusing magnets for television picture-tubes.*

Magnetic electron lenses

For the focusing of the electron beam in television picture-tubes, magnetic lenses are commonly used. The focusing action results from the magnetic field which has rotational symmetry about the axis of the lens. The focal length f is given by

$$\frac{1}{f} = \frac{0.0347}{V} \cdot \int_{z=-\infty}^{z=+\infty} H_z^2 \, dz , \quad . \; . \; . \; (1)$$

where V is the potential difference traversed by the electrons before they enter the lens and H_z is the field strength on the axis of the lens [1]).

If an electromagnetic lens is used, both V and H_z vary somewhat with the supply voltage. The circuit can be designed such that these variations largely compensate each other. This is of course no longer possible when permanent magnets are used because only V then varies with the supply voltage. Nevertheless, permanent magnet lenses are usually preferred nowadays to the electromagnetic type of lens. There is no need to supply the considerable power which is required to energize an electromagnetic lens, so that the power pack can be simpler. At the same time one source of heat in the apparatus is eliminated. Finally, the cost of manufacture of permanent magnet lenses is considerably less than that of electromagnetic lenses.

In spite of the above-mentioned advantages, the conventional permanent magnet lens has the serious disadvantage that the field is spread out along the axis much further than that of an electromagnetic lens of the same strength. The reasons why such an extended field is undesirable will first be discussed. The fundamental difference between electromagnetic and permanent magnet lenses will then be considered more closely. Finally it will be shown how the stray field of permanent magnet lenses can be considerably restricted. It will become evident that, for this purpose, a magnetic material of high coercivity is necessary. Some of the "Ticonal" magnet steels satisfy this requirement. However, an important economic advantage is gained by the use of Ferroxdure, the new permanent-magnet ceramic, one of whose distinctive properties is a high coercivity [2]). In addition to the low cost of Ferroxdure, its use leads to a very simple construction.

Some consequences of a strong stray field

The electron lens is situated on the neck of the tube between the ion trap and the deflection coils (*fig. 1*). Especially when a short-neck tube is used [3]), these three components are so close together that the field from a conventional permanent magnet lens still has a considerable value at the ion trap and the deflection coils (*fig. 2*). The result is that if the position or strength of the lens is changed, the angle through which the ion trap deflects the electron beam is also changed. The beam then goes eccentrically through the tube, so that the raster, which is "written" by the electron beam on the screen of the tube is no longer well-centered. Furthermore, part of the beam will now be cut off by the diaphragm (*fig. 3*), so that the brightness

*) Also known by the name "Magnadur".

[1]) With the constant used, the formula is expressed in rationalized Georgi units. V is expressed in volts, z and f in metres and H_z in A/m (1 A/m = 1.251×10^{-2} oersted).
The validity of eq. 1 is limited to the case of a "thin" lens, i.e. a lens in which the axial extent of the field is small compared with the object and image distances. If the lens is not thin, other terms must be added to the expression (1). As H_z appears, in these terms too, only in even powers, the focal length is always independent of the sign of H_z, the importance of which will be seen presently.

[2]) J. J. Went, G. W. Rathenau, E. W. Gorter and G. W. van Oosterhout, Ferroxdure, a class of new permanent magnet materials. Philips tech. Rev. **13**, 194-208, 1952.

[3]) The reasons why the neck of a picture tube should be kept short, and means by which this may be achieved were discussed recently in this Review. See J. L. H. Jonker, A short length direct-view picture-tube, Philips tech. Rev. **14**, 361-367, 1953 (No. 12).

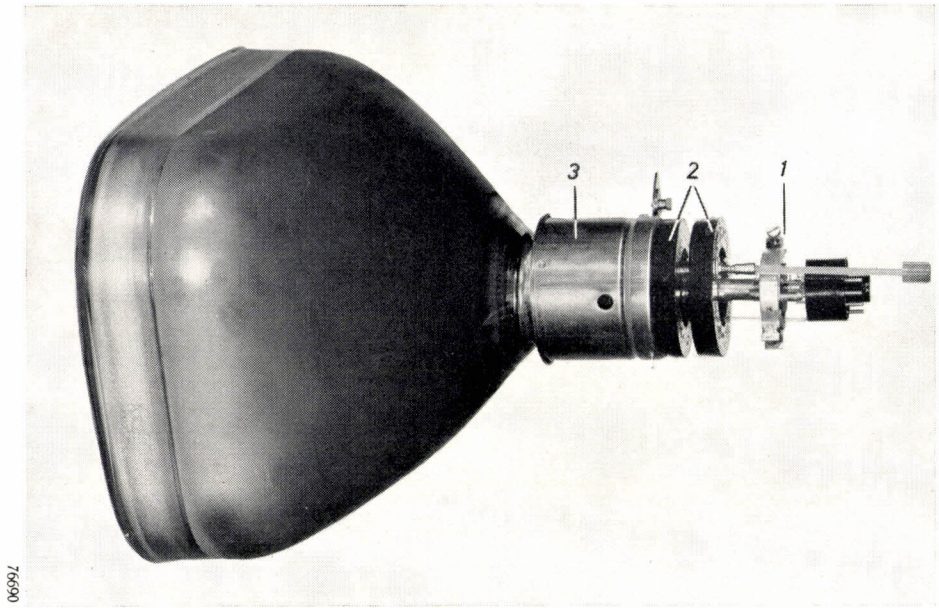


Fig. 1. Television picture-tube (diagonal = 36 cm). Mounted on the neck are: *1* ion trap magnet, *2* electron lenses, *3* deflection coils.

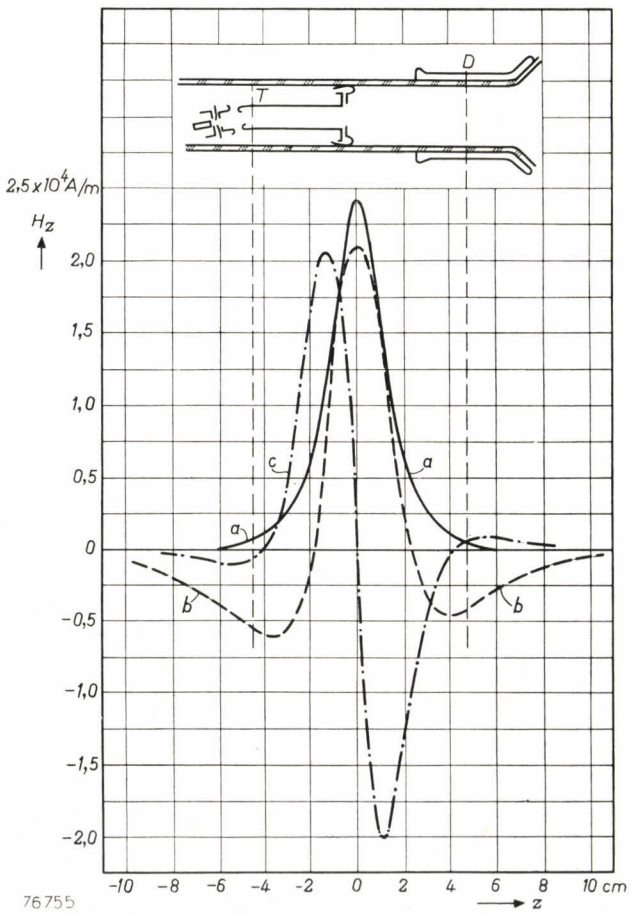


Fig. 2. The form of the field along the axis of a picture-tube, due to *a*) an electromagnetic lens, *b*) a conventional permanent magnet lens, and *c*) a permanent magnet lens of the new type. The conventional permanent magnet lens (*b*) gives rise to a field with a considerable "tail", which is still quite strong at the ion trap *T* and at the deflection coils *D*.

diminishes. If the field of the focusing magnet extends to the deflection coils, a rotation is caused as well as a distortion of the picture, which changes when the lens is adjusted. By a renewed adjustment of the ion trap magnet and of the deflection coils,

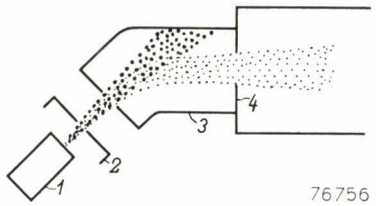


Fig. 3. Schematic representation of an ion trap. The electron beam emitted by the cathode *1* always contains negative ions which would damage the screen of the tube. A local magnetic field, perpendicular to the plane of the drawing, bends the electron beam so that it passes through the aperture in the diaphragm *4*. The ions (large dots) are negligibly affected and thus fall on the anode *3*. *2* is the control grid.

these distortions can be corrected again, but the mutual interaction between the electron lens and the ion trap magnet and the deflection coils is very undesirable; apart from the inconvenience of the successive adjustments, the adjustment for the optimum light spot on the screen is upset.

Electromagnetic and permanent-magnet electron lenses

An electromagnetic lens consists of a coil enclosed in a soft iron casing provided with an internal air gap (*fig. 4a*). The strength of the lens depends on

the current through the coil. A permanent magnet lens of conventional construction consists of an axially magnetized ring of "Ticonal" or a similar magnet steel, provided with pole pieces (*fig. 5a*).

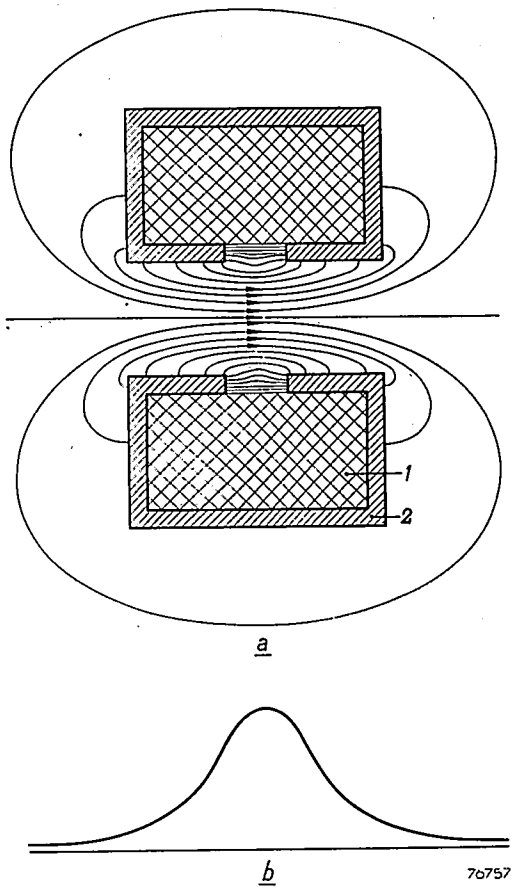


Fig. 4. *a*) Schematic representation of the construction of an electromagnetic electron lens showing the lines of force. The coil *1* is surrounded by a soft iron casing *2*, which has an air gap. The strength of the lens is controlled by varying the current through the coil.
b) The form of the field along the axis of the lens shown in (*a*).

The strength of such a lens can be varied by means of an adjustable soft iron ring. The more this ring bridges the gap between the pole pieces (the more the air gap is short-circuited), the weaker the lens. In figs. 4*b* and 5*b* are plotted the field strengths along the axis for the two types of lens. The essential difference is that with an electromagnetic lens the field strength is nowhere negative, while that of the permanent magnetic lens necessarily has both positive and negative values. Considering the area between the curve and the z-axis as positive above the axis and negative below it, the total area for a permanent magnet lens amounts to zero, whilst for an electromagnetic lens it is proportional to $n \cdot i$ (n is the number of turns of the coil and i is the current). This area is, in fact, a measure of the line integral of the field through the lens along the axis. From Maxwell's theory, in the case of permanent magnets,

$$\int_{z=-\infty}^{z=+\infty} H_z \mathrm{d}z = 0,$$

since no electric currents are flowing. For an electromagnetic lens on the other hand we have

$$\int_{z=-\infty}^{z=+\infty} H_z \, \mathrm{d}z = n \cdot i.$$

The strength of a lens is determined by the form of H_z^2 (see Eq. (1) and footnote [1])), so that, as far as its focusing action is concerned, it is of no consequence that the line integral of the field of the permanent magnet is zero. With a Ferroxdure lens the importance of this fact is even more striking than with a conventional lens, as will be seen presently.

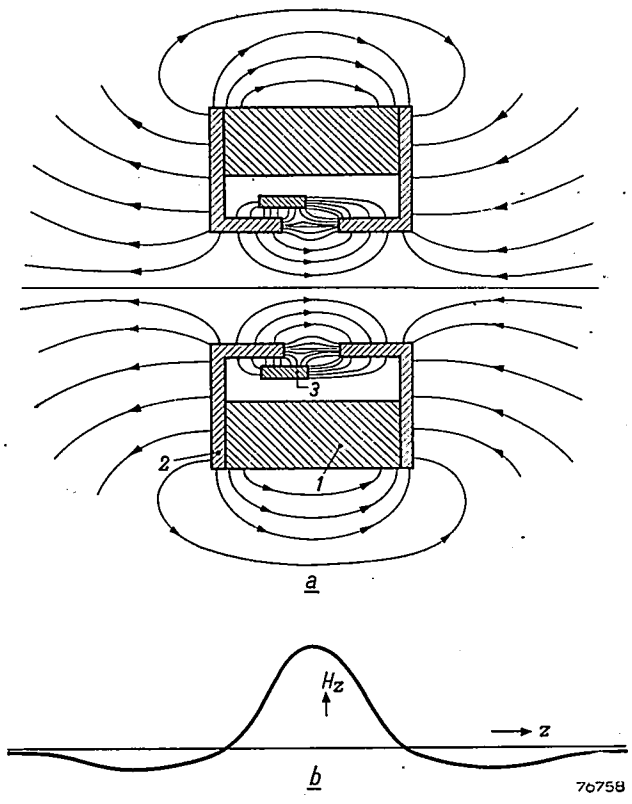


Fig. 5. *a*) Schematic representation of the construction of a permanent magnet lens showing the lines of force. *1* annular piece of magnet steel, *2* pole pieces, *3* moveable soft-iron ring by means of which the strength of the lens may be varied. *b*) Form of the field along the axis of this lens. The algebraic sum of the areas under the curve above and below the *z*-axis is here zero.

A better physical picture of the essential difference between an electromagnetic lens and a permanent magnet lens may be obtained as follows. First consider first the field of a cylindrical film of current (of length L and diameter D), e.g. that which results from passing a current through a coil of one layer of thin wire(of diameter small compared to D). This field can be calculated by elementary methods. In *fig. 6* the field along the axis

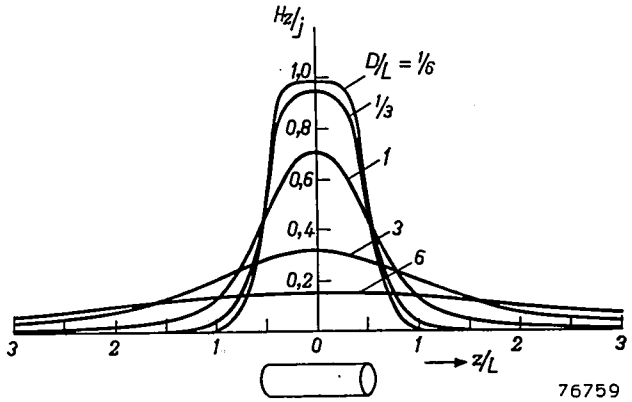


Fig. 6. Calculated distribution of the axial field of cylindrical current films of length L and diameter D. The position of the current film is shown by the sketch just under the graphs. The greater the parameter D/L, the more the field is spread out along the axis. Along the abscissa is plotted H_z/j, where j is the current per unit length of the current film. If j is expressed in A/m, then H is also given in A/m (in rationalized Giorgi units, 1 A/m $\approx$ 1/80 oersted).

is plotted for different values of the ratio D/L. It can be seen that for constant length L, the field is more concentrated when the current film diameter D is smaller. Neglecting the influence of the soft-iron casing of the electromagnetic lens, the field of the whole coil can be obtained by the superposition of the fields of the separate current films with different values of D/L. The resulting field will be roughly bell-shaped, as in the curves of fig. 6. For a permanent magnet lens the situation is different. Neglecting the influence of the pole pieces, the lens has the form of a ring-shaped magnet. It is assumed that the ring is uniformly magnetized in an axial direction and that the magnetization is due to circular atomic currents which lie in a plane perpendicular to the axis of the magnet (*fig. 7*). The electric charges, except those which move at the inner and outer surfaces of the ring, neutralize each other's influence outside the magnet itself. The net effect of the whole magnetization, as far as the space outside the

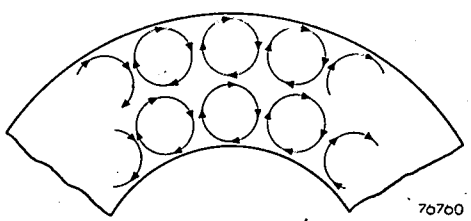


Fig. 7. The magnetization of a uniformly, axially magnetized ring can be thought of as originating in circular atomic currents in the plane perpendicular to the axis of the ring. The field outside the magnetized material is then equivalent to that of two current films flowing in opposite directions along the inner and outer surfaces of the ring.

magnet is concerned, is the same as the effect of two current films of equal strength per unit length flowing in opposite directions along the inner and outer surfaces. The superposition of the fields of these current gives the field of the magnetized ring (*fig. 8*).

Some idea of the effect of a soft iron casing placed round the current coil, in the case of the electromagnetic lens, can be obtained by regarding the total field as due to the superposition of the original field of the coil and the field of the casing, which becomes magnetized by the field of the coil. Due to the casing the field becomes more concentrated and less spread out along the axis (see the lines of force in fig. 4*a*). If the casing is made without an air gap, a field would still exist outside the casing ($\oint H_l \, dl = n \cdot i$ remains valid!)

The form of the field of a permanent magnet lens with pole pieces can be visualized by considering the lens as a magnetic circuit with two parallel air gaps — a narrow gap on the inside and a wide gap on the outside. The field in the latter gap is responsible for the negative parts of the field on the axis (fig. 5.) The negative part of the field cannot be avoided, because the value of the line integral along the axis must always be zero. Should one try to avoid the negative field, for example by making the outer air gap smaller with the aid of pole pieces, then at the same time the useful field of the inner air gap would be weakened. Closing one of the air

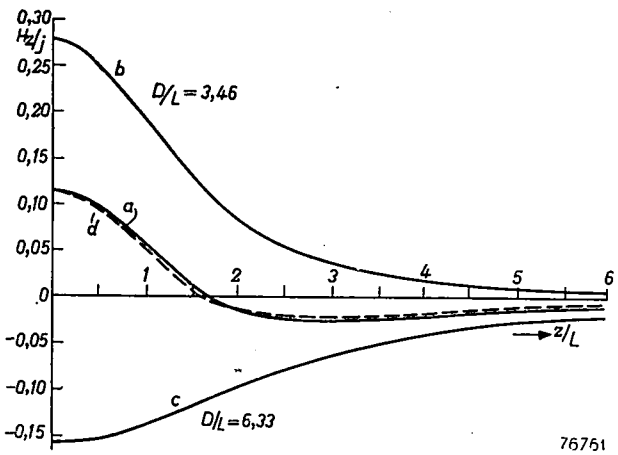


Fig. 8. Theoretical distribution of the field along the axis of a homogeneously magnetized ring (curve *a*), obtained by superposing the fields due to the current films on the inner (curve *b*) and outer surfaces (curve *c*), where D/L is 3.46 and 6.33 respectively. (These values apply to the Ferroxdure rings used for focusing in television picture-tubes.) Curve *d* is the measured field of such a ring. The maxima of the theoretical and the measured curves have been made coincident. The slight deviation between the two curves is attributable to inhomogeneous magnetization in the axial and radial directions. The curves are symmetrical on either side of the vertical axis.

gaps would have the result that the field on the axis would disappear completely. The task of the pole pieces in the conventional permanent magnet lens is principally to enable the lens strength to be varied by means of the moveable ring (fig. 5*a*), and the smoothing out of rotational asymmetries of the magnetization of the magnet ring.

Focusing with Ferroxdure magnets

To reduce the stray field, a lens hasb een developed which consists of two flat rings of Ferroxdure, coaxially mounted and axially magnetized in

opposite directions (*fig. 9*). Apart from the central region, the fields now largely neutralize each other (*fig. 10*). The total area under the curve $H_z = f(z)$, of course, remains zero, for this is the case for any combination of permanent magnets. As has already been pointed out, this is not in any way opposed to the satisfactory operation of this lens, because the strength of the lens is determined exclusively by the form of H_z^2 (dotted line, fig. 10*b*).

reduce the magnetization to zero is about 2.4×10^5 A/m (3000 oersted), while that of Ticonal E amounts to only 0.53×10^5 A/m (660 oersted).) Compared with Ticonal E, Ferroxdure is therefore much less demagnetized by the field of the magnetic poles which establish themselves at the end faces of the ring. With a given volume of magnetic material and a given internal radius, if the axial length L is made smaller (and so the radial dimension h greater,

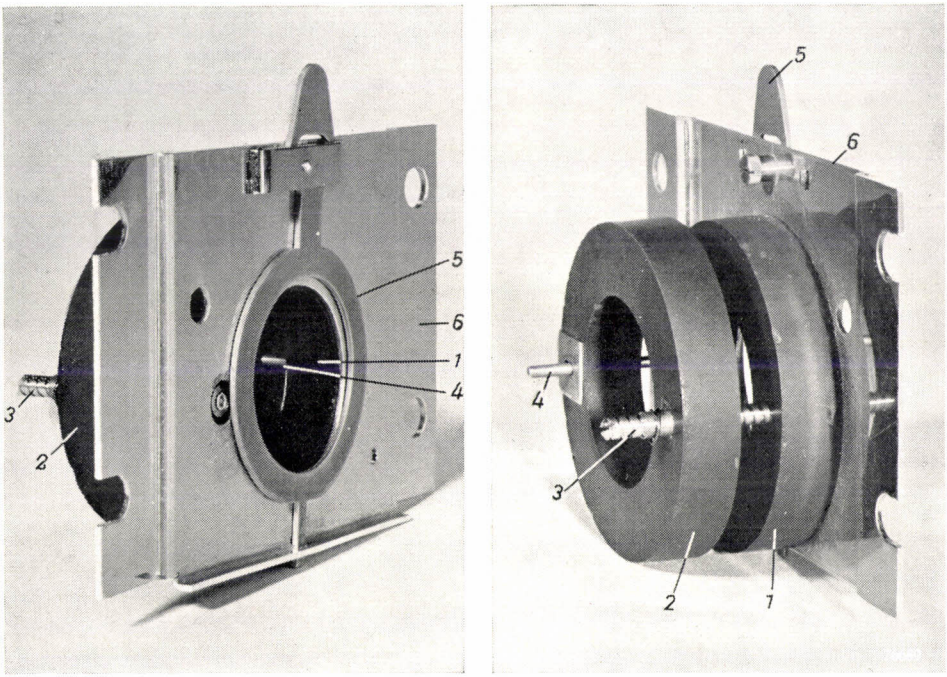


Fig. 9. Two photographs of a focusing unit made up of two Ferroxdure rings (*1* and *2*) axially magnetized in opposite directions. The distance between the rings is adjusted by the screw *3*. *4* is a guide pin, *5* an adjustable soft iron ring with which the electron beam can be centered. The plate *6* is of aluminium. The focusing unit seen in fig. 1 is of slightly different construction.

As the value of $(BH)_{max}$ is much smaller for Ferroxdure than for the magnet steels commonly used in conventional permanent magnet lenses, e.g. "Ticonal" E, it is not surprising that a greater total volume of Ferroxdure is needed than of Ticonal "E", to produce a lens of a given strength [4]). In spite of this fact and in spite of the division of the material into two separate rings, the whole lens is no longer than the conventional lens, thanks to the flat form of the rings. This flat shape is possible owing to the great coercitivity of Ferroxdure. (The field strength which would be necessary to

fig. 11) then the poles which produce the demagnetization field come closer together and the strength of this increases. In the case of Ticonal E, this field would soon reduce the magnetization to only a fraction of the remanence: if two flat, oppositely magnetized rings of Ticonal E were assembled to form a lens, they would destroy each others magnetization almost completely. Evidently this material cannot be applied in the case of flat rings. With Ferroxdure, on the other hand, however thin the ring is made, the magnetization (*fig. 12*) never falls below 80% of the remanence [5]). In the extreme case, when the thickness

[4]) A. Th. van Urk, The use of modern steels for permanent magnets, Philips tech. Rev. **5**, 29-35, 1940. Some further remarks concerning the product $(BH)_{max}$, may be found in the article referred to in note [2]), on p. 116 and p. 202.

[5]) Lenses consisting of two oppositely magnetized rings can be constructed of "Ticonal", provided that Ticonal K is used; for completeness its demagnetization curve is also given in fig. 12. The application of Ferroxdure, however, leads to a more economic solution.

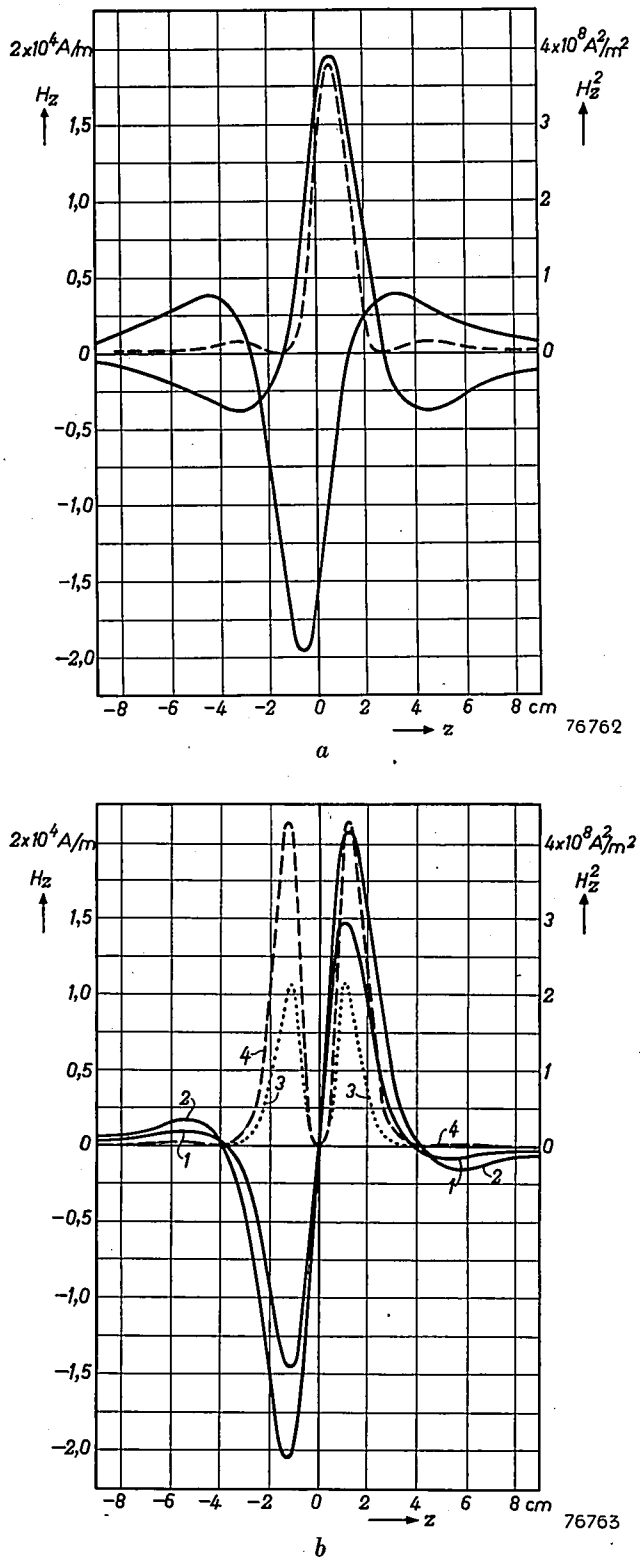


Fig. 10. *a*) The separate axial fields of two oppositely directed Ferroxdure rings are depicted here in the one figure. The curves are displaced with respect to each other by a distance corresponding to the thickness of one ring. For one of the rings (the right-hand one) the form of H_z^2 is drawn in dotted lines.
b) Curve *1* shows the resultant field along the axis of a lens with the two rings placed face to face; curve *2* shows the field when the rings are separated by 10 mm. The dotted curves *3* and *4* show the form of H_z^2 along the axis for these two cases.

of the ring is made equal to zero, the working point (i.e. the point on the demagnetization curve corresponding to the magnetic condition of the material) is found to be the intersection of the line $J = -\mu_0 H$ with the demagnetization curve (μ_0 is the permeability of free space). The disastrous result for Ticonal E can be seen in fig. 12, where the magnetization drops from a very high value for a

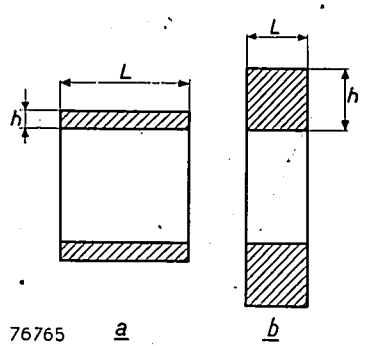


Fig. 11. Cross-sections of ring-shaped magnets for television tube focusing.
a) For a lens of the conventional type using magnet steel.
b) For a lens of the new type, using Ferroxdure.

long magnet, to the point P_1 for a magnet of zero length; on the other hand the change in the magnetization of Ferroxdure ($R - P_2$) is very small indeed.

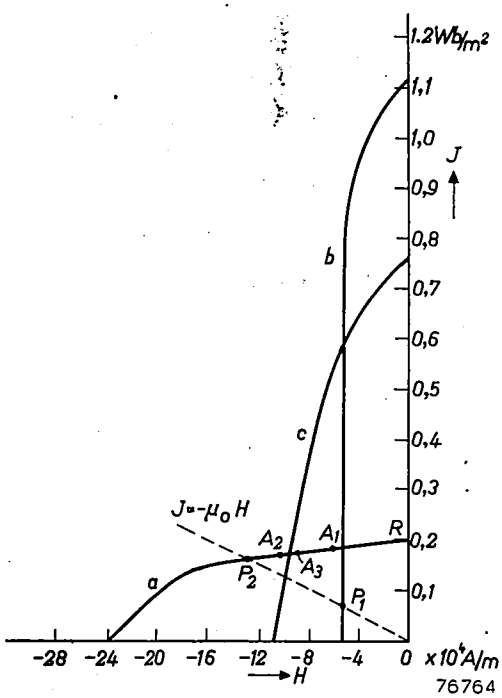


Fig. 12. Demagnetization curves of Ferroxdure (curve *a*), "Ticonal" E (curve *b*) and "Ticonal" K (curve *c*). Note that the magnetization J is plotted along the vertical axis, and not the magnetic induction. The point R represents the remanent magnetization of Ferroxdure. With a focussing unit as in fig. 9 the working point of the Ferroxdure lies between A_2 and A_3, depending on the distance between the two rings. A_1 is the working point for a single ring not influenced by the field of another magnet. P_1 and P_2 are the working points for very thin flat discs of Ticonal E and Ferroxdure respectively.

Even with Ferroxdure, however, the rings cannot be made thinner indefinitely. The opposite poles on the two end faces will neutralize each other's external fields more and more as the rings are made thinner. Accordingly, the field along the axis decreases quickly as L becomes considerably smaller than h. For the actual rings used, $L \approx h$.

The ceramic Ferroxdure rings can more easily be made homogeneous than rings made from magnet steel. Hence the magnetization shows the same rotational symmetry as the ring and no soft iron pole pieces are necessary to ensure a field with rotational symmetry.

The strength of the Ferroxdure lens is varied by adjusting the distance between the two rings (fig. 9). As the distance is increased, there is less mutual neutralization of the main fields, and thus the lens becomes stronger. The neutralizing of the stray field is then less complete, but it is still adequate, even over the whole range of adjustment of the lens strength, which is at least as great as that of a conventional permanent magnet lens, and often greater. ("Adequate" neutralization implies that the operation of the ion trap magnet and of the deflection coils are only slightly influenced by the electron lens. The adjustment of these components can then be made independently). In a given case, it proves possible to obtain good focusing of the electron beam for any acceleration potential between 10 and 20 kV (see Eq (1)), by varying the separation of the two rings between 0 and 12 mm.

The two rings, magnetized in opposite directions and situated in each other's field, exercise on each other a demagnetizing field. If the separation of the rings is such that they exert no demagnetizing field on each other, the working point lies approximately at the point A_1 (fig. 12). As the rings are brought closer together, the working point moves towards A_2. With the normal spacing of the rings, the working point lies between A_2 and A_3. The position of the working point is important because, with Ferroxdure, the magnetization depends on the temperature, and the temperature coefficient varies with the position of the working point. Near the point R the temperature coefficient is about -0.2% per °C, but in the region between A_2 and A_3 it is only about half this value. With some thought, television sets can be so designed that the rise of temperature of the magnet rings remains so small that the influence on the picture definition is imperceptible.

In practice the picture is never accurately centered on the screen of the tube without some adjustment. In the Ferroxdure lens, a thin soft iron ring is placed against the first magnet (fig. 9). By moving this soft iron ring, a movement of the picture across the screen can be achieved.

Summary. By the application of the distinctive properties of Ferroxdure, the new ceramic material for permanent magnets, it is possible to construct a simplified permanent magnet lens for the focusing of the electron beam in television picture-tubes. The lens is characterized by a stray field along the axis which is considerably less than that of permanent magnet lenses of the conventional type. The lens consists of two axially magnetized rings of Ferroxdure which are mounted together, like poles facing each other. The stray fields of the two rings then largely neutralize each other. The strength of the lens is adjusted by varying the mutual separation of the rings. Centering the picture on the screen is accomplished simply by moving a soft iron centering ring.

Adjustment of the lens strength does not affect the adjustment of the ion trap magnet, and causes only a negligible rotation of the picture. The adjustment of the ion trap magnet, the lens itself and the deflection coils is thereby simplified. As a result of the simple construction of the lens and of the low price of Ferroxdure, the new electron lens is also cheaper than existing lenses for this purpose.

VOL. 15 No. 8-9, pp. 221-264 FEBRUARY-MARCH 1954

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
THE PHILIPS INDUSTRIES

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, NETHERLANDS

THE FLYING-SPOT SCANNER

by F. H. J. van der POEL and J. J. P. VALETON.

621.385.832:621.397.611.2:
535.373.3

In laboratories, factories engaged in the manufacture of television equipment and in studios, it is often desirable to have available a well-defined and reproducible television signal. The "flying spot scanner" supplies such a signal, starting in the first instance from a flat, transparent object. The principle of the flying-spot scanner dates from the last century; the article below discusses a modern application of the principle.

The flying-spot scanner is an apparatus for generating a television signal from a flat object. If, in particular, this object is transparent, such as a photographic plate, film or lantern slide, or a microscope slide, the flying-spot scanner gives a signal of very good quality. In the following description a stationary, transparent object is assumed, e.g. a lantern slide.

The principle of the flying-spot scanner is given in *fig. 1*. A point source of light is projected on to the transparency. The source is made to trace out a frame or raster of the required number of lines

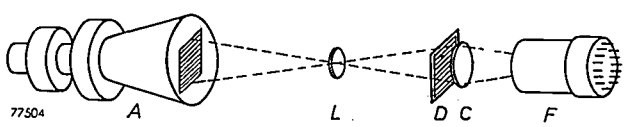


Fig. 1. Principle of the flying-spot scanner. The electron beam of a special cathode-ray tube *A* (the scanning tube) describes a frame on the screen, which is projected on to the flat transparent object *D* by the lens *L*. The condenser lens *C* collects the light passing through and throws it on to the photo-cathode of the multiplier tube F.

so that its image traces an identical frame on the transparent object. The light passing through the latter is concentrated on a photo-electric cell by means of a condenser lens, so that a photo-current is generated which at every moment is proportional to the transparency of the object at the point defined by the spot. This current is passed through a resistor, the voltage across which, after amplification, represents the required picture signal.

The photo-cathode gives off a very small current, viz. of the order of 10^{-9} A. By the use, however, of a photo-electric cell with built-in secondary emission amplification (photo-multiplier tube), the final signal current has a value in the region of 0.5 mA. The advantage of this method of current amplification over conventional methods is that a more favourable signal-to-noise ratio is obtained. This fact largely explains the better quality of the signal given by the flying-spot scanner as compared with that of the normal camera tubes.

In the flying-spot scanner to be discussed (*fig. 2*), the light spot on a specially designed cathode-ray tube serves as the light source. To trace out the raster, the electron beam is deflected as in a television receiver picture-tube, but of course, its intensity is kept constant. An early type of flying-spot scanner in which the movement of the light source was obtained by mechanical means was described in this Review in 1937 [1]).

In the present article, after some remarks on the optical system, the afterglow of the fluorescent material on the screen of the scanning tube is discussed. In particular, the specific properties of the phosphor are set forth.

This is followed by a discussion on the non-linear amplification of the signal (gamma-correction) which is necessary to faithfully reproduce on the picture tube the contrasts of the object.

[1]) H. Rinia and C. Dorsman, Television system with Nipkow disc, Philips tech. Rev. **2**, 72-76, 1937. H. Rinia, Television with Nipkow disc and interlaced scanning, Philips tech. Rev. **3**, 285-291, 1938.

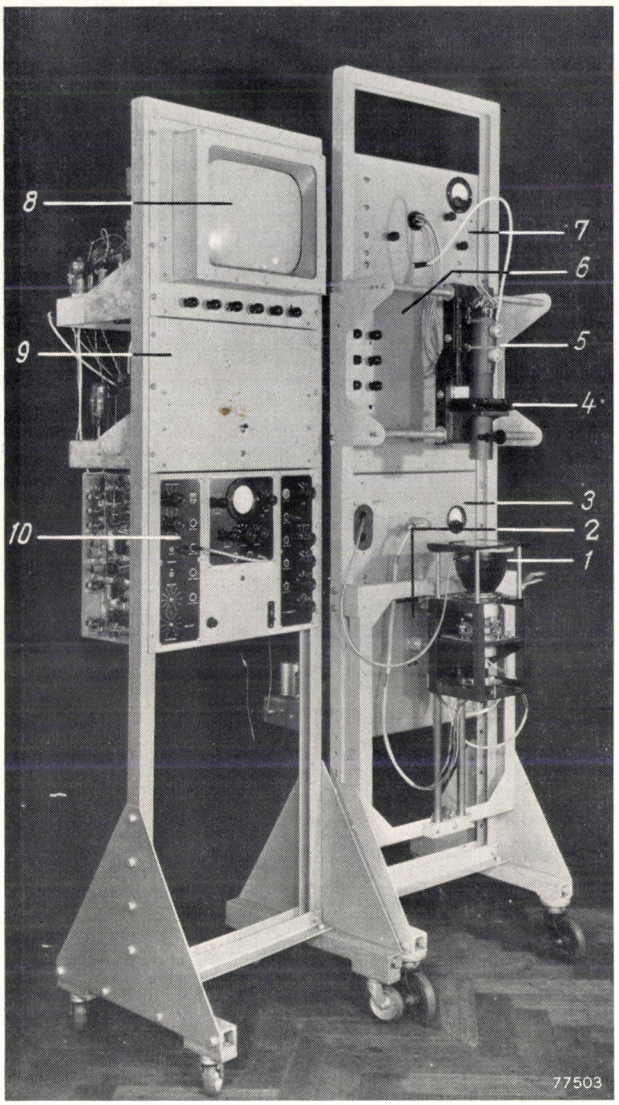


Fig. 2. The flying-spot scanner built in the laboratory. *1* the scanning tube MC 13-16; *2* and *3* panels containing power packs and deflection generators for the scanning tube; *4* the transparent object; *5* multiplier tube; *6* gamma corrector; *7* afterglow compensator; *8* the picture tube; *9* power packs for 8; 10 standard signal generator GM 2657: this supplies the necessary synchronization signals.

An important point of difference between the flying-spot scanner and a camera-tube (e.g. an iconoscope) is that in the former no accumulation effect occurs, in contrast to a camera-tube, where the potential pattern is built up by the illumination during the entire period (1/25 sec) between two consecutive scannings of the same picture-element. With the flying-spot scanner, the signal corresponding to a specific picture-element is due solely to the illumination at the moment that this element is scanned (approx. 10^{-7} sec).

An advantage of the flying-spot scanner over the iconoscope is that the former gives no spurious signals. Camera tubes such as the iconoscope give a background signal in the absence of light, which is superimposed on the picture signal when the tube is exposed [2]).

[2]) See, e.g. P. Schagen, H. Bruining and J. C. Franken, The image iconoscope, a camera tube for television, Philips tech. Rev. **13**, 119-133, 1951.

The optical system

A photographic enlarging lens with an aperture of f. 4 is used to give a sharp undistorted image of the frame on the object. Enlarging lenses are so corrected that they give the best results at small enlargements and with blue light, which predominates in the light emitted by the scanning tube of our installation (type MC 13-16). Thus they are specially suitable for our purpose. Since enlarging lenses are specially designed to give a flat image of a flat object, the scanning tube has been given a flat screen.

In order to obtain as large a photo-current as possible in the photo-electric cell, and hence as favourable a signal-to-noise ratio as possible, it would be preferable to use a lens with an even greater aperture. Photographic camera lenses, which are obtainable with a greater relative aperture, are less suitable for our purpose, however, than enlarging lenses. Thanks to the excellent properties of the phosphor in the scanning tube, very good signal-to-noise ratios are nevertheless obtained (see later).

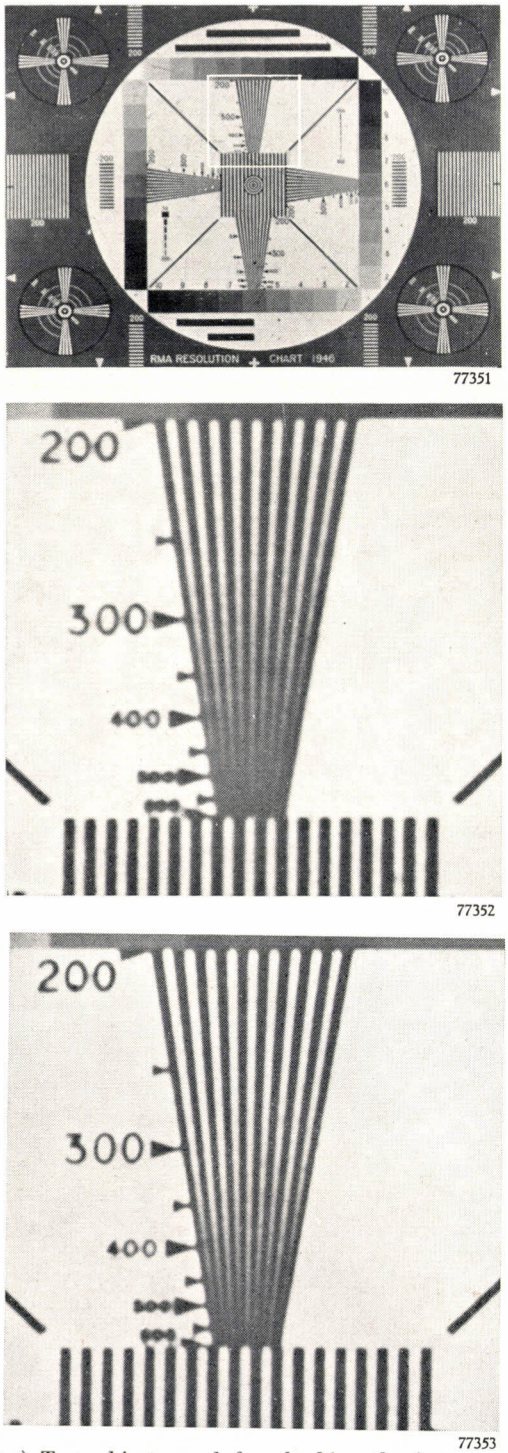
The condenser lens placed directly behind the object is larger than the latter, so that it collects all the light passing through the object. The photocell is so placed that the condenser lens forms an image of the first lens that falls completely on the photo-cathode. The light that reaches the photocell from any arbitrary point of the object is then spread over this image, so that no trouble is caused by local differences in the sensitivity of the photo-cathode.

Compensation for the afterglow of the phosphor

An important factor in the flying-spot scanner with a cathode-ray tube source, is the afterglow of the phosphor, as this influences the picture quality. As a result of the afterglow, in addition to light from the object element being scanned at a given instant, some light from object elements which have already been scanned falls on the photo-cathode. In the case of a sudden transition from dark to light (or vice versa) in the object, the image signal only gradually attains the new value, so that the transition in the reproduced picture is also gradual. This causes a certain lack of sharpness of the picture in the line direction (*fig. 3*), a phenomenon that also occurs in the use of video amplifiers with too small a frequency band.

The effect of the afterglow can be compensated in a fairly simple manner by modifications to the circuitry. Before discussing this, the influence of

the afterglow on the signal will be examined somewhat more closely [3]).



a

b

c

Fig. 3.*a*) Test object, used for checking the image quality obtained with the flying-spot scanner.
b) The part of the object outlined in white in *a*, reproduced on the picture tube without afterglow compensation.
c) As (*b*), but now with afterglow compensation. The edges of the vertical stripes are sharper than in (*b*); the difference is, however, not very striking, owing to the very short afterglow time (10^{-7} sec) of the phosphor of the scanning tube.

[3]) See also J. J. Müller, Die Korrektur des Nachleuchtens bei der Kathodenstrahlabtastung, Hochfrequenztechnik und Elektroakustik **54**, 111-115, 1939, where a similar treatment of the afterglow is given.

Quantitative consideration of the afterglow

It is known that for many phosphors the emitted light decreases as an exponential function of time. If therefore, the surface struck by the electron beam emits a light flux φ_0 at the instant t_0, the light flux emitted at a later period t will be decreased to the value

$$\varphi_t = \varphi_0 \, e^{-\frac{t-t_0}{\tau}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The time τ during which the light flux decreases to 1/e of the initial value is termed the *afterglow time.*

There is some absorption in the optical system, so that only an amount $k\varphi_t$ of the light flux is transmitted; moreover, of this amount, only a fraction $a\ (t_0)$ corresponding to the object element scanned at the instant t_0 is passed by the object and falls on the photo-electric cell.

If σ is the sensitivity of the photo-electric cell, then the light flux φ_t makes a contribution $\sigma k a(t_0)\varphi_t$ to the photo-current at the moment t.

To find the total photo-current at an instant t, consider the lines traced by the electron beam on the fluorescent screen at the constant speed V, as divided up into short lengths ξ, equal to the breadth of the spot (assumed to be square in shape). The contributions of all these lengths are then summed, including also the previous lines scanned up to the moment t. Assuming that ξ is small with respect to the distance over which the afterglow is perceptible, this summation can be carried out as a simple integration with respect to distance x along the line (*fig. 4*). If at the instant t, the electron beam is at the point x_t, then the current at this instant is

$$i'(t) = \sigma k \int_{\infty}^{x_t} a(t_0) \frac{\varphi_t}{\xi} \, dx,$$

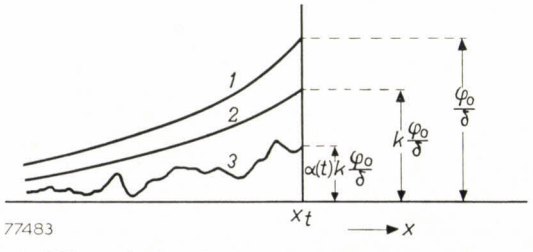


Fig. 4. Effect of afterglow of the phosphor in the scanning tube. The electron beam strikes the line at the point x_t. The light flux per unit length φ_t/ξ, which at the instant t, is transmitted by a line of the scanning tube, is given by curve *1* as a function of the position of point x. Part of the light is absorbed by the lens system, so that an amount indicated by curve *2* falls on the object. Of this amount, a varying fraction, determined by the transparency of the object, arrives at the photo-electric cell (curve *3*). The area below curve *1* represents the total light flux emitted by the scanning tube; the area under curve *3* is the total light flux which strikes the photo-electric cell.

which, disregarding the factor σ, represents the area under the curve *3* in fig. 4. (The dash above the i serves to indicate that the signal current is distorted as a result of the afterglow. The same convention will be used below for other quantities.) As $x = Vt_0$, this integration with respect to distance can be replaced by an integration with respect to time.

The result is obtained in a useful form by introducing the quantity Φ, which is the total light flux at an instant t emitted by the *whole* screen. Its value is obtained by integration of the light flux over the whole area of the screen (area under curve *1*, fig. 4): thus $\Phi = V\tau\varphi_0/\xi$. Combining this with equation 1, substituting for φ_t, and changing the variable to t, the equation for i' becomes

$$i'(t) = \frac{\sigma k \Phi}{\tau} \int_{-\infty}^{t} a(t_0)\, e^{-\frac{t-t_0}{\tau}}\, dt_0. \quad . \quad . \quad (2)$$

An *ideal* phosphor, i.e. a phosphor without afterglow, would deliver a signal:

$$i(t) = \sigma k \Phi\, a(t). \quad . \quad . \quad . \quad . \quad . \quad (3)$$

The actual phosphor (with afterglow time τ) gives rise to the same voltage across the photo-cell resistor R_m, as would an ideal phosphor if the resistor were shunted (*fig. 5*) by a capacitance C_m given by

$$R_m C_m = \tau. \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

This may be seen as follows. The undistorted current $i(t)$, (eq. 3) due to the ideal phosphor, flowing through the parallel circuit of R_m and C_m, makes a contribution $dq_0 = i(t_0)dt_0$ to the charging of the condenser in the time interval t_0, $t_0 + dt_0$. In accordance with the well-known law of discharge of a condenser, this part of the charge, at a later instant t, will have dropped to

$$dq(t) = dq_0\, e^{-\frac{t-t_0}{R_m C_m}} = i(t_0)\, e^{-\frac{t-t_0}{R_m C_m}}\, dt_0.$$

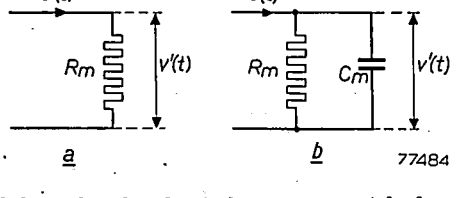


Fig. 5. With a phosphor having an exponential afterglow of afterglow time τ, the distorted signal current supplies the same voltage $v'(t)$ across a resistance R_m (fig. 5*a*) as the undistorted current $i(t)$ would supply across a parallel connection (*b*) of R_m and C_m, where $R_m C_m = \tau$.

The charge $dq(t)$ makes a contribution $dv'(t) = dq(t)/C_m$ to the voltage; thus the total voltage at the capacitor at the moment t is:

$$v'(t) = \frac{1}{C_m} \int_{-\infty}^{t} i(t_0)\, e^{-\frac{t-t_0}{R_m C_m}}\, dt_0.$$

By combining this result with (3), (4) and (2), the voltage due to the ideal phosphor across R_m-C_m is

$$v'(t) = R_m\, i'(t), \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

which is clearly the voltage across the resistor R_m alone due to the afterglow phosphor (fig. 5).

Application of the afterglow compensation

If the signal voltage $v'(t)$ obtained across R_m is fed to the control grid of a pentode (slope S) having an impedance Z_k in the cathode circuit (*fig. 6a*) then, provided Z_k meets certain requirements, there

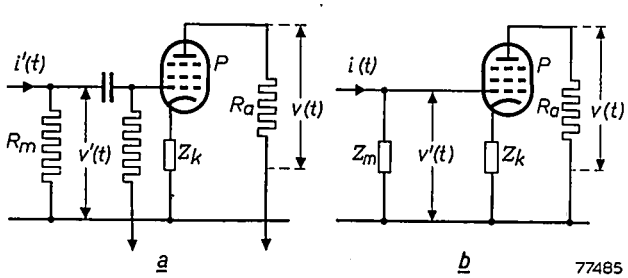


Fig. 6. The afterglow can be compensated with the circuit shown in (*a*). The photo-multiplier tube has an anode resistance R_m, across which the current $i'(t)$, distorted by the afterglow, develops a voltage $v'(t)$. Across R_a there is an undistorted voltage, provided Z_k satisfies certain requirements, which can be derived with the aid of the equivalent circuit (*b*).

will be a voltage across the anode resistor R_a from which the distortion due to afterglow has disappeared. The requirements concerning Z_k can be derived by regarding the phosphor as ideal and the resistor R_m as being replaced by an impedance Z_m (fig. 6*b*) consisting again of the parallel connection of R_m and C_m. Consider the undistorted current $i(t)$ which now flows through Z_m as being analyzed into its Fourier components. The component with the frequency $\omega/2\pi$, denoted $i(\omega)$, gives rise to a component $v(\omega)$ in the voltage across R_a. A simple calculation shows that:

$$v(\omega) = \frac{R_a}{\frac{Z_k}{Z_m}\left(1 + \frac{1}{S Z_k}\right)}\, i(\omega).$$

The impedance Z_k is chosen such that it is the same function of frequency as Z_m; this is achieved for Z_k

analogously with Z_m, by connecting in parallel a resistance R_k and a capacitance C_k, where $R_kC_k = \tau$. In the coefficient of $i(\omega)$ in the above expression, therefore, only one frequency-dependent term occurs, viz. $1/SZ_k$. However, S and Z_k can be made so large that throughout the whole frequency range this term is negligible with respect to unity. Hence the relation between $v(\omega)$ and $i(\omega)$ is independent of the frequency. The voltage $v(t)$ is then proportional to the undistorted signal current $i(t)$, i.e. the distortion caused by the afterglow has been eliminated.

Phosphors with non-exponential afterglow

The afterglow behavior of the majority of phosphors is not exactly an exponential function as assumed above; in many cases the true afterglow phenomenon can be fairly accurately described as the sum of two, three or still more of these functions. The phosphor may then be imagined composed of a corresponding number of components with afterglow times τ_1, τ_2, etc., which give rise to light fluxes of Φ_1, Φ_2, etc. The current supplied by the photo-electric cell therefore consists of the sum of the currents which would be generated by each of the components separately, and across the resistor R_m there is a voltage equal to the sum of the voltages corresponding to each of these currents. Separate RC networks must now be incorporated in Z_m (fig. 6*b*) for each component of the phosphor, of RC values equal to the afterglow periods of the corresponding phosphor component, i.e.

$$R_{m1}C_{m1} = \tau_1, \quad R_{m2}C_{m2} = \tau_2, \text{ etc.} \qquad (6a)$$

These RC networks must be connected in series so as to give a total voltage equal to the sum of the voltages corresponding to the various phosphor components. The magnitudes of the component voltages are in correct proportion when

$$\frac{R_{m1}}{\Phi_1} = \frac{R_{m2}}{\Phi_2} = \ldots = \frac{R_m}{\Phi_{tot}}, \qquad (6b)$$

where

$$\Phi_{tot} = \Phi_1 + \Phi_2 + \ldots$$

The current $i_1'(t)$, originating from the first phosphor component is found by substituting Φ_1 and τ_1 for Φ and τ in (2). In formula (3), however, Φ represents the total light flux of all the phosphor components together, thus here Φ must be replaced by Φ_{tot}. With the modified formulae (2) and (3) the expression (5) becomes

$$v_1'(t) = \frac{R_{m1}}{\Phi_1}\,\Phi_{tot}\,i'(t),$$

which is the voltage corresponding to the phosphor component with afterglow time τ_1 and light flux Φ_1.

If (6b) is satisfied, then $v_1'(t) = R_m i_1'(t)$, which is equal to the voltage created by the first phosphor component across R_m.

The requirements which the impedance Z_k (fig. 6) has to meet remain the same as in the case of a phosphor with a truly exponential afterglow.

Z_k must show the same dependence on the frequency as Z_m, and must therefore consist, like Z_m, of a number of RC networks connected in series. The RC network in Z_k that compensates for a phosphor component with an afterglow time τ_1, must have an RC value equal to τ_1 etc., whilst the resistances in Z_k must be in the ratio $\Phi_1 : \Phi_2 : \ldots$ Moreover, these resistances must be chosen so great that $1/SZ_k$ is negligible with respect to unity in the frequency range used.

Another, and in principle, more simple method of compensating the afterglow effect is to feed the distorted signal current directly into a correction impedance Z_c. With a phosphor of exponential afterglow, it is found that a series connection of a resistance R_c and self inductance L_c must be taken for Z_c, such that $L_c/R_c = \tau$. With a non-exponential afterglow, for each phosphor component such an LR network must be incorporated, and all members connected in parallel. The resistances must then be inversely proportional to the corresponding Φ values. The self-capacitances of the coils and the stray capacitances of the multiplier tube, however, cause difficulties, so that in practice the method described earlier is more satisfactory.

The signal-to-noise ratio

In spite of the fact that afterglow compensation has been achieved relatively simply, it must not be concluded that no difficulties remain. In the first instance, the signal-to-noise ratio is unfavourably influenced by the unavoidable compensation for the afterglow; furthermore, after sharp transitions from white to black, so-called "noise smears" become visible.

In order to obtain some insight into the cause of the deterioration in signal-to-noise ratio, consider a phosphor with exponential afterglow. A glance at fig. 5 shows that the distortion of the signal current results from a reduced signal at the high frequencies (influence of the capacitor C_m). The working of the compensation circuit depends upon the fact that it has to give an amplification increasing with the frequency ($\omega/2\pi$), viz. proportionately to $\sqrt{1+\omega^2\tau^2}$. Superimposed on the signal current is a noise current created by incidental fluctuations in the emission of electrons in the multiplier tube. This gives rise to the occurrence of spots in the image. The average value of the noise current is zero. In accordance with the theory of noise, the noise current contains

equal Fourier components of all frequencies up to the highest passed by the amplifier (flat noise spec-spectrum, *fig.* 7). The compensation circuit, which gives an amplification proportional to $\sqrt{1+\omega^2\tau^2}$, also amplifies the noise proportional to this factor.

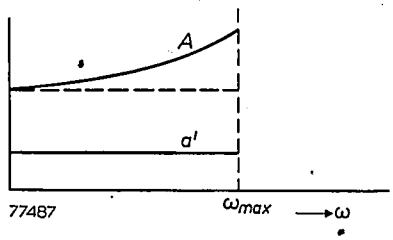


Fig. 7. Noise spectrum before and after compensation of the afterglow. The amplitude factor a' of the Fourier spectrum of the noise current is given as a function of the angular frequency by a horizontal straight line. This is valid up to the highest frequency passed by the amplifier ($\omega_{max}/2\pi$). After the frequency-dependent amplification, required for compensating the afterglow, a' becomes the amplitude factor A, which increases with the frequency. Using a phosphor without afterglow, the form of A would be given by the broken horizontal line.

The greater τ, the steeper the noise spectrum rises with frequency. If an ideal phosphor (without afterglow) were used, the amplification would be independent of the frequency, so that after amplification, the noise spectrum would again be given by a horizontal straight line (shown dotted in fig. 7), From this it can be seen that the longer the afterglow time, the more noise occurs, since the amplification must then be made more dependent on the frequency.

The signal-to-noise ratio which serves as a measure of the quality of signal, is defined as the ratio of the signal current to the r.m.s. value of the noise current. The theory of the noise of a multiplier tube confirms the plausible assumption that the signal-to-noise ratio will be more favourable for a large than for a small signal current. Since the signal current is proportional to the light flux falling on the photo-cathode, which itself is proportional to the efficiency of the phosphor, a high efficiency has a favourable influence on the signal-to-noise ratio. The detrimental effect of long afterglow time on the noise in the whole image can therefore be compensated by a high efficiency η. It may be shown that, as far this phenomenon is concerned, for not too small values of $\tau(\tau > 3 \times 10^{-7}$ sec) the quantity $\sqrt{\eta}/\tau$ can be regarded as a quality factor for the phosphor. For very small values of τ, the influence of τ is less than is expressed by this factor.

The r.m.s. value of the noise current which is superimposed on the signal current of a multiplier tube, is proportional to the square root of the signal current. The signal-to-noise ratio is consequently likewise proportional to this square root. Thus with a strong signal, the signal-to-noise ratio is higher than with a weak one; this was already pointed out in the previous paragraph and still holds good, in spite of the fact that with a strong signal, the noise itself is stronger. This fact causes the "noise smears", which are the second undesirable consequence of the afterglow.

If a light object element in the transparency (strong signal) is followed suddenly by a dark element (weak signal) then, as a result of the afterglow, the signal current in the multiplier tube decreases exponentially, instead of as a sudden drop in level(*fig.* 8*a*). The r.m.s. value of the noise current also decreases exponentially (fig. 8*b*). The compensation circuit restores the sudden jump in the signal current (fig. 8*c*), but the shape of the curve of the noise current remains unchanged (fig. 8*d*). The signal-to-noise ratio is thus particularly low for a moment, and then rises gradually to the value corresponding to the dark part of the image (fig. 8*e*). When using a phosphor with a long afterglow time, (with frequency-dependent amplification) "noise smears" are therefore seen in those dark parts of the image which have been scanned immediately after a light object element in the transparency.

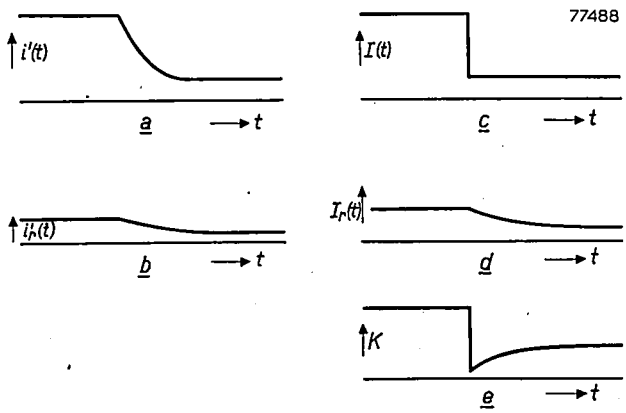


Fig. 8. The creation of "noise smears" after sharp transitions from light to dark. The signal current i' has the shape drawn in (*a*) (i.e. exponential), due to the afterglow. The r.m.s. value i_r' of the noise current, which is superimposed on i', is proportional to $\sqrt{i'}$ and has an even longer and flatter trailing edge than i' itself (*b*). As a result of the frequency-dependent amplification, i' is changed into the form I (fig. *c*), but i_r' is changed into I_r without the shape being modified (*d*). The signal-to-noise ratio $K = I/I_r$ has, therefore, a very low value directly after the transition (*e*).

The noise smears gradually decrease in the direction of the scanning movement.

In the case of a sudden jump from dark to light in the object, the effect is reversed; the signal-to-noise ratio is then for a moment relatively too high. Naturally, this is not detrimental.

With a phosphor having, e.g. an afterglow time of 10^{-5} sec, the "noise smears" have a length of approx. 1/6th of the total image width after a white-black jump. In principle the same effect occurs for a phosphor with a very short afterglow time, e.g. 10^{-7} sec; the length, however, of the noise smears is then of the same order of magnitude as the size of the image elements, and the smears are therefore invisible.

From this it can be seen that even though phosphors have the same quality factor $\sqrt{\eta}/\tau$, the one with the shorter afterglow time is to be preferred.

Closer consideration of the quality factor $\sqrt{\eta}/\tau$

As has already been mentioned, the amplitude factor $a'(\omega)$ of the spectrum of noise current in the multiplier tube is independent of the frequency (fig. 7) and proportional to the square root of the instantaneous value of the signal current i', i.e.

$$a'(\omega) \propto \sqrt{i'}.$$

After the compensation for the afterglow, $a'(\omega)$ changes into:

$$A(\omega) = a'(\omega)\sqrt{1+\omega^2\tau^2} \propto \sqrt{1+\omega^2\tau^2}\,\sqrt{i'}.$$

It can be proved, quite generally, that the r.m.s. value of a current with amplitude factor $A(\omega)$ is given by

$$\sqrt{\int_0^\infty \tfrac{1}{2}A^2(\omega)\mathrm{d}\omega}\,.$$

Since only the frequencies passed by the amplifiers are of importance to us, the upper limit of integration can, in our case, be replaced by ω_{max}. For the r.m.s. value I_r of the noise current after the compensation of the afterglow, we obtain by integration:

$$I_r \propto \sqrt{\omega_{max}\,(1+\omega_{max}^2\tau^2)}\,\sqrt{i'}.$$

For simplicity, assume that a part of the object with constant transparency is scanned. The generated signal current i is then constant. (Because now there is no difference between distorted and undistorted signal currents, i may be written for i'.) For the signal-to-noise ratio K in the current after the compensation, we have

$$K \propto \frac{\sqrt{i}}{\sqrt{\omega_{max}\,(1+\omega_{max}^2\tau^2)}}.$$

Since i is proportional to η, the efficiency of the phosphor, it follows that

$$K \propto \frac{\sqrt{\eta}}{\sqrt{\omega_{max}\,(1+\omega_{max}^2\tau^2)}}.$$

If $\omega^2_{max}\,\tau^2 \gg 1$, this reduces to

$$K \propto \frac{\sqrt{\eta}}{\tau\,\omega_{max}^{3/2}}.$$

In the case of a television system with 625 lines, $\omega_{max}/2\pi =$ 5 Mc/s, from which it follows that τ must be considerably greater than 0.6×10^{-7} sec for this approximation to be valid.

We have therefore, that K is approximately proportional to $\sqrt{\eta}/\tau$. Further, K is approximately inversely proportional to $\omega_{max}^{3/2}$, and since ω_{max} is proportional to N^2 (N represents the number of lines), K is roughly inversely proportional to N^3. With a greater number of lines, therefore, the phosphor must meet much higher requirements, or more noise must be tolerated in the image. The signal-to-noise ratio calculated for the apparatus illustrated in fig. 2, with a transparency $a = 0.6$ of the object, has the very high value of roughly 120, thanks to the very short afterglow time of the phosphor used ($<10^{-7}$ sec.). In the darkest parts of the image, where the transparency is, say, 100 times less, the signal-to-noise ratio is $\sqrt{100}$ times less, i.e. still 12.

The gamma corrector

After the compensation of the afterglow effect, the flying spot scanner gives a signal v_l that is linearly proportional to the transparency a_p of the object (p = positive), i.e.

$$v_l \propto a_p \quad \ldots\ldots\ldots \quad (7)$$

The brightness B_w on a receiver picture-tube, however, is not proportional to the supplied voltage v_w; it is approximately represented by a power function

$$B_w \propto v_w^{\gamma_w}, \quad \ldots\ldots \quad (8)$$

in which γ_w has approximately the value 2.5. It is obvious that it is desirable for the brightness at each point on the picture tube to be proportional to the transparency at the corresponding point of the object, i.e. we require that

$$B_w \propto a_p \quad \ldots\ldots\ldots \quad (9)$$

To achieve this, the signal v_l is passed through a "gamma corrector" before transmission, in which it suffers linear correction, so that the signal v_w on the picture tube is proportional to $v_l^{\gamma_c}$. One then finds, using (8) and (7),

$$B_w \propto v_l^{\gamma_c\gamma_w} \propto a_p^{\gamma_c\gamma_w}. \quad \ldots \quad (10)$$

Comparison of (10) with (9) at once shows that γ_c must be equal to $1/\gamma_w$ for (9) to be satisfied.

It is not always possible to realize exactly the proportionality expressed in (9). This fact is connected with brightness range or *contrast ratio* of the object. The contrast ratio of an object (or image) is the ratio of the highest level of brightness to the lowest level of brightness. For a transparency, we define this quantity as the ratio between the maximum and the minimum optical transmission. The maximum contrast ratio attainable in practice on the screen of a picture-tube is not greater than 100. On the one hand the maximum brightness cannot, of course, be raised above a certain value,

while on the other hand the minimum brightness is also limited, even if there is no external lighting, e.g. in a completely darkened room. Even then, light from the lighter parts of the image reaches the darker parts via reflection. If the contrast ratio C of the transparency is greater than 100, it must be reduced to, let us say, c, in the general case. The condition (9) cannot, therefore, be fulfilled, and to obtain as faithful an image as possible, it is best if the reduction in contrast is distributed over the whole brightness range in such a way that in place of (9) we obtain the relation:

$$B_w \propto a_p^{\gamma_{\text{total}}}, \quad \ldots\ldots \quad (11)$$

where

$$\gamma_{\text{total}} = \frac{\log c}{\log C}. \quad \ldots\ldots \quad (12)$$

The fact that the relationship between B_w and a_p must have precisely this form is linked with the characteristics of the human eye.

It is an empirically established fact that the impression which the eye obtains from a transition from light to dark in a picture is not determined by the variation of the brightness b with position, viz. not by $\mathrm{d}b/\mathrm{d}x$, but by the *relative* brightness variation per unit length, $(\mathrm{d}b/\mathrm{d}x)/b$. A similar brightness difference between two adjacent points therefore appears to the eye of proportionately lower contrast as the local brightness level is higher (this is the reason why, e.g. the auditorium is darkened for the showing of films).

In the reproduction of the image, for example, on the screen of a picture tube, it is further found that for a transition from light to dark, one gets exactly the same impression if $(\mathrm{d}B/\mathrm{d}x)/B = \beta\,(\mathrm{d}b/\mathrm{d}x)/b$ (B and X in the reproduction correspond with b and x in the original image; β is a constant which may differ from unity.) If $X = mx$, i.e. if m is the linear magnification of the reproduction with respect to the original image, then $\mathrm{d}B/B$ must be equal to $m\,\beta\mathrm{d}b/b$. By integration it is found from this that for a faithful reproduction ($m\beta = \gamma_{\text{total}}$):

$$B \propto b^{\gamma_{\text{total}}} \quad \ldots\ldots\ldots \quad (13)$$

If, e.g. $\gamma_{\text{total}} = 2.3$, then it is said that the "gamma" of the apparatus with which the reproduction is made, is 2.3.

If, as with the flying-spot scanner, one starts from a transparency as object, then the optical transmission a_p of this transparency assumes the part of the brightness b in the original image, and (13) is therefore equivalent to (11).

Inserting actual values in (13) of the brightnesses in the original and in the reproduced image, first for the lowest brightness levels and then for the maximum brightness levels, gives two equations, which divided by each other and equated to the contrast ratio of both images $\log c/\log C$, leads to equation (12).

The requirement (11) thus takes the place of (9). Comparison of (11) with (10) at once shows that for (11) to be satisfied, we must have

$$\gamma_{\text{total}} = \gamma_c \cdot \gamma_w. \quad \ldots\ldots \quad (14)$$

In order to be able to adjust γ_{total} to the required value (12), γ_c must be variable. If the contrast ratio of the transparency is 100 or less, then the same contrast ratio can be realized in the image, i.e. one can select γ_{total} equal to unity, as already mentioned. Since in our case $\gamma_w = 2.5$, it is required that $\gamma_c = 0.4$. If the contrast ratio in the transparency is greater than 100, then this factor must be reduced in the image; therefore γ_{total} must be smaller than unity. It is also sometimes required to increase somewhat the constrasts of transparencies that have been printed too "soft", i.e. to increase the contrast factor. In this case γ_{total} must be greater than unity. The apparatus described is based on the requirement that γ_c must be variable between 0.2 and 0.6.

Construction of the gamma corrector

The electron beam in the scanning tube is suppressed during the flyback period. During this period the signal given by the flying-spot scanner corresponds to absolute black in the transparency. This period of blackness is used to fix the absolute black level on the picture-tube too; the black level is therefore independent of the adjustment of the gamma corrector. In practice, it is necessary that the white level on the picture-tube is also independent of the adjustment of the gamma corrector: a variation of the level with γ_c would make it difficult to judge the effect of γ adjustment on image quality. To achieve this independence, the series of characteristics $v_c = A_c v_l^{\gamma_c}$ (see *fig. 9a*) must intersect each

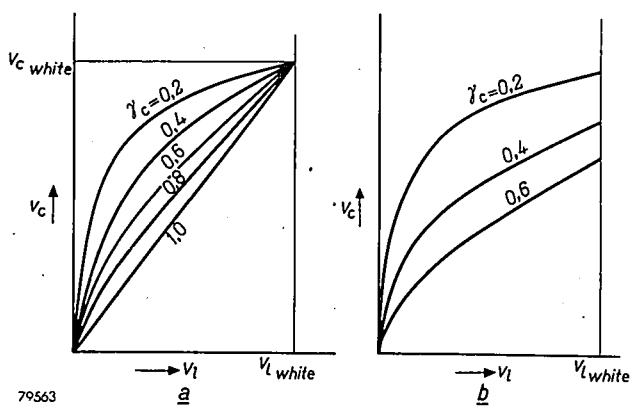


Fig. 9 *a*) Characteristics required of the gamma corrector (output voltage v_c as a function of the input voltage v_l). The curves are power functions of the form $v_c = A_c \cdot v_l^{\gamma_c}$. The input signal $v_{l\,\text{white}}$ corresponds to "white" in the image. It is required that, with variations of γ_c, the point with coordinates $v_{l\,\text{white}}$, $v_{c\,\text{white}}$, remains invariant; to achieve this, A_c must depend upon γ_c in a definite manner.
b) Example of the gamma corrector characteristics if no special measures are taken to make $v_{c\,\text{white}}$ independent of γ_c. The variation of $v_{c\,\text{white}}$ with γ_c now interferes with the judgment of the effect of a variation of γ_c.

other at the fixed point $v_{l\,\text{white}}$, $v_{c\,\text{white}}$. This implies that the proportionality constant A_c must be a definite function of γ_c, viz. $A_c = (v_{c\,\text{white}})\,(v_{l\,\text{white}})^{-\gamma_c}$. In many gamma corrector circuits, no special provisions are made in this respect, so that A_c is a different function of γ_c than that desired; the characteristics of the corrector then have the forms as shown in fig. 9*b*.

In the circuit to be described this requirement has been met in a very simple manner. The signal v_l coming from the flying-spot scanner is fed to the control grid of the pentode P (*fig. 10*), with a

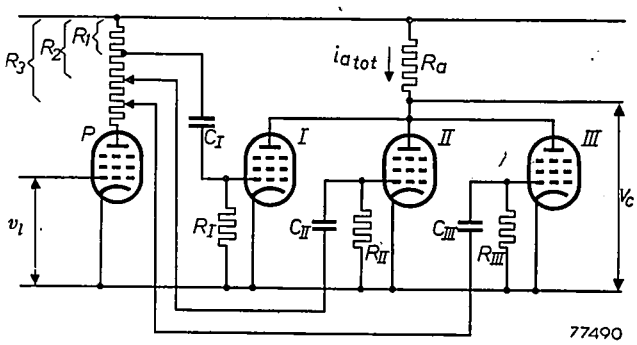


Fig. 10. Circuit diagram of the gamma corrector. The signal v_l, from the flying-spot scanner, corrected for afterglow, is applied to the grid of the pentode *P*, which, with the aid of a special circuit (not shown here), is so adjusted that the anode current has a low, fixed value when the signal $v_l = v_{l\,\text{black}}$ (flyback period). The shape of the overall characteristic ($i_{a\,\text{total}}$ as a function of v_l) can be modified by varying R_2 and R_3. This curve is constructed in fig. 11.

polarity such that the anode current of this tube increases with increasing v_l, i.e. with increasing brightness in the object. Moreover, with the aid of a special circuit (not mentioned here or in fig. 10), this tube is so adjusted that when the flying-spot scanner gives the signal corresponding to absolute black (flyback periods), the anode current of *P* has a fixed, small value. The relation between the input voltage v_l and the output voltage v_c of this circuit then has the desired form (fig. 9*a*) to a very good approximation. The value of γ_c is set with the resistors R_2 and R_3, while the constant white level is obtained by giving R_1 a suitable fixed value. Further details are given in figs 10 and *11*.

In the case of the signal $v_{l\,\text{black}}$, the anode current of *P* is a minimum, and the control grid potentials of the pentodes *I*, *II* and *III* thus have their maximum values, which are equal to the common cathode potential (if they exceeded this, grid currents will flow and charge the capacitors C_I, C_{II} and C_{III} until the control grids were at cathode potential).

In the graph in fig. 11*a*, the voltage *v* is plotted along the abscissa. The origin of the co-ordinates has been chosen at the point $v_{l\,\text{black}}$. The control grid potentials v_{gI}, v_{gII} and v_{gIII} of the three pentodes are plotted on the negative part of the ordinate. (The origin corresponds to the cathode potential.) As explained above, the grids have cathode potential when $v_l = v_{l\,\text{black}}$, and the curves for v_g as a function of v_l in the fourth quadrant thus pass through the origin. Further, all three are of the same shape as the i_a-v_g characteristic of the tube *P*, v_{gI}, v_{gII} and v_{gIII} being linearly proportional to the anode current through this tube. The slopes are in the ratio of $R_1 : R_2 : R_3$.

In the third quadrant is plotted the i_a-v_g characteristic of the pentodes *I*, *II* and *III*, which for simplicity are regarded as identical. In fig. 11*a*, R_1 has been so selected that at $v_{l\,\text{white}}$, the anode current of tube *I* is exactly zero. (It is not essential that, at $v_{l\,\text{white}}$, the anode current should be *exactly* zero.) The curves in the first quadrant have been constructed from those in the fourth quadrant with the aid of the i_a-v_g characteristic shown in the third quadrant. These curves (first quadrant) indicate the relation between v_l and the anode currents in the tube *I*, *II* and *III*. By addition of the three curves, the total anode current $i_{a\,\text{total}}$ through the common anode resistance R_a is obtained as a function of v_l. If the tappings R_2 and R_3 are adjusted, the points P_2 and P_3 shift along the $v_{l\,\text{white}}$ axis. The curves i_{aII} and i_{aIII} (in the first quadrant) then turn about the point O_1, and the portions *II* and *III* of the $i_{a\,\text{total}}$ curve turn about the points O_2 and O_3 respectively. Thus the

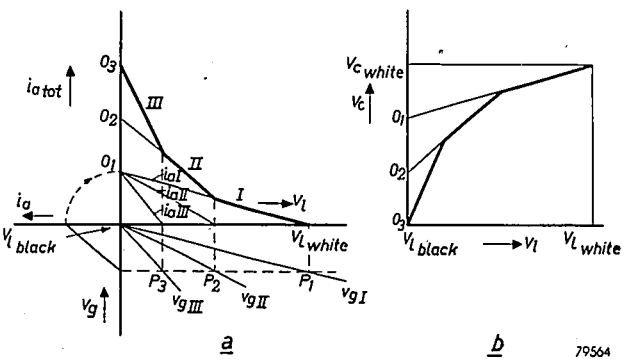


Fig. 11. Construction of the overall characteristic of the gamma corrector circuit shown in fig. 10.
a) In the fourth quadrant (lower right), the control grid voltages v_{gI}, v_{gII} and v_{gIII} of the pentodes *I*, *II* and *III* are plotted as functions of v_l. The third quadrant contains the anode-current/control-grid voltage characteristic of each of these three pentodes (assumed identical). From the above curves, the anode currents i_{aI}, i_{aII} and i_{aIII} are constructed in the first quadrant as a function of v_l. The total anode current $i_{a\,\text{total}}$ is obtained by addition.
b) The characteristic of the gamma corrector of fig. 10. As the output voltage v_c is opposite in sign to $i_{a\,\text{total}}$ ($v_c = -i_{a\,\text{total}} R_a$), the characteristic required is obtained from the $i_{a\,\text{total}}$-v_l curve by taking its mirror image in the abscissa.

shape of the $i_{a\,\mathrm{tot}}$ curve is variable, but the extremity at black (O_3) remains in its place. Similarly the extremity at white remains fixed (provided $v_{l\,\mathrm{white}}$ corresponds to a point on the fixed part I of the curve; to ensure this, the anode current, for $v_{l\,\mathrm{white}}$, should be zero, or a small value). Because the tube characteristic from which the $i_{a\,\mathrm{tot}}$ curve is derived are not exactly straight, nor the cut-off value sharply defined, the actual curve is not sharply bent, but rather, bends gradually. For the output voltage v_c of this gamma corrector, $v_c = -i_{a\,\mathrm{total}} \times R_a$. (The minus sign results from the phase reversal which occurs in the tube when the output signal is taken from the anode).

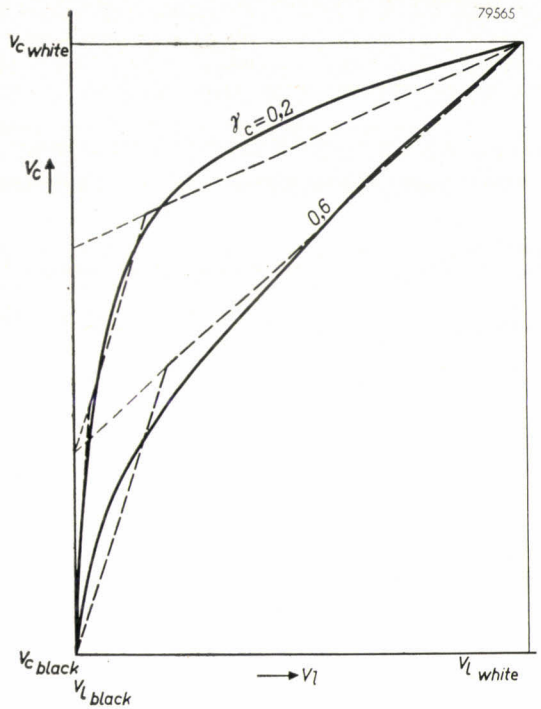


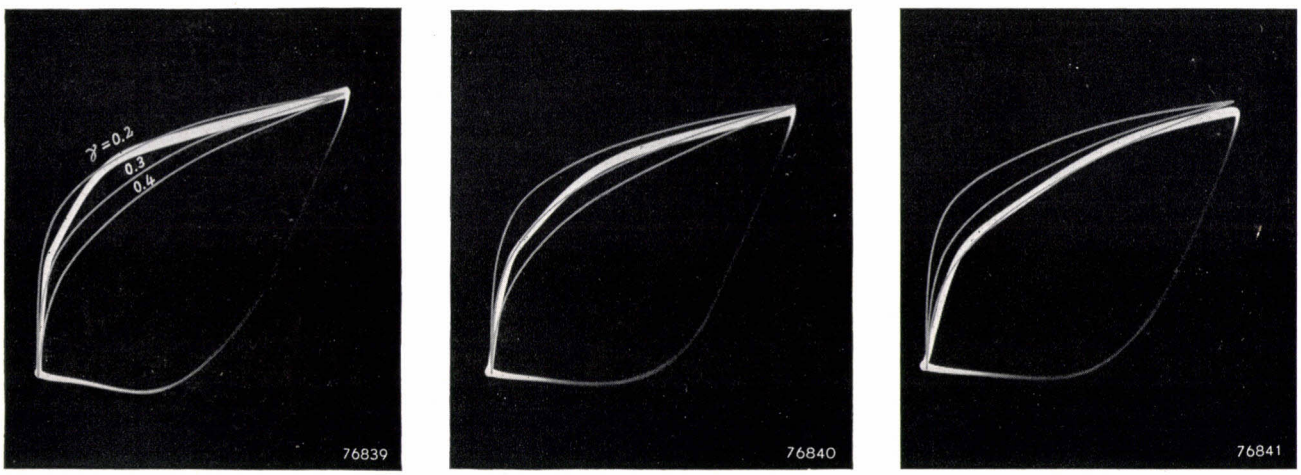
Fig. 12. Approximation to the required characteristics with the gamma corrector. The full lines show the desired characteristics, for $\gamma_c = 0.2$ and 0.6. The broken lines represent the approximations to these curves; in practice the sharp bends are rounded off, cf. fig. 13.

The shape of the curve which determines the relation between v_c and v_l is thus obtained by making a mirror image of the $i_{a\,\mathrm{total}}$ curve (fig. 11*a*) with respect to the abscissa (fig. 11*b*). *Fig. 12* shows (for two values of γ_c) how this shape can be fitted to the one required by adjusting R_2 and R_3.

The characteristic of the gamma corrector can be displayed on the screen of a cathode-ray oscilloscope. A sawtooth voltage serves as v_l, and is also used for the horizontal deflection on the oscilloscope, while the output voltage v_c controls the vertical deflection. In *fig. 13* some curves are reproduced which were obtained by this method, while *fig. 14* gives an impression of the influence which the gamma correction exercises on the picture. By connecting four or more pentodes in parallel instead of three, a still better approximation to the required characteristic is possible.

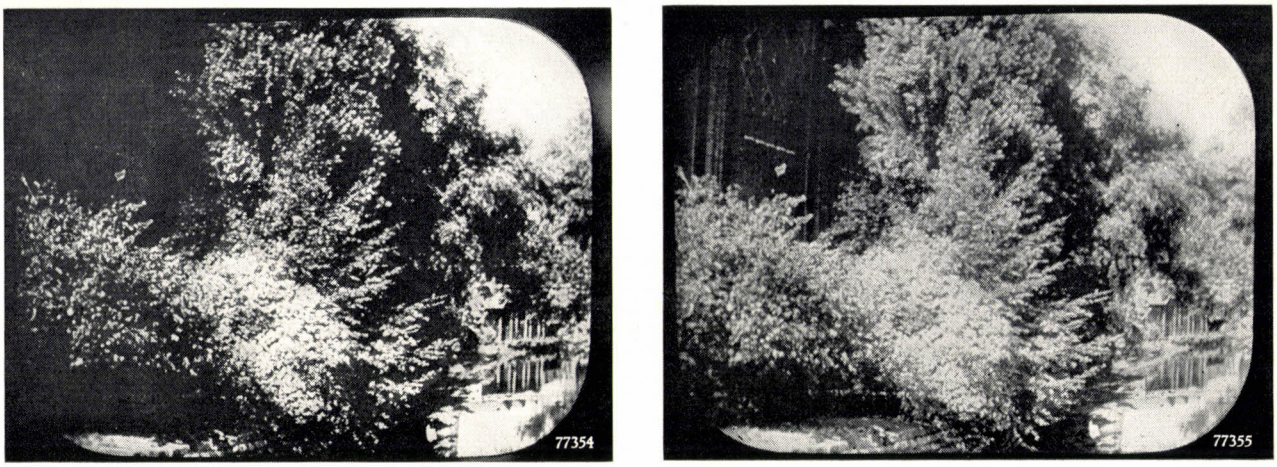
The reproduction of negatives

It is perhaps surprising that with the flying-spot scanner as described, it is also possible directly to reproduce the positive picture by scanning a negative. In order to see how this may be done, first consider again a positive object transparency. It should be noted that what is required in the final image is a faithful reproduction, with tone contrast, of the *original scene* printed on the transparency, rather than a true reproduction of the transparency itself. To make the transparency, it is usual to first make a negative of the original scene by normal



a *b* *c*

Fig. 13. Oscillograms showing the approximations of the gamma corrector to the exact characteristics, for different values of γ_c. The exact curves are drawn in thin lines for $\gamma_c = 0.2$, 0.3 and 0.4.



a *b*

Fig. 14. Image on the picture tube (*a*) without and (*b*) with gamma correction. The contrast ratio of the object was so great that the gamma of the corrector had to be set to approximately 0.3.

photographic methods. It is a characteristic of the photographic process that the transparency a_n in the negative is related to the brightness B_s of the corresponding part of the scene according to

$$a_n \propto B_s^{\gamma_n}.$$

Here, γ_n has a value between —0.4 and —0.8. That γ_n is negative follows from the fact that a high brightness in the scene corresponds to a low transparency in the negative [4]). The positive print (= the "transparency") is made photographically from the negative; the transparency a_p of the positive is related to a_n according to a similar expression,

$$a_p \propto a_n^{\gamma_p}.$$

Here again γ_p is negative.

We have already seen that a_p is transformed into the brightness B_w on the picture-tube via a power function (10). The complete transformation of B_s to B_w occurs, therefore, as a product of power functions, the total transformation being represented by:

$$B_w \propto B_s^{\gamma_{\text{total}}}, \quad \ldots\ldots \quad (15)$$

where γ_{total} is equal to the product of the exponents of the subsequent transformations. Therefore

$$\gamma_{\text{total}} = \gamma_n \cdot \gamma_p \cdot \gamma_c \cdot \gamma_w.$$

γ_{total} must be positive, because a negative γ_{total} would mean that a negative image appears on the tube (i.e. a large B_s would give rise to a small B_w, eq. (15)). γ_n and γ_p are both negative, and γ_w is positive, so that the γ_c of the gamma corrector must be made positive.

If now we use a negative transparency as object in the flying spot scanner,

$$\gamma_{\text{total}} = \gamma_n \cdot \gamma_c \cdot \gamma_w.$$

In order now to obtain a positive γ_{total}, γ_c must be negative, which means that the characteristic of the gamma corrector should be as in *fig. 15*. This form

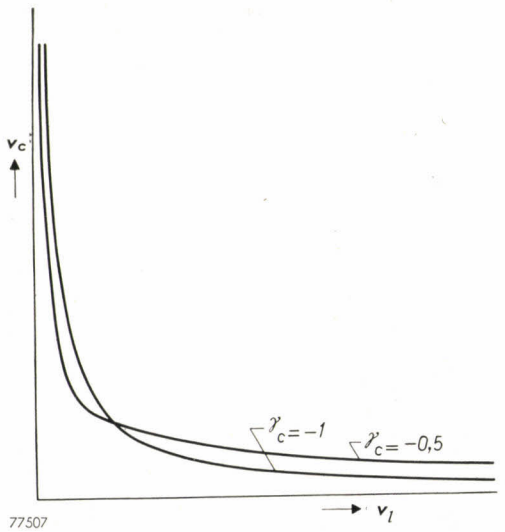


Fig. 15. The function $v_c \propto v l^{\gamma_c}$ for negative values of γ_c.

may be derived from the characteristic of the (positive object) gamma corrector (fig. 11*b*) by taking the image of the latter in a line parallel to the abscissa. The "mirroring" is achieved by adding an extra tube behind the gamma corrector to give a phase reversal. With the adjustments provided on the gamma corrector, the shape of the ideal charac-

[4]) The reader who is acquainted with photographic processes will, perhaps, take offence at the introduction of negative gammas. The negative values are the result of the use of the transparency a, in place of the blackening usual in photography.

teristic can now be attained to a sufficient approximation.

There is still a complication in that, during the flyback times, no current flows in the photo-electric cell; for negatives this corresponds to the brightest white. In order to suppress the flyback on the picture tube, image blanking pulses having an amplitude equal to the maximum "white" signal must therefore be added to the signal.

It might at first appear that the gamma correction of a negative object could also be achieved by reversal of the *input signal* of the gamma corrector instead of the output signal. Upon closer examination, however, this is found not to be the case, since reversal of the signal means that v_c changes into $v_0 - v_c$ where v_0 is a constant reference level. Reversal thus corresponds to subtraction of a constant, the gamma corrector then raising the signal to the power γ_c. Because these operations are non-commutative, the order in which they are carried out will influence the final result.

The apparatus discussed is useful whenever an image signal of very good quality is required, viz. in laboratories and in factories engaged in the manufacture of TV apparatus. A direct, practical application in the TV studio is the transmitting of lantern slides, such as the test pattern, or the meteorological map, and of drawings and photographs for the illustration of talks. An important application is undoubtedly the transmitting of films. The complications encountered in the latter will be discussed in a later issue of this Review.

Summary. For the generating of TV signals from flat transparent objects (lantern slides, transparencies, microscope slides etc.), a flying-spot scanner has been developed, in which the light source is the light spot of a cathode-ray tube (the scanning tube) specially adapted for the purpose. Images of very good quality can be obtained with this apparatus: distortion-free, very sharp and with good gradation. The influence of the unavoidable afterglow of the phosphor of the scanning tube can be compensated by an amplification which varies in a specified way with the frequency. Since, however, the signal-to-noise ratio is unfavourably influenced by this, the scanning tube is provided with a phosphor having a very short afterglow time. After compensation of the afterglow, the flying-spot scanner supplies a signal that is proportional to the transparency of the object to be reproduced. Because of the non-linear characteristic of receiver picture tubes, the signal must be corrected in order to obtain a faithful picture on the TV screen. This is done with the aid of a so-called gamma corrector which, moreover, is designed to provide additional correction in the case of a transparency which is too contrasty or too soft. A particularly simple and practical circuit has been designed for this gamma corrector. By means of a simple modification to the gamma corrector, negatives can be scanned directly and reproduced as positive images.

A CATHODE-RAY TUBE FOR FLYING-SPOT SCANNING

by A. BRIL, J. de GIER and H. A. KLASENS. 621.385.832:621.397.611.2:535.373.3

The cathode-ray tube developed for the flying-spot scanner differs in many respects from oscilloscope tubes and television picture-tubes. Of special interest are the construction of the window and the choice of the phosphor used for the fluorescent screen.

In the preceding article [1]) details are given of the flying-spot scanner, by means of which television signals may be generated from transparent flat objects such as lantern slides, films or microscope slides. A special cathode-ray tube scans a frame (raster) of constant intensity on its flat screen. An optical system projects this image on to the flat transparent object. The light passing through the object is modulated in intensity according to its transparency, and falls on a photo-electric cell. The current furnished by this cell, after amplification and the application of certain corrections, forms the television signal which is to be transmitted. At the receiving end, the signal is transformed into an image in the normal way, by means of a cathode-ray tube operating synchronously.

The particular demands made by the system on the cathode-ray tube, which produces the flying-spot light source at the transmitting end, have led to a design which is specially suitable for this purpose, and to a special choice of phosphor. Some details of the tube will now be given.

General construction of the tube

The cathode-ray tube is constructed on the same principle as cathode-ray tubes for projection television receivers [2]). In both cases a bright frame of relatively small dimensions is desired.

The optical requirements for the flying-spot frame however differ in some respects from those for the tube used for television projection. In the latter a spherical image is required to suit the Schmidt optical system used [3]). In the flying-spot scanner, the required light flux is smaller, so that a lens assembly can be used. Since the most suitable lenses are those corrected to produce a flat image from a flat object (see I), the cathode-ray tube is in this case provided with a flat window. This has the additional advantage that the phosphor coating can be applied more easily using a precipitation method, which can be important in connection with the choice of the phosphor.

Use of the lens assembly means that the dimensions of the cathode-ray tube need no longer be kept small, as is necessary in the television projection tube. A somewhat larger tube has the advantage that greater tolerances can be permitted in the dimensions and that the definition of the light spot does not need to be so high. As in the projection tube, magnetic focusing and magnetic deflection are used, with the result that the internal construction becomes simple. The focusing coil has been specially designed to ensure that the spot is sharply focused right to the corners of the frame. *Fig. 1a* shows the dimensions of the tube. The neck diameter is increased from 21.5 to 36 mm, and the window has a diameter of approx. 217 mm, which is double that of the window of the MW 6-2 projection tube. The maximum dimension (diagonal) of the frame is approx. 120 mm. The size of the raster, however, is not limited to this maximum — it may sometimes be convenient to use a smaller raster. This is accomplished simply by reducing the current in the deflection coils.

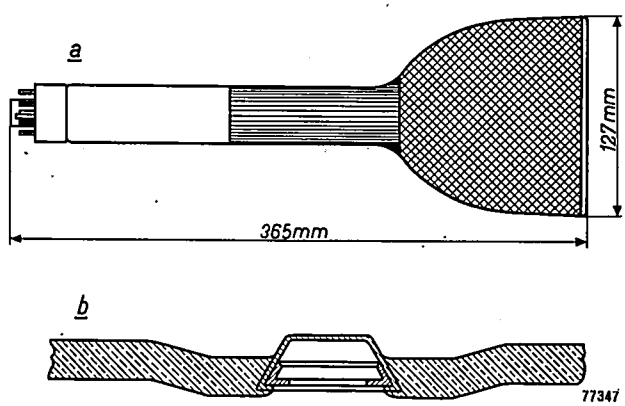


Fig. 1. *a*) Dimensions of the tube for flying-spot scanner. *b*) Lead-in terminal for the anode potential.

[1]) This issue, pp. 221-232, F. H. J. van der Poel and J. J. P. Valeton, The flying-spot scanner. Referred to in this article as I.

[2]) See J. de Gier, Philips tech. Rev. **10**, 97-104, 1948.

[3]) P. M. van Alphen and H. Rinia, Philips tech. Rev. **10**, 69-78, 1948.

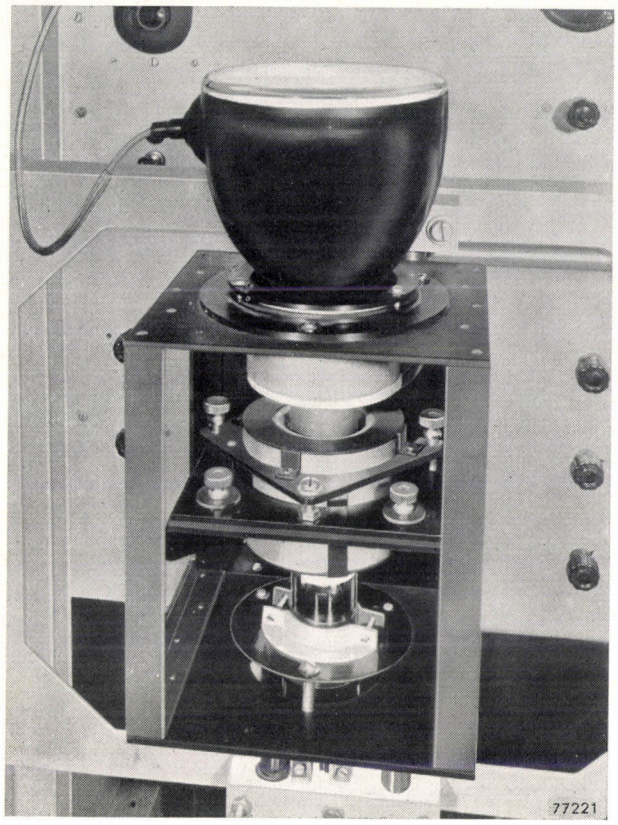


Fig. 2. The cathode-ray tube for the flying-spot scanner (type No. MC 13-16), fitted with its focusing and deflecting coils.

Fig. 2 is a photograph of the tube, showing the focusing and deflection coils.

The construction of the electron gun

The electron gun of the flying-spot scanner is constructed on the principle of the projection tube gun [4]). It is in principle a triode gun, consisting of an indirectly heated cathode, a grid and an anode. There is also a screening electrode, the so-called spark trap, placed between the grid and the anode, but this does not fundamentally affect the operation of the gun. The spark trap is actually connected with the cathode, and serves merely to prevent an undesirable discharge between the cathode and the anode, should any gas unexpectedly become liberated in the tube. The electrode diameters are about $1^1/_2$ times as large as those of the television projection tube, as a result of which spark-over is very unlikely.

The potential between cathode and anode is the same as that used in the projection tube, i.e. approximately 25 kV, and the intensity of the beam current normally used lies in the neighbourhood of 0.1 mA. This current intensity may be adjusted by varying the grid potential. In a triode, the characteristic is determined by the gradient of the potential at the cathode surface, which depends on the distance from the anode and the diameter of the grid hole. Since the screen of this tube is larger, it is also permissible to use a somewhat larger spot of light. For this reason the grid hole is somewhat enlarged (0.6 instead of 0.5 mm) and the anode distance is chosen so that the desired characteristic is obtained. In contrast to the projection tube, where the beam current intensity is modulated by the signal and thus varies continuously, the scanning tube uses a constant current intensity. The i_a-v_g-characteristic of the tube thus plays only a secondary rôle. The supply of current to the indirectly heated cathode and of various potentials to the cathode, grid and spark trap is through the base of the tube. The anode potential is supplied through an external terminal near the screen: the anode itself is connected by spring contacts to a narrow ring-shaped layer of silver, hard-baked on the inside of the neck. A thin layer of aluminium completely covering the inner side of the neck and the bulb makes contact with this ring and with the anode lead-in on the bulb. No disturbing effects occur due to optical reflection by this layer, since the screen is also coated with a thin, non-transparent aluminium layer (screen mirror).

In order to ensure a good contact between the inner layer of aluminium and the anode terminal, the former is in turn covered with a thin layer of colloidal graphite ("Aquadag"). In the projection tube, the lead-in electrode is surrounded on the outside by a small glass tube which serves to prevent flash-over between the electrode and the conducting outer wall which is earthed. In the flying-spot scanner tube, however, the outer wall is not conducting, and the lead-in electrode consists of a small metal insert, sunk into the wall of the tube (fig. 1*b*). The supply cable is provided with a spring clip which snaps firmly into this "cavity contact". The danger of flash-over has been obviated by improving the insulating properties of the outside by covering, it with an insulating, water-repellant lacquer coating. The neck of the tube is, however, covered on the outside with a conducting layer of graphite, as in the projection tube.

The window

The window of the tube must meet special requirements. As already mentioned, it is a flat window. Furthermore, since the depth of focus of the lens system is much greater than that of the Schmidt assembly of the projection tube, because of its

[4]) For details of the construction of this electron gun, see the article referred to in footnote [2]).

smaller relative aperture, special care must be taken with the outside of the window. This side is almost as sharply reproduced on the object as the phosphor screen itself. Care must therefore be taken that the outside has no spots or scratches, since these would then be visible on the television image.

Special care must also be given to the choice of the glass. The window shows a tendency to discoloration, not only under the influence of the primary electrons (depth of penetration about 8μ), but also under the influence of the electrons which originate in the glass as a result of the soft X-rays generated by the electron bombardment. The discoloration by the X-rays is reversible, i.e. it disappears in the course of time, especially at high temperatures. The discoloration by the primary electrons is not reversible, but it can be minimized by using glass of high electrical resistance, and containing no easily reducible oxides. The discoloration by X-rays can largely be prevented by using a special glass containing cerium [5]) which is used in the tube described.

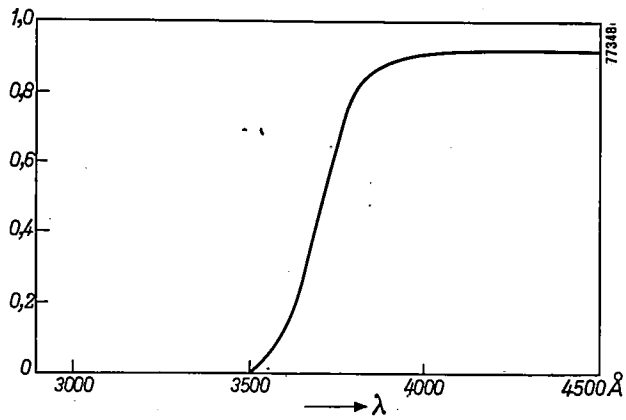


Fig. 3. The transmission of the cerium-containing glass used, as a function of the wavelength.

In *fig. 3* the transmission of this glass is shown as a function of the wavelength. From this it is evident that the transmission drops sharply at wavelengths <3900 Å. This, however, presents no difficulty since the phosphor used has its maximum emission at wavelengths >3900 Å.

Choice of the phosphor

The flying-spot scanning system makes heavy demands on the phosphor. Each point of the raster is struck by the cathode-rays for approximately 0.5×10^{-7} sec. It has been found that, in view of the occurrence of "noise smears" (see I, p. 225), the afterglow time of the phosphor should preferably not exceed 10^{-7} sec or must, at all events, be of the same order of magnitude.

In the previous article (I, p. 226) it is also mentioned that, if the afterglow time is longer than 3×10^{-7} sec, the quality of a phosphor for flying-spot scanning may be judged by the factor

$$Q = \eta/\tau^2, \quad \ldots\ldots\ldots \quad (1)$$

where η is the efficiency of the phosphor and τ is the afterglow time (in phosphors with an exponential decay in intensity, τ is the time in which the emitted light falls to a fraction $1/e$ of the initial value).

In general, phosphors can be divided into three groups [6]):

1) The phosphors in which the decay of the fluorescence is determined by the recombination of electrons and ionized centres. In this case, the decay in intensity is not exponential with time, and furthermore is dependent on the intensity of the excitation. Examples of these phosphors, which are in practice used for flying-spot scanning, are ZnO and ZnS.

Pure ZnO shows an ultra-violet emission at $\lambda = 3900$ Å; ZnO with excess zinc has a green emission with a maximum at 5050 Å. The decay of the ultra-violet emission is very rapid and for the most part takes place within 10^{-6} sec. The light emitted, however, is to a large extent absorbed by the phosphor itself. The efficiency is therefore small, viz. about 0.2%. A further consequence of the light absorption is that small variations in thickness of the phosphor layer give rise to large variations in intensity. Great demands are therefore made on the homogeneity of the layer. The efficiency of green luminescent ZnO is much greater, being in the most favourable cases about 7%, but the decay period is about ten times as long. Hence, in view of equation (1), ultra-violet fluorescent ZnO is preferable.

2) A second group of phosphors are those in which the fluorescent properties are due to certain groups of atoms, such as tungstates, molybdates, zirconates, titanates, and uranyl compounds. These phosphors show an exponential decay, but have fairly long afterglow times (10^{-4} to 10^{-6} sec), as a result of which they are less suitable for our purpose.

---

[5]) J. de Gier and J. A. M. Smelt, USA patent 2477329. The composition of this glass is: SiO_2 66%, B_2O_3 2%, Na_2O 5%, K_2O 10%, BaO 15%, CeO_2 2%. Other compositions are also possible using CeO_2.

[6]) See A. Bril and H. A. Klasens, New phosphors for flying-spot cathode-ray tubes, Philips Res. Rep. **7**, 421-431, 1952, (No. 6). See also F. A. Kröger, Applications of luminescent substances, Philips tech. Rev. **9**, 215-221, 1947, and J. H. Gisolf and W. de Groot, Philips tech. Rev. **3**, 241-247, 1938.

An interesting phosphor belonging to this group is non-activated zirconium pyrophosphate, ZrP_2O_7, with an afterglow time of 2×10^{-6} sec. This phosphor has an emission band in the ultra-violet with a maximum at $\lambda = 2850$ Å, and a fairly high efficiency (3.5 %). It is therefore eminently suited for ultra-violet microscopy with the aid of flying-spot scanning. The short wavelength of the light has here a twofold advantage: the resolving power of the microscope is increased, and many constituents of living cells, such as e.g. nucleic acids, are rendered visible by their absorption of the ultraviolet (the same constituents are almost transparent to visible light and thus remain invisible).

(3) A third group of phosphors which are important for our purpose is that in which the fluorescence is due to ions with incompletely filled shells, such as Mn^{2+}, Mn^{4+}, Cr^{3+}, Sb^{3+}, Pb^{2+}, Tl^{+}, Bi^{3+}, and ions of the rare earth metals. These phosphors also show an exponential decay in intensity. The decay period is determined partly by the nature of the ion (i.e. the nature of the electron transition which determines the emission) and partly by the crystal lattice containing the ions. An example showing the influence of the crystal lattice is ZnF_2-Mn, whose afterglow time is ten times as long as that of Zn_2SiO_4-Mn (Willemite), although in both cases the light is due to the same electron transition within the Mn^{2+} ion.

Table I. Afterglow time of a group of phosphors activated by ions with incompletely filled shells.

Activator	Basic material	Afterglow time sec
Mn^{2+}	Silicates, Phosphates, Fluorides	10^{-1} — 10^{-3}
Mn^{4+}	$Mg_2Ti\,O_4$	10^{-3}
Cr^{3+}	Al_2O_3	10^{-2}
Sn^{2+}	β-$Ca_2P_2O_7$	7×10^{-6}
	$NaCaPO_4$	6×10^{-6}
Sb^{3+}	Apatites	5×10^{-6}
	MgS	10^{-6}
Tl^{+}	NaI	$<10^{-6}$
	$Ca_3(PO_4)_2$	10^{-6}
Pb^{2+}	$BaSO_4$	10^{-6}
Bi^{3+}	β-$Ca_2P_2O_7$	2×10^{-6}
Ce^{3+}	Phosphates, Silicates	$< 4 \times 10^{-7}$

The afterglow time of a number of phosphors of group 3 is given in *table I*. Apart from the afterglow time, the efficiency and the colour (wavelength) are important when making a choice. The phosphors mentioned in the table which are activated by cerium, deserve special consideration. They have afterglow times of 10^{-6} to 10^{-7} sec, and also have a high efficiency up to 4% under excitation by cathode rays [7]).

In *fig. 4* the spectral distribution of the emission from a number of cerium phosphors is shown. These all show a maximum emission at wavelengths of less than 3800 Å. Thus they cannot be used in combination with the above-mentioned glass containing cerium, which absorbs light of these wavelengths to a large extent.

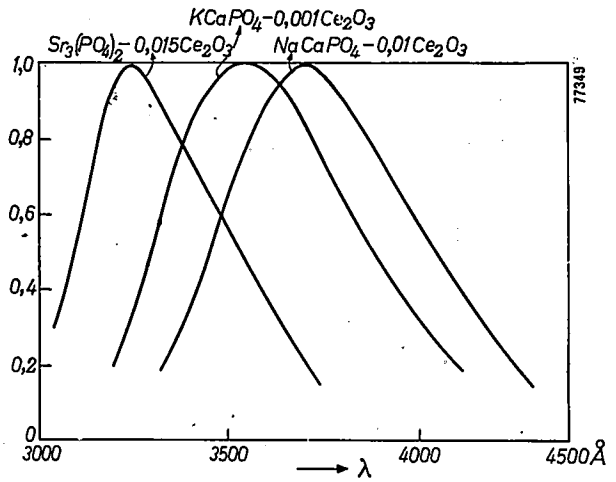


Fig. 4. Spectral distribution of the emission of a number of phosphors containing cerium.

Use is therefore made in the flying-spot tube, of a special phosphor activated by cerium, namely $2CaO.Al_2O_3.SiO_2$-0.015 Ce_2O_3. The basic material of this phosphor is also known as the mineral gehlenite. This phosphor has not only a favourable afterglow time (approximately 10^{-7} sec) and a fairly high efficiency, but also a favourable spectral distribution. The emitted light shows a maximum at approximately 4000 Å, and the spectrum extends to beyond 4500 Å, so that the light has a bluish-violet colour.

With this spectral distribution it is important to bear in mind not only the transmission of the window of the tube, but also that of the glass objective lens which projects the flying spot image onto the object slide, and that of the condenser lens which concentrates the light from the object onto the photo-electric cell. The condenser lens is made of a transparent plastic material ("Perspex") which has a better transmission at these wavelengths. The spectral sensitivity of the caesium-antimony photo-

[7]) See A. Bril and H. A. Klasens, Intrinsic efficiencies of phosphors under cathode-ray excitation, Philips Res. Rep. 7, 401-420, 1952 (No. 6); and The efficiency of fluorescence in cathode-ray tubes, Philips tech. Rev. 15, 63-72, 1953 (No. 2).

electric cell is also important. In *fig. 5* all these curves are shown with the emission curve of the phosphor itself. It is seen that the maximum of the spectral distribution of the gehlenite emission

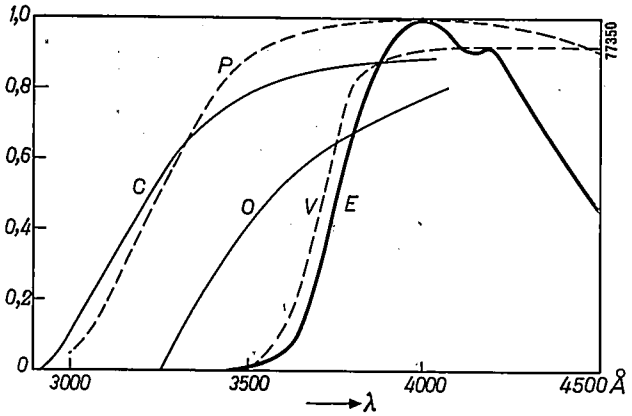


Fig. 5. Spectral distribution E of the emission of the gehlenite-phosphor. Also plotted are the spectral sensitivity P of the caesium-antimony photo-electric cell, and the transmission of the window of the cathode-ray tube (V, same curve as in fig. 3), the objective lens (O), and the "Perspex" condenser lens (C).

practically coincides with that of the spectral sensitivity of the photo-electric cell, and that the absorption by the lenses and the window of the tube is not serious.

Measurement of the afterglow time

The short afterglow times of the cerium phosphors are, in general, difficult to measure accurately. The measurement is best done in the flying-spot scanner itself. With a test slide (I, fig. 3) mounted in the apparatus, the signal generated gives a reasonable image even without any afterglow compensation (I, fig. 3*b*), although the definition of the black-white transitions along each scanning line still leaves something to be desired. Afterglow compensation is then introduced by means of the two RC-networks discussed in I (p. 224, fig. 6); the values of R and C are adjusted until the best result is obtained. In the case of the gehlenite phosphor, the values found for the products R_1C_1 and R_2C_2 amount to 10^{-7} and 2.5×10^{-6} sec, respectively, while $R_1/R_2 \approx 33$. From this it may be concluded that the greater part of the radiation has an afterglow time of approximately 10^{-7} sec. It must, however, be taken into account that in this case the condition mentioned in I is no longer satisfied, namely that the time during which the energy is supplied to the phosphor (approximately 0.5×10^{-7} sec, see above) must be small compared with the afterglow time [8]). As a result of this, the equation (2) derived in I, giving the signal current generated by the photocell, is not precisely true. The same applies to equations (4); (6a) and (6b). With the help of a special circuit with considerably greater scanning speed (and amplifier of correspondingly greater bandwidth) such that the above requirement was satisfied, somewhat modified RC-values have been found. In this way it was ascertained that the most important component has an afterglow time of only 0.3×10^{-7} sec.

In a brand new tube, the blurring cannot be eliminated entirely by means of two RC-networks (*I*, fig. 6). A third RC-network would have to be added in which the values of R and C were continually varied. Only after the tube has been in use for some hours is compensation possible with only two RC-networks, so that the third circuit can be dispensed with. In this time, the efficiency of the phosphor falls to about half its initial value and then remains practically constant. It would seem that freshly prepared phosphors contain a long afterglow component of high efficiency which is destroyed by the working of the tube.

The tubes are therefore artificially aged during manufacture, by exposing the phosphor to cathode-rays for a few hours. After this treatment, the afterglow and the efficiency undergo practically no further change and the remaining afterglow can be compensated by a circuit with two RC-networks.

[8]) In I, p. 223, this condition is formulated thus: the dimension of the light spot must be small with respect to the distance over which the phosphor has a perceptible afterglow.

Summary. The development of a special cathode-ray tube for the flying-spot television scanner is described. The particular requirements to be met are discussed, viz. the mechanical construction, the electron gun, the optical requirements and the properties of the glass, and finally the choice of the phosphor, which in this case must have a very short afterglow time (10^{-6} to 10^{-7} sec). In the tube described a gehlenite phosphor (calcium aluminium silicate) activated by cerium is used.

A REMARKABLE ETCHING OF COPPER

620.183.23

The colourful picture shown on the next page is a photograph of the polished surface of a piece of pure copper (O.F.H.C. quality), which had been immersed for 10 seconds in a silver nitrate solution [1]). The etched surface was photographed by means of normal optical microscopy with incident polarized light, polarizer and analyzer being nearly "crossed". The enlargement is well over 150 ×.

The lively colours help to make the polycrystalline structure of the metal stand out very clearly. It is not possible to account for all the details of the phenomena responsible for these colours, but they most likely originate from interference due to double refraction in a thin layer at the surface. It may be useful to devote a few explanatory notes on this phenomenon.

The preparatory process to which a metal is subjected before undergoing a microscopic examination, comprises two stages. The surface is first ground and polished until it is very smooth and flat. It is then immersed in an etching liquid; owing to the fact that the crystallites at the surface are differently orientated, they are attacked by the reagent at different rates, which produces a differential effect and reveals the structure to the eye.

Polishing, as it used to be done, was always a mechanical process. If applied to soft metals of high purity, this process results in a strongly distorted surface structure, which, moreover, contains considerable quantities of the polishing agent. As far as the commonly used, empirical etching methods are concerned, this method of polishing presented no difficulties, as in the subsequent stage of the treatment the metal was dissolved to a depth where no distortion existed. There are other etching methods, however, in which the surface metal is not "eaten" away, but partly replaced by another material whereby the surface is sometimes *raised*. Such a method, based on the principle of *electrochemical displacement*, has been applied in the case of the specimen shown here. When a soft metal like copper is etched in this way, it is essential that the metal be polished electrolytically (or chemically), not mechanically [2]). Electrolytically polished metals yield an almost ideal surface for microscopic examination, being entirely free from scratches, impurities and distortions.

In the present instance, we have used the process of electrochemical displacement, which, it is hoped, may open the possibility of introducing a method of quantitative evaluation of the etching effect and permit quantitative conclusions to be drawn from the etchings obtained (although this objective is still a long way off). The principle of electrochemical displacement is that of the replacement of a noble metal (in this case silver) in the form of ions in aqueous solution, by a less noble metal (the copper to be etched) placed in the same solution. The reaction takes place as follows:

$$Cu + 2\,Ag^{+} \rightleftarrows Cu^{++} + 2\,Ag.$$

Under appropriate conditions, the silver will then be precipitated on the copper as a thin, adhering layer. In principle, the reaction will continue until a certain equilibrium ratio has been attained between the concentrations of the silver and the copper ions in solution. It should be possible, therefore, to regulate the reaction of the etching liquid on the metal by means of a single parameter.

The interchange process will, of course, slow down, or even stop, as the copper becomes covered by the deposited silver. This raises no difficulty from the point of view of metallographic examination, however, as it is particularly the initial stage of the reaction which leads to the crystallites being made visible, when there exist the biggest differences in the thickness of the silver layer due to differing reaction rates. Incidentally it should be noted that the silver is probably precipitated in such a reactive form it reacts with the oxygen in the air at room temperature, so that actually the test specimen is covered with a thin layer of silver oxide instead of pure silver.

The above brief description outlines the manner in which the layers are produced, which give rise to the colours in the picture. It should further be pointed out that it is not always necessary to use polarized light. Sometimes the colours are visible even in ordinary light. These colours will change when the light reflected by the surface is viewed through an analyzer — because, of course, reflection causes partial polarization. However, by polarizing the incident light, the colour effect becomes much more striking.

When etching pure copper with concentrated nitric acid — which simply eats away the metal — the crystallites also appear coloured, but only if observed in polarized light. This etching method is characterized by the formation of numerous etching

---

[1]) Composition of the solution: 0.1% $AgNO_3$ by weight and 10% HNO_3 by volume (specific gravity = 1.4) in distilled water. The etching took place at room temperature.

[2]) P. A. Jacquet, Bull Soc. Chim. France, 5e série **3**, 705, 1936.

figures (small pits of particular geometrical shapes) on the crystallites. It is evident that in this case the high degree of polarization occurring as the result of multiple reflection in the pits is mainly responsible for the colourful appearance.

In order to verify the explanation proposed above — which attributes the colour effect, in the case of etching by means of silver nitrate, to the presence of a thin anisotropic, double-refracting layer, and in the case of deep-etching with nitric acid, to the special topography of the surface, — both specimens were vacuum-coated with a very thin layer of silver[3]). The layer deposited was so thin that it resembled the original shape of the surface in every detail, any specific reflection phenomena thus being preserved; on the other hand, it was thick enough to fully

[3]) E. C. W. Perry and J. M. Lack, Nature **167**, 479, 1951.

absorb any light not reflected on the outside, the optical effects of an underlying anisotropic layer thus being suppressed completely. As was to be expected, this treatment obliterated the colour effect in the case of the first specimen etched by electrochemical displacement, whereas no noticeable change was perceived in respect of the second specimen (etched with nitric acid).

For a more complete explanation of the colour phenomena, more detailed observation would be necessary on the composition, structure and thickness of the layers produced by the process of electrochemical displacement.

In order to avoid misunderstanding, it should further be pointed out that a distinction must be made between metals having a cubic crystal structure, as e.g. copper, silver, iron etc., and those of non-cubic structure, such as bismuth, antimony, uranium, etc. In the case of the latter, light striking even a smooth and clean surface at right angles is reflected anisotropically from the differently orientated crystallites, so that a colourful picture of the structure can be obtained without any previous etching [4]).

J. J. de JONG.

[4]) A fine example of this is provided by the pictures of electrolytically polished bismuth, shown in the article by B. W. Mott in Endeavour **12**, 154-161, 1953 (fig. 2a-d), which also describes in detail the procedure followed.

A SINUSOIDAL *RC*-OSCILLATOR FOR MEASUREMENTS IN THE FREQUENCY RANGE 20-250,000 c/s

by J. D. VEEGENS and E. PRADO.

621.396.615.1.029.4

Oscillators whose frequency is determined by resistance-capacity networks (RC-oscillators), usually generate non-sinusoidal oscillations, e.g. saw-tooth or square waveforms. It is also possible, however, to design RC-oscillators which generate practically sinusoidal oscillations. For a wide frequency coverage, they even offer certain advantages over LC-oscillators.

The RC-oscillator dealt with in this article has been designed for measuring frequency characteristics. At frequencies between 20 c/s and 250 kc/s it is capable of supplying a sinusoidal voltage that is not only practically undistorted, but also of very constant amplitude.

Electrical networks, such as amplifiers, radio receivers, filters, cables, etc. can be considered as 4-pole networks or quadripoles. The output voltage of a quadripole, in general, differs from its input voltage; apart from amplification or attenuation, distortion or demodulation may have occurred, to an extent dependent on the frequency and the amplitude of the input voltage. For the examination of quadripoles, an oscillator is required which supplies a voltage whose frequency and amplitude are adjustable to known values. The distortion present in this signal, which can never be avoided completely, should also be known.

In this article, a measuring oscillator is described for the frequency range 20-250 000 c/s, a range that includes the audio frequencies and a considerable part of the carrier-telephony range.

Alternating voltages can be generated by resonant circuit valve oscillators. The frequency of this type of oscillator is determined by a tuned circuit comprising a self-inductance and a capacitor; in most cases the coil has a fixed self-inductance, and the capacitor is of the variable type. A direct application of this principle to the range of audible or even lower frequencies, however, involves practical difficulties and drawbacks:

1) The frequency range covered by a circuit with a constant self-inductance and a tuning capacitor of the usual maximum capacitance (e.g. 1000 pF) is rather small, the ratio of its limits being about 1 : 3. The total frequency range has thus to be divided into a large number of ranges, which necessitates repeated switching.

2) The coils required for generating low frequencies must have a very high self-inductance. Not only are these coils large and rather expensive, but they present difficulties with regard to the stability of their self-inductance.

In order to avoid these difficulties, many audio-frequency generators employ the heterodyne principle. Here the required frequency is obtained by mixing the voltages from two oscillators, having different, fairly high frequencies [1]). With this

[1]) See, e.g. L. Blok, A tone generator, Philips tech. Rev. **5**, 263-269 1940; also J. de Jong, Maintenance measurements on carrier telephony equipment, Philips tech. Rev. **8**, 249-256, 1946.

system, a large frequency range can be covered without switching, and no coils of high self-inductance are necessary. It has, however, disadvantages of another nature. The main objection is that at low frequencies the relative frequency stability is, in the nature of things, small, and can be kept within reasonable limits only by setting extremely high standards for the tuning elements of both oscillators (the temperature coefficient of the self-inductance of the coils and of the capacitance of the capacitors must be very small and very constant). A further drawback of heterodyne oscillators is the rather complicated circuit (two oscillators, a mixing valve, an amplifier and the means for suppressing undesired frequencies).

For these several reasons, another alternative has been adopted in the design of the present apparatus.

The *RC*-oscillator

Valve oscillators employing resistance-capacity networks in place of a tuned circuit of inductances and capacitors, have been known for some time [2]). A circuit which was to prove of great practical importance in the development of *RC*-generators of this type, has been dealt with in publications dating back to 1938 and 1939, and it is this type of circuit [3]) that is applied in the measuring oscillator described here.

Various other circuits are possible; for example, those in which a T-filter or double T-filter is used in place of the *RC*-filter to be discussed later, and also those in which the *RC*-network is incorporated in the feedback circuit.

The circuit represented in *fig. 1* shows a resistance-coupled amplifier (A) consisting of two stages, which together effect a phase shift of 360° between input and output voltage. A positive feedback is obtained by connecting the amplifier output to the input via a filter consisting of the resistors R_1 and R_2 and the capacitors C_1 and C_2. The system will oscillate at that frequency for which the loop gain is exactly 1. As will be demonstrated later, this is the case at a frequency f_0, given by:

$$f_0 = \frac{1}{2\pi\sqrt{R_1 R_2 C_1 C_2}},$$

whilst, if suitable values are chosen for the filter elements (viz. $R_1 = R_2$ and $C_1 = C_2$), an amplification factor of only 3 is sufficient for the amplifier.

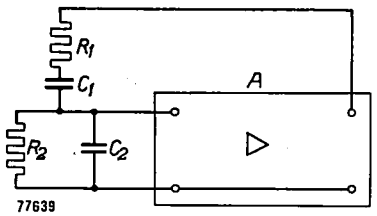


Fig. 1. Basic circuit of an *RC*-oscillator. The amplifier A obtains a positive feedback via a filter consisting of the resistors R_1 and R_2 and the capacitors C_1 and C_2. A condition for oscillation is that the phase shift between input and output voltage of the amplifier amounts to 360°.

By simple means, this type of oscillator can be given the following favourable properties:

a) The frequency, even at low values, can be kept constant within a few per cent, or better; the stability depends mainly on the resistors and capacitors involved, and is influenced by the valves and the other circuit elements to a far smaller extent.

b) The output voltage can be kept constant throughout a very large frequency range, e.g. with a frequency ratio of $1 : 10^5$.

c) Distortion can be kept very small, e.g. 0.5 - 3 parts per thousand.

d) By means of fixed resistors and normal variable capacitors, a large frequency range can be obtained, e.g. up to 1 : 10 (although it may be preferable to divide the overall frequency coverage into somewhat smaller ranges, to avoid crowding of the scale).

The frequency-determining element

If an alternating voltage v_1, having an angular frequency ω, is applied between terminals *1* and *3* of the filter shown in *fig. 2*, then a voltage v_2 is created between terminals *2* and *3*, which is related to v_1 as:

$$a = \frac{v_2}{v_1} = \frac{1}{1 + \frac{R_1}{R_2} + \frac{C_2}{C_1} + j\left(\omega R_1 C_2 - \frac{1}{\omega R_2 C_1}\right)}.$$

This relationship can be simplified for the case where $R_1 = R_2$ ($= R$) and $C_1 = C_2$ ($= C$). In this case $f_0 = 1/2\pi RC$, and writing $f = \omega/2\pi$, we have:

$$a = \frac{1}{3 + j\left(\frac{f}{f_0} - \frac{f_0}{f}\right)}.$$

[2]) J. van der Mark and B. van der Pol, The production of sinusoidal oscillations with a time period determined by a relaxation time, Physica **1**, 437-448, 1934.

[3]) H.H. Scott, A new type of selective circuit and some applications, Proc. Inst. Rad. Engrs. **26**, 226-235, 1938. F. E. Terman and co-workers, Some applications of negative feedback, Proc. Inst. Rad. Engrs. **27**, 649-655, 1939.

For this case, *fig. 3* shows the absolute value of α and the corresponding phase angle φ as functions of the ratio f/f_0. The quantity $|\alpha|$ has a maximum value of $^1/_3$ when $f = f_0$, and at that value $\varphi = 0$; hence, v_2 is in phase with v_1. It will be clear that the

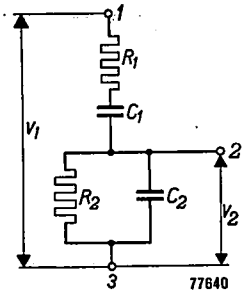


Fig. 2. The filter R_1-C_1-R_2-C_2- of fig. 1, with the input voltage v_1 and the output voltage v_2.

oscillator circuit of fig. 1 will be able to oscillate at this frequency f_0, if the amplification factor is 3, so that for $|\alpha| = ^1/_3$ the loop gain becomes 1.

There are, however, various factors causing f to deviate slightly from f_0, to an extent which differs for the various ranges of the total frequency range. If these ranges have been so chosen that the corresponding frequencies in each range are in a ratio of exactly 10 or 100 to those of the next, then one frequency scale will be possible, provided that the deviation between f and f_0 is small. It is therefore important to find the causes of this deviation and, so far as possible, to remove them.

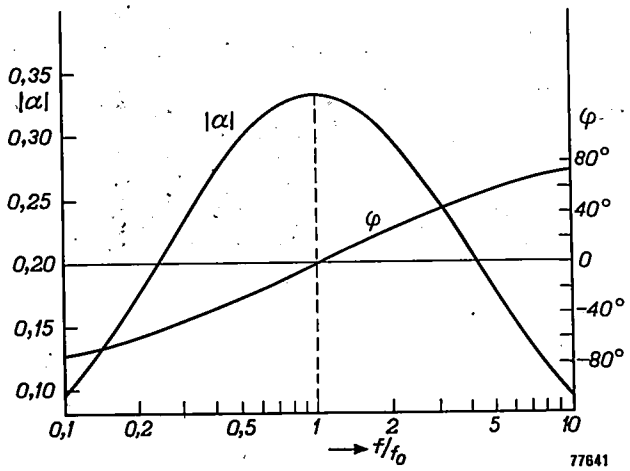


Fig. 3. Graph representing the properties of the filter of fig. 2, in which $R = R_1 = R_2$ and $C = C_1 = C_2$. The absolute value $|\alpha|$ of the voltage ratio v_2/v_1 and the phase shift φ between v_2 and v_1 are plotted vertically. The ratio f/f_0 is plotted horizontally. f is the frequency of v_1 and v_2, and $f_0 = 1/2\pi RC$.

1) If the amplifier shows a phase shift deviating from 360° at the frequency to be generated, this is automatically compensated by the fact that f deviates from f_0 in such a way that the *total* phase shift in amplifier plus filter will again amount to exactly 360°. It can be adduced from the φ-curve in fig. 3 that for a phase difference of 1° the frequency will deviate from f_0 by 3%.

2) Stray capacitances, too, cause a slight discrepancy between f and f_0. In *fig. 4*, three stray capacitances C', C'' and C''' are represented. C' (the capacitance between P and earth) is by far the most important of the three; it causes a frequency deviation $\Delta f = 0.87\, f_0 C'/C$. Since C' has different values according to the particular non-variable resistors R, switched into the circuit, f will have different values according to the frequency range.

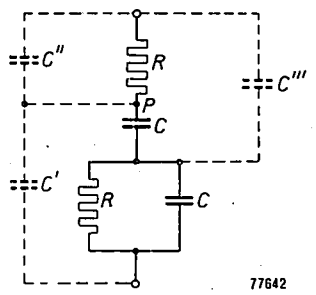


Fig. 4. The filter of fig. 2 with its three stray capacitances C', C'' and C''', of which C' is the most important.

A similar effect may be caused by faulty insulation between the various points of the filter. Here too, we find that point P (fig. 4) is particularly sensitive. Great care should thus be taken that this point has a constant, low stray capacitance and is well insulated.

3) The internal resistance R_i of the amplifier as seen from the output terminals has some influence on the frequency f, in accordance with the relation: $f = f_0/\sqrt{1+(R_i/R)}$. This effect can be compensated in the manner indicated in *fig. 5*, by incorporating a resistance $\frac{1}{2}R_i$ between points *3* and *4*. In the *RC*-oscillator described here, however, another procedure has been adopted: R_i has been kept so

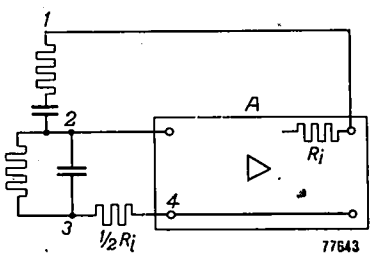


Fig. 5. The influence of the internal resistance R_i can be eliminated by incorporating a resistor with a value $\frac{1}{2}R_i$ between points *3* and *4*.

small that its influence (even at the smallest values of R) can be considered negligible. This point will be dealt with later in this article.

4) A further discrepancy between f and f_0 is due to the fact that the time constant of the resistors R is not exactly zero (the time constant is the ratio between the stray self-inductance and the resistance); thus resistors with the smallest possible time constant should be used.

The filter elements

If the *RC*-oscillator is to produce a variable frequency, then the filter elements have to be made variable. Either the resistors R can be made continuously variable and the capacitors C adjustable in steps, or the capacitors can be made continuously variable and the resistors adjustable in steps. As already mentioned, the latter system has been chosen, for the following practical reasons:

a) Identical variable capacitors — ganged in pairs, such as are required here — are components commonly used in radio manufacture. Their capacitance can be made stable (within 1%) and continuously variable, with a low temperature coefficient (below 100×10^{-6} per °C). In addition, they are robustly constructed.

b) In each frequency range covered by a fixed value of R, the impedance of the filter (between points *1* and *3*, fig. 2) has a constant value, viz. $3R/(1+\mathrm{j})$, so that the output of the amplifier is connected to a constant load.

The use of variable capacitors has, on the other hand, the disadvantage that their capacitance is rather low, so that for generating low frequencies high resistances are required: e.g. for $C = 500$ pF and $f = 20$ c/s, R has to be approx. 16 MΩ. This means that carbon resistors have to be used, which cannot meet such high stability standards as wire resistors. Carbon resistors can be assumed to have a stability of approx. 1% and a temperature coefficient of approx. -5×10^{-4} per °C.

Features of the measuring oscillator, type GM 2317

The measuring oscillator, type GM 2317, shown in *fig. 6*, has been designed in accordance with the above-mentioned principles. The resistors of the *RC*-filter are mounted inside a drum which can be rotated with respect to the fixed terminals of the capacitors. In order to avoid a crowded scale, each frequency range has been confined within limits 1 : 5. The scale division is practically logarithmic, i.e. all frequencies can be read to about the same relative accuracy. For insulation, use has been made throughout of a ceramic material coated with a water-repellent lacquer, so that no water-film can be formed. This is particularly important at low frequencies (i.e. at high R values,

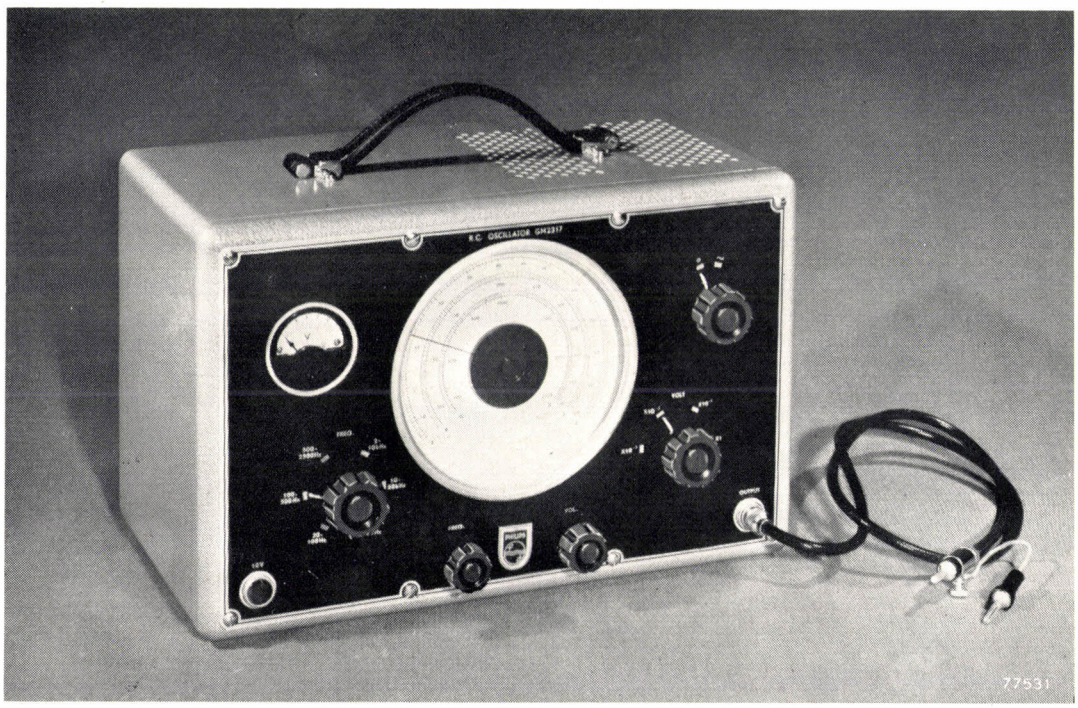


Fig. 6. Photograph of the measuring oscillator, type GM 2317.

where a very good insulation is essential to keep the frequency sufficiently stable. For the same reason the variable capacitors are contained inside a metal housing, to avoid the penetration of dust, which would lower the insulation resistance in humid conditions. The container moreover serves as a screening against the stray field of the supply transformer, which might otherwise induce undesirable hum voltages.

The amplifier

The main requirement to be met by the amplifier is that the phase shift between input and output voltage should deviate very little from 360° throughout the entire frequency range. It is a well-known fact that good results in this respect can be effected by a system of negative feedback [4]). The parameters of this feedback circuit should be such that whatever the frequency, positive feedback never occurs, as this would incur the risk of "squegging". Further advantages of negative feedback are that the distortion of the amplifier is reduced, as well as the influence of supply voltage fluctuations and that of variations in the slope of the valves.

a factor of 3 (provided that the conditions $R_1 = R_2$ and $C_1 = C_2$ are accurately satisfied); hence a very large negative feedback can be applied. The amplification without feedback would amount to approximately 2000 ×.

It may be noted that the *RC*-filter, together with the resistors r_2 and r_3, may be considered as a Wien-Robinson bridge.

The two-stage amplifier is followed by a cathode-follower stage, which has the purpose of keeping the internal resistance R_i, seen from the output terminals, at a low value (approximately 100 Ω). Consequently, the influence of R_i on the frequency is very small (R having a minimum value of about 6000 Ω).

Amplitude limiting

The method of limiting the amplitude of the oscillation is of great importance in oscillator design, since it affects:

1) the distortion, which has to be kept as small as possible; this is particularly important with *RC*-oscillators because of the poor selectivity of the *RC*-filter;

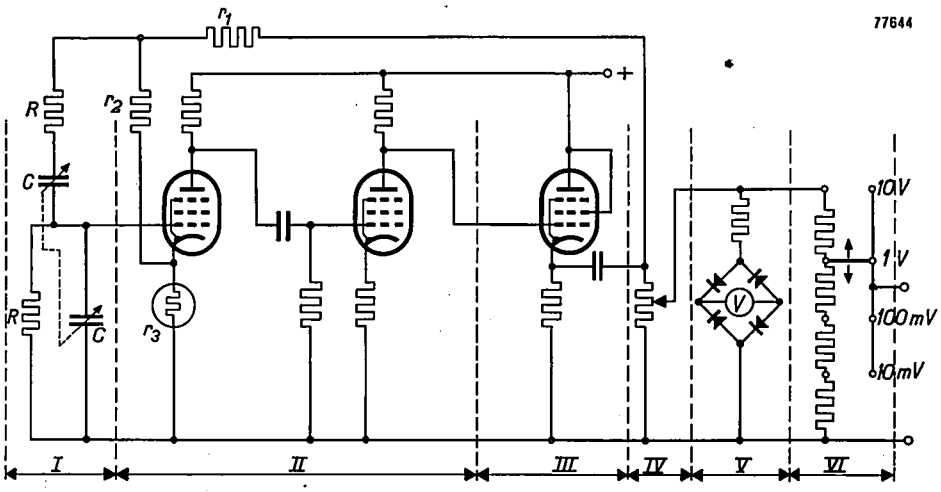


Fig. 7. Circuit diagram of the measuring oscillator GM 2317 (somewhat simplified). Section *I* contains the *RC*-filter with two resistors *R*, adjustable in steps, and two variable capacitors *C*. Section *II* represents the two-stage amplifier, with a large negative feedback via the resistors r_1, r_2 and r_3, the last one being a filament lamp (type 8099). Section *III* is the cathode-follower output stage. Section *IV* is a continuously adjustable attenuator, section *V* a voltmeter and section *VI* an attenuator adjustable in calibrated steps.

Fig. 7 shows a somewhat simplified circuit diagram of the oscillator. The negative feedback is effected via the resistors r_1, r_2 and r_3. The last, r_3, is a filament lamp; the reason for this will be discussed later. As mentioned before, the necessary amplification of the two stages need not exceed

2) the constancy of the amplitude throughout the entire frequency range; it is desirable that the amplitude should be as much as possible independent of frequency, supply voltages and valve characteristics.

A suitable method for limiting the amplitude is one which employs a thermal device, comprising a non-linear resistance whose value varies with

[4]) See e.g. B. D. H. Tellegen, Inverse feedback, Philips tech. Rev. 2, 289-294, 1937.

the mean value of the output voltage. Such a thermal control may be effected either by a resistor with negative temperature coefficient (N.T.C.-resistor), or by a resistor with positive temperature coefficient, e.g. a type of filament lamp. The latter

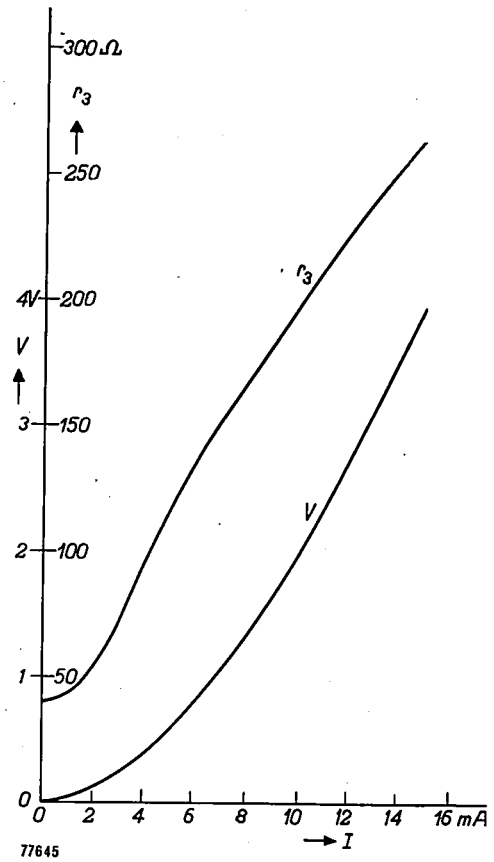


Fig. 8. Characteristics of the regulator tube, type 8099. V is the voltage and r_3 is the resistance, plotted as functions of the current I.

is used in the oscillator GM 2317 (r_3, fig. 7), since it has been found that in this case a less distorted and more reproducible signal can be obtained. An increase of the output voltage increases the current through the tube and thus raises the temperature of the filament, and hence its resistance; consequently, the negative feedback is increased, thus counteracting the initial rise of the output voltage. An important factor here is the thermal inertia of the resistor. This should not be too great, as then a state of equilibrium would be reached too slowly, but on the other hand, nor should it be too small, as this may give rise to distortion (distortion will occur if the resistance shows discernible fluctuations within one cycle of the generated alternating voltage).

The working temperature of the filament should preferably be fairly high, since this prevents changes in the ambient temperature exerting a disturbing influence on the output voltage. The construction of such a lamp-type resistor should also be such as to suppress any tendency towards microphony. These requirements are met by the lamp-type resistor (regulator tube), type 8099. *Fig. 8* shows the voltage and the resistance of this tube, plotted as functions of the current. The properties of this tube ensure certain essential features of the oscillator: the output voltage (10 V) is constant throughout the entire frequency range within a few %, distortion throughout the greater part of the frequency range is less than 0.3 %, and mains voltage fluctuations and alternations in the slope of the amplifier tubes have only a very slight influence on the output voltage.

Voltmeter and attenuators

The output voltage can be continuously adjusted between 0 and 10 V by means of a potentiometer (section *IV*, fig. 7) and read from the voltmeter *V*, consisting of a series resistor and a moving-coil micro-ammeter connected across a rectifier bridge (four germanium diodes). The voltage can be attenuated to $^1/_{10}$, $^1/_{100}$ or $^1/_{1000}$ by means of a second voltage divider (section *VI*, fig. 7).

Summary. An *RC*-oscillator (type GM 2317) is described, which acts as a source of sinusoidal voltage for measurements on 4-pole networks of all kinds. The oscillator consists of a two-stage amplifier followed by a cathode-follower stage, and a network effecting the positive feedback. This network is a filter composed of two resistors of equal value and two capacitors of equal value, and is the frequency-determining network. The resistors are adjustable in steps. For each resistance value the frequency is adjustable between limits having a ratio of 1 : 5, by means of the continuously variable capacitors. The total range covers all frequencies between 20 and 250,000 c/s. A large negative feedback is applied in the amplifier, with the result that the output is practically undistorted and only very slightly influenced by fluctuations of the supply voltages or by variations of the valve properties. The cathode resistor of the first amplifying valve is a special filament lamp (type 8099), the resistance of which increases with the current in such a way that throughout the whole frequency range the output voltage (10 V) is kept constant within a few %. Distortion, over the greater part of the frequency range, is less than 0.3%. The instrument is provided with a continuous attenuator, a voltmeter, and an attenuator adjustable in calibrated steps.

LATTICE IMPERFECTIONS AND PLASTIC DEFORMATION IN METALS

I. NATURE AND CHARACTERISTICS OF LATTICE IMPERFECTIONS, NOTABLY DISLOCATIONS.

by H. G. van BUEREN

548.4:539.374:669

Not more than 30 or 40 years ago our knowledge of the physical properties of metals was based almost entirely on experience. After the first world war a change came, brought about by a gradual expansion in the study of the physics of metals, the object of which was to justify theoretically the observed characteristics. Recently, much progress has been made in the field of plastic properties of metals. The new conception, that the plastic deformation of materials is intimately connected with the occurrence and concentration of imperfections in the regular structure of the atoms, that is, of lattice imperfections, has been found to be very fruitful. It can be applied not only to metals, but also to other materials.

Introduction

At the beginning of this century it was still the general belief that the more important physical properties of crystalline substances could be explained exclusively in terms of the periodic arrangement of the component atoms. It was generally thought that any defects in that structure (irregularities in the crystal lattice) had but little bearing on the characteristics, and these irregularities were accordingly disregarded; the crystal lattices were held to be perfect.

Although this standpoint originally met with considerable success, e.g. in the explanation of X-ray diffraction patterns and in the theories of cohesion between the atoms in crystals, it was soon found that there were many phenomena which would definitely not answer to theoretical considerations on that basis. Particularly in the study of transport phenomena such as the conduction of heat and electricity and diffusion, insurmountable difficulties were encountered.

The first and most important step towards a better understanding of these phenomena can be said to have been the recognition of thermal lattice vibrations. The atoms at the lattice points of a crystal vibrate, practically as harmonic oscillators, about their position of equilibrium, and the energy with which they do this, and hence the amplitudes, increase considerably with the temperature.

Although this conception led to important additions to the theory of certain transport phenomena, many effects were still without any satisfactory explanation, in particular, those which relate to the mechanical properties of crystals. Ultimately, therefore, it was found necessary to take into consideration other imperfections in the crystal structure. To enter into a general discussion on the influences of lattice imperfections upon the physical and chemical properties of crystals, would not be practicable within the scope of this article. We shall therefore limit our discussion to two articles dealing with some of the phenomena connected with the plastic deformation of metals. For a review of the many other domains in which lattice defects are of interest, the reader may refer to a recent article in this Review by G. W. Rathenau [1]), and to the comprehensive survey by F. Seitz [2]).

Some phenomena related to the plastic deformation of metals

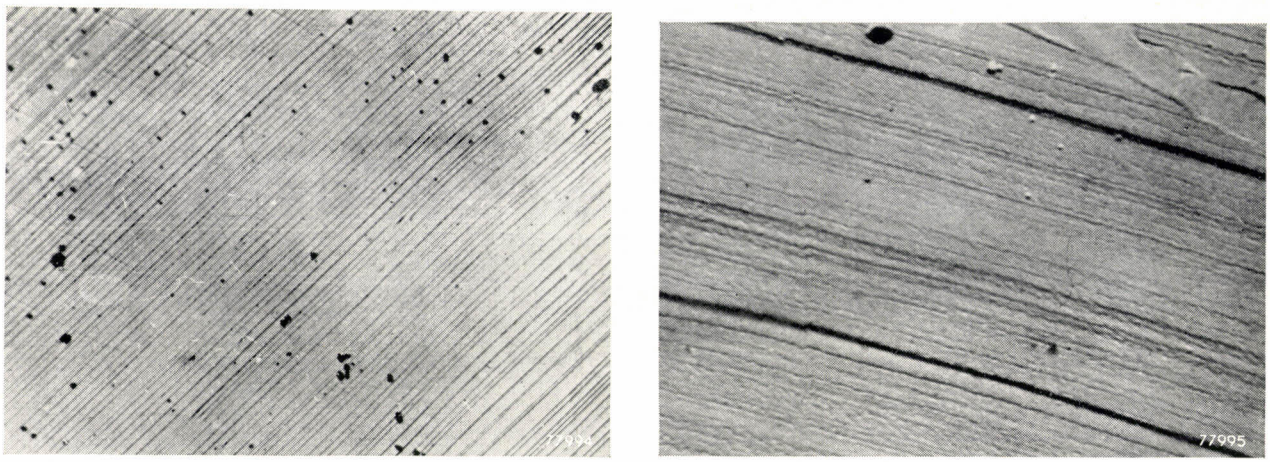
Microscopic examination of the surface of metals which have been polished and subsequently deformed, usually reveals a pattern of fine, more or less straight lines (*fig. 1a*). The more the metal is deformed, the more clearly these so-called *slip lines* become visible and the greater is their number. The orientation of these lines appears almost invariably to correspond to the planes in the crystal in which the atoms are most closely packed. This relationship is found to be the most marked in metals with the closest packed crystal structure. In metals having body-centred cubic lattices which are not packed as closely as possible, the slip lines are somewhat irregular, but the relationship between the slip lines and the crystal planes still remains.

Examination of the slip lines under high magnification by means of the electron microscope, shows that what appear to be single lines when seen with lower magnification, are in many cases groups of some

---

1) G. W. Rathenau, Philips tech. Rev. **15**, 105-113, 1953 (No. 4).

2) F. Seitz. Imperfections in nearly perfect crystals. John Wiley, New York, 1952, p. 3-76.



a *b*

Fig. 1. *a*) Micrograph of the polished surface of a metal single crystal subjected to 7% deformation. Magnification approx. 200 × (R. W. Cahn, J. Inst. Metals *79*, 129-158, 1951). *b*) Slip band on the surface of a polished, deformed crystal as seen under the electron microscope. Magnification 25,000 × (A. F. Brown, Advances in Physics, **1**, 421-479, 1952, No. 4).

tens to hundreds of adjacent lines (fig. 1*b*), which may be termed *slip bands*. Single lines are seen as well, however.

A study of the slip lines or bands has revealed the fact that these are in effect "steps" in the surface of the metal, of which the height may vary from some tens to some thousands of times the atomic spacing. The conclusion to be drawn is clear, viz. that in the deformed metal, translations take place along certain crystallographic planes whose orientation corresponds to the direction of the slip line.

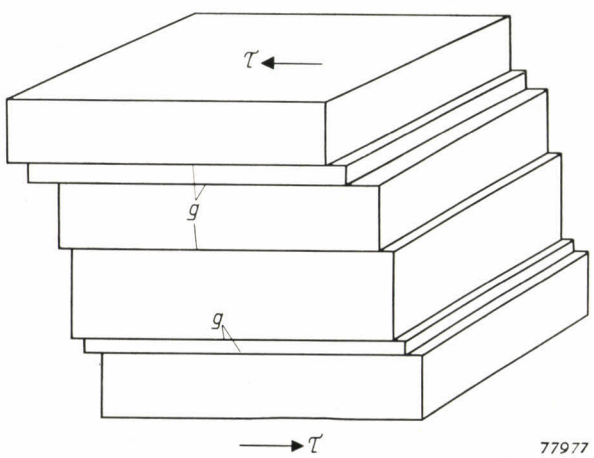


Fig. 2. Diagrammatic representation, on a highly exaggerated scale, of slip in a metal crystal. g slip planes; τ shear stress.

This mechanism is sketched in *fig. 2*, which refers to simple slip lines and is, of course, drawn on an exaggerated scale. This deformation mechanism, which appears to be of a very universal nature, is known as *gliding* or *slip*; the lattice plane along which the slip takes place is termed the *slip plane*, and the direction of the translation, the *slip direction*. The slip plane and direction together constitute the *slip system*.

The theoretical critical shear stress τ_{cr}, that is, the minimum shear stress necessary to produce a slip translation of this kind, is of the order of 0.1 G, where G is the modulus of rigidity, or torsional modulus of the material.

This result is obtained in the following manner (vide Frenkel [3]).

Take the case of two neighbouring rows of atoms in a simple crystal. Let a denote the spacing of the atoms in the undistorted condition. The force required to move the one row an infinitesimal distance over the other depends upon the displacement that both rows have already undergone. In the undeformed state, which is the state of equilibrium, this force is zero. Each time the rows are displaced an integral number of times $a/2$ from the undeformed state, an equilibrium state is restored, as a symmetrical configuration of the lattice is again reached. It would appear to be a reasonable assumption that the relationship between the shear stress (τ) and the relative displacement (x) might be approximated by a simple periodic function:

$$\tau = k \sin \frac{2\pi x}{a} .$$

With only a small displacement this becomes $\tau = k.2\pi x/a$. In this case also Hooke's law applies, which states that:

$$\tau = Gx/a .$$

From this, $k = G/2\pi$; hence:

$$\tau = \frac{G}{2\pi} \sin \frac{2\pi x}{a} .$$

If τ is larger than $G/2\pi$, it follows from this formula that the displacement x of the atoms is unrestricted and slip will occur. Closer investigation yields a value for the theoretical critical shear stress slightly lower than $G/2\pi$, although still of the order of 0.1 G.

[3]) J. Frenkel, Z. Physik. **37**, 572, 1926.

For most metals, G lies between some thousands and some ten-thousands of kilograms/mm^2, from which a theoretical critical shear stress of a few hundreds or a few thousands of kilograms/mm^2 follows. It is a well-known fact, however, that metals can be very much more easily deformed than would appear from these values. The observed critical shear stress of well-annealed single crystal occurs between 0.1 and 10 kg/mm^2; annealed polycrystalline metals yield somewhat higher values but still very much less than the theoretical shear stress.

To explain this ease of deformation, it was postulated that a particular kind of lattice defect, namely *dislocations*, occur very frequently in every crystal. It was very soon apparent that such dislocations do play a dominant rôle in the process of deformation of a metal.

We have referred above to the critical shear stress of well-annealed materials. The annealing minimizes the consequences of any previous deformations which would otherwise manifest themselves, for example, by the increase which occurs in the stress required to deform a metal as the deformation itself increases (*fig. 3*). This effect is known as *work hardening* (it does not occur in all materials; substances such as pitch, for example, undergo no work hardening). A severely deformed and unannealed metal may have a critical shear stress many tens of times higher than that in the annealed state.

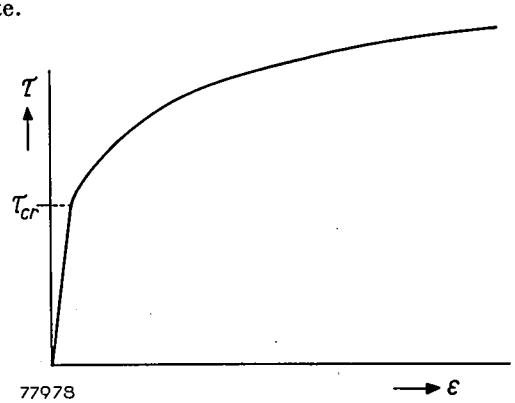


Fig. 3. Diagram showing the relationship between the deformation ε and the shear stress τ in a pure metal. The straight part at the commencement of the curve represents the elastic deformation. When the critical shear stress τ_{cr} is exceeded, plastic deformation sets in, this being accompanied by strain hardening, as a result of which the characteristic assumes a parabolic form (in cubic metals). In practice, irregular curves are often encountered, but the general form remains the same.

As we shall see presently, work hardening in metals is attributable mainly to the mutual interaction of the dislocations associated with the deformation.

The extent to which a metal can be deformed is dependent not only on the previous deformation; variations in the temperature will also affect the ductility, albeit to a relatively small degree. Thus, materials which are brittle when cold are nearly always ductile at elevated temperatures. Foreign atoms or conglomerates in a metal also affect the ductility appreciably; the various hardening processes for metals are based on this fact [4]).

Lastly, it appears that the plastic deformation of a metal affects its general physical properties in varying degrees. Amongst other things, the electrical conductivity will change, and it is found that the study of this and other subsidiary consequences of plastic deformation can reveal valuable information about the behaviour of lattice imperfections.

In this section we have surveyed only the ground on which our further considerations are to be based. Before entering into a detailed discussion of the relationship between the mechanical properties of metals and lattice imperfections in crystals, however, it will be necessary for us to say a little more about present-day conceptions of the nature and behaviour of possible kinds of lattice imperfections. The rest of the present article is accordingly devoted to this subject.

Possible kinds of lattice imperfections

Imperfections in a crystal lattice may be one-, two- or three-dimensional. Also, singularities may occur at the lattice points themselves, in which case we might speak of zero-dimensional, or point defects. The three-dimensional defects such as macroscopic holes or inclusions, precipitates etc, and two-dimensional faults which may be taken to include the crystal boundaries and surface layers, can be dealt with quite briefly. In recent years, several articles have appeared in this Review on the subject of three-dimensional imperfections [4])[5]); as to the two-dimensional defects, these can often be successfully interpreted as more or less ordered associations of linear and point defects.

Dislocations belong to the category of linear imperfections. Point defects may be taken to include vacancies, i.e. lattice points where an atom is missing, and further, interstitial atoms (atoms at intermediate points in the lattice) and impurities (foreign atoms). Whereas the concept of dislocations was originally introduced to provide an explanation of the mechanical properties of crystals, vacancies etc. have been postulated to promote a deeper insight into diffusion effects and other trans-

---

[4]) J. L. Meijering, Hardening of metals, Philips tech. Rev. **14**, 203-211, 1953 (No. 7).

[5]) J. D. Fast, Ageing phenomena in iron and steel after rapid cooling, Philips tech. Rev. **13**, 165-171, 1951.

port phenomena. It is only in recent years that a close relationship has been found to exist between the different kinds of lattice imperfections.

The conception of linear lattice irregularities, i.e. dislocations, was introduced in the theory of metals independently by Taylor, Orowan and Polanyi [6]) in 1934. Since 1939, when Burgers made the first detailed theoretical study of the behaviour of dislocations [7]), the work of N. F. Mott and his co-workers at Bristol has been mainly responsible for the almost general acceptance of the hypothesis of dislocations. The work of the British scientists gave the impetus to the first direct experimental pointers to the existence of dislocations in crystals,

Dislocations

The dislocation concept links up very closely with the most important deformation mechanism in metals, namely the above-mentioned "slip". *Fig. 4* illustrates the present conception of this mechanism. The two halves of the crystal do not move over each other as a whole; this, as already mentioned, would require a much higher shear stress than that which is indicated by experiment.

The displacement commences at one side (left, in fig. 4) of the crystal and is propagated very rapidly to the other side. A situation whereby all the atoms simultaneously assume non-equilibrium positions never arises; the deformation at a given

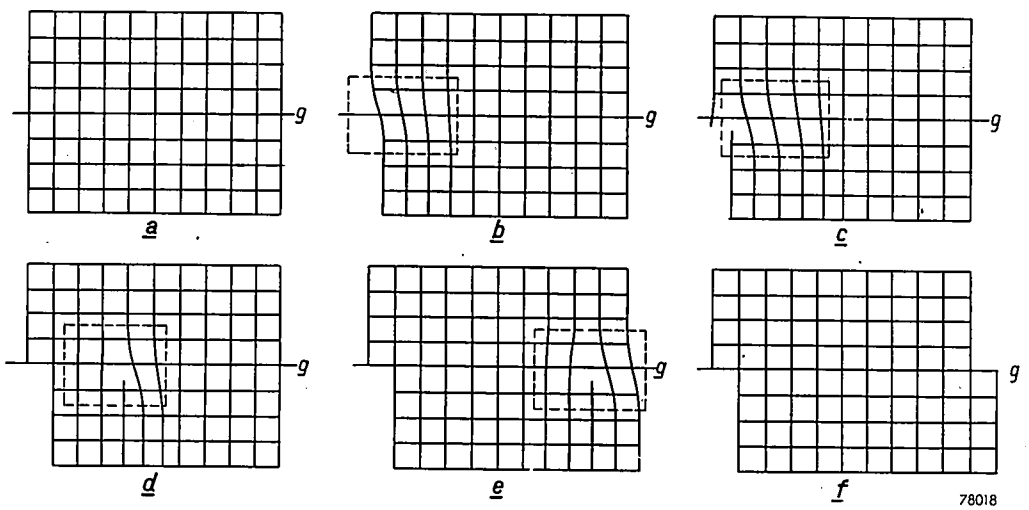


Fig. 4. Diagram representing on an atomic scale the occurrence of slip in crystals, according to the dislocation mechanism. The full lines represent the lattice planes. The two parts of the crystal do not slip simultaneously as a whole along the slip plane *g*; the deformation is localized within a small zone (shown in dotted-line rectangles). Only the cross-section in one atomic plane is illustrated, but it should be visualised as extended without limit perpendicular to the plane of the drawing. Movement of the dislocation from left to right (in the sequence of figures *a* to *f*) completes the slip.

viz. the spiral growth of crystals from supersaturated vapour [8]), as well as the recently discovered sub-structures of photo-sensitive silver bromide crystals [9]). Further direct evidence for the existence of dislocations in metals has been gathered by workers in the United States, from the study of crystal boundaries [10]).

instant is always localized within a zone of some 3 or 4 times the atom spacing. This zone, of which fig. 4 shows only a cross section, i.e. one atomic layer (this layer should be regarded as arbitrarily extended in a direction perpendicular to the plane of the drawing), contains a linear lattice defect, and it is this that has come to be known as a dislocation.

The displacement of the two halves of the crystal in relation to each other has already taken place to the left of the dislocation in fig. 4*d*, but not on the right-hand side. The deformation of the crystal as a whole approaches completion as the dislocation moves to the right along the slip plane; once it has arrived at the opposite side of the crystal (fig. 4*f*) the shear is complete. In this example it is

[6]) G. I. Taylor, Proc. Roy. Soc. A **145**, 362 1934; E. Orowan, Z. Phys. **89**, 634, 1934; M. Polanyi, Z. Phys. **89**, 660, 1934.
[7]) J. M. Burgers, Proc. Kon. Ned. Akad. Wet. Amst. **42**, 293 and 377, 1939.
[8]) F. C. Frank, Advances in Physics. **1**, 91, 1952. See also fig. 2 in the article referred to in footnote 1.
[9]) J. M. Hedges and J. W. Mitchell. Phil. Mag. **44**, 223-224, 1953.
[10]) See, e.g. W. T. Read, Dislocations in crystals, McGraw Hill, New York, 1953.

seen that plastic deformation due to slip is simply the movement of a dislocation.

The critical shear stress is now the force required to initiate a dislocation and cause it to be propagated. To produce this effect, only *a fraction* of the total number of atoms in the region of the slip plane need be simultaneously in a condition of increased energy, so that it follows that this force is very much less than in the case of a perfect lattice.

Dislocations occur not only in metal crystals; in principle they may arise in any crystalline substance. Whether or not they play an important part in the behavior of the material, depends very largely on the crystal structure and the cohesion between atoms. Recent investigations have indicated that in ionic crystals and in non-polar semiconductors, dislocations probably play quite an important role, especially with regard to the optical and electrical properties of these materials. The study of dislocations in non-metals is still in an early stage, however, and we shall therefore not concern ourselves with it here.

the Burgers vector with respect to the dislocation axis. The dislocation referred to in our discussion on the mechanism of slip, is the *edge dislocation*, as characterized by a Burgers vector perpendicular to the dislocation axis. *Fig. 5a* illustrates an edge dislocation in perspective, and it will be seen that, in order to complete the translation of the one part of the crystal over the other, the edge dislocation must be propagated in a direction at right angles to its axis, that is, parallel to its Burgers vector. Fig. 5*b* depicts a cross-section of an edge dislocation perpendicular to the dislocation axis. On close inspection of this diagram (and also of fig. 4), it will be seen that an edge dislocation can be considered as being initiated by the introduction of an additional plane of atoms perpendicular to the slip plane in one of the halves (in this case the upper half) of the crystal. As a result of this, variations in density

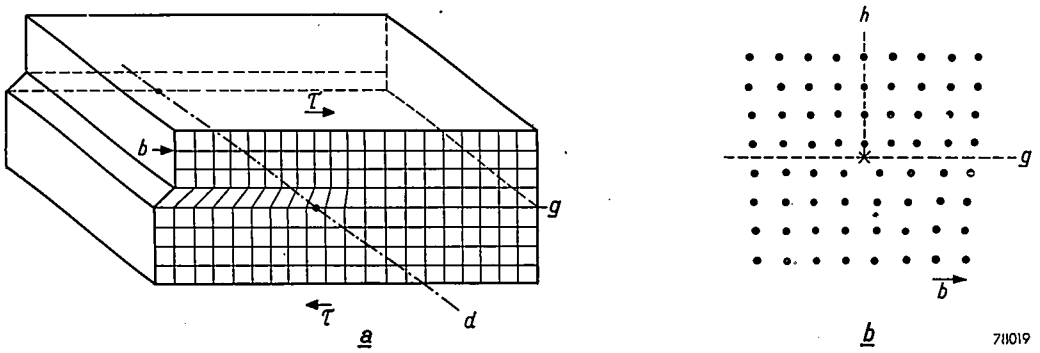


Fig. 5. *a*) Perspective diagram of an edge dislocation which has progressed half way through a crystal. The chain-dot line *d* is the dislocation axis; the light lines represent the atomic planes. The Burgers vector *b* is oriented in the direction of the shear stress τ which strives to complete the slip. To do this, the dislocation must move in the direction of its Burgers vector, along the slip plane *g*. *b*) Cross-section of an edge dislocation perpendicular to the dislocation axis as indicated by $\times$; *g* is the slip plane, *b* the Burgers vector; *h* extra atomic half-plane.

Types of dislocation

Let us now consider the concept of dislocations in a broader sense. The axis of the linear lattice imperfection (which need not necessarily be a straight line) is called the dislocation axis. The dislocation separates the part of the lattice which has already slipped, from the part which has not yet undergone displacement. The degree and direction of the translation in the displaced zone is represented by a vector *b*, known as the *Burgers vector*. Usually, the extent of the translation at a dislocation will be equal to the spacing of the lattice plane in the corresponding crystallographic direction, and the magnitude of the Burgers vector is then equal to this distance. The direction of the vector corresponds to that of the slip direction.

It is thus possible to differentiate between various types of dislocation according to the orientation of

of the crystal lattice are produced around the axis of an edge dislocation; in fig. 5*b* the upper half of the crystal in the region of the dislocation is more closely packed than normally, the lower half less so.

An exactly similar distortion of the crystal, having the same slip plane as in fig. 5, and also produced by the movement of an edge dislocation, can be thought of as being initiated by the introduction of an extra plane of atoms in the *lower* half of the crystal. The direction in which this dislocation would have to be propagated in order to complete the movement is then reversed in sign, and the more densely and less densely packed regions about the dislocation axis change places. We then speak of a dislocation of opposite sign to that of the original.

A second kind of dislocation is the so-called *screw-dislocation*, the Burgers vector of which is parallel to the dislocation axis. A dislocation of

this kind is shown in perspective in *fig. 6a*. The regions in which slip has occurred (right-hand side) and has not occurred (left-hand side) are now separated by a linear lattice imperfection parallel to the direction of slip, in contrast to the edge dislocation, which is oriented perpendicular to the slip direction. Once more, in order to complete the translation, the dislocation must move in a direction perpendicular to its axis, but here this means perpendicular to the Burgers vector. In consequence of the peculiar mutual orientations of the Burgers vector and the dislocation axis, the screw dislocation has no specific slip plane, as this plane is defined as the plane through the dislocation axis and the Burgers vector. Whereas the edge dislocation, in order to accomplish slip, must move over a very has been accomplished equal in extent to the Burgers vector in the direction of the dislocation axis. In other words, a set of parallel atomic planes (perpendicular to the dislocation axis) in the undistorted crystal, forms, on the introduction of a dislocation, a helical surface having the dislocation line as axis. The atomic structure around a screw dislocation is depicted in fig. 6*b*. It is not possible to visualise the screw dislocation as being produced by introducing an additional atomic plane, and there is accordingly no question of any appreciable variation in density within the crystal, as occurs around edge dislocations.

A sign can also be attributed to the screw dislocation, since here again an equivalent translation can be brought about in two different ways: the

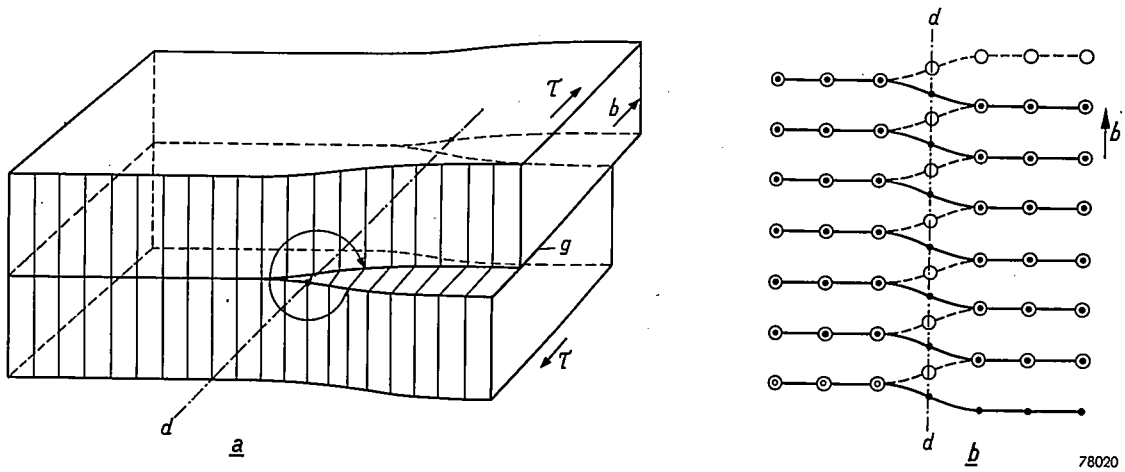


Fig. 6. *a*) Perspective diagram of a screw dislocation displaced half way along its slip plane *g*. The chain-dot line *d* is the dislocation axis. τ is the shear stress which tends to move the dislocation along the slip plane, perpendicular to the Burgers vector *b* and to the shear stress itself, in order to complete the slip. The curved arrow round the dislocation axis indicates that a complete revolution about the dislocation axis within the same atomic plane does not result in a return to the point of origin; the atomic planes describe a helical surface about the dislocation axis. *b*) Atom positions in the lattice planes above and below the dislocation axis *d* and parallel to that axis. The open circles represent atoms in the upper plane and the dots those in the lower plane. *b* is the Burgers vector.

definite slip plane, the screw dislocation is free to select its own direction of propagation (provided it is perpendicular to its axis).

This is not the only point of difference between the two kinds of dislocation. From fig. 6*a* it will be seen that a screw dislocation can be imagined as formed by making an incision of some depth (in the perfect crystal), this being followed by displacement of the part above the incision, parallel to the bottom of the cut, through a distance equal to the atomic spacing. The bottom of the cut is the dislocation axis. If we now describe a path round the dislocation axis, keeping within the same atomic plane, it will be seen that on completion of one whole turn we do not return to the point of origin as in the undistorted crystal, but that a translation screw dislocation can transform the atomic planes into either right-hand or left-hand helical surfaces. Fig. 6 shows a right-hand screw. Right and left-hand screw dislocations must move in opposite senses in order to produce the same translation.

Edge and screw dislocation represent only the extreme instances of the general conception of dislocations in which the angle between the Burgers vector and the dislocation axis is arbitrary.

The character of a dislocation need not necessarily be the same throughout its length: it may vary between one point and another. Moreover, a dislocation element may change its character in the course of propagation through the crystal. A probably common form of dislocation is the *dislocation loop* depicted in *fig. 7*, the feature of

which is that within the loop the crystal has slipped and that outside it no slip has occurred, since by definition, a dislocation separates a zone in which slip has taken place from another in which it has not. Everywhere within the loop, the relative amount of slip must be of the same magnitude and in the same direction; otherwise the loop would reveal branches. In other words, around a closed dislocation loop without any branch points, the Burgers vector — which determines the translation — is at all points constant in magnitude and direction. Only certain parts of the loop, as indicated by E and S in fig. 7, conform to the requirement of a purely edge or screw type of dislocation; at all other points the character of the dislocation is composite. The dislocation therefore nearly everywhere includes elements with at least some edge-character, i.e. differences in density are present, and the movement will be confined to a definite slip plane.

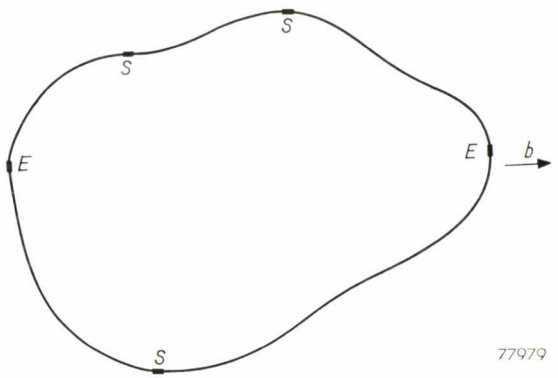


Fig. 7. Dislocation loop. b is the Burgers vector which is constant along the whole loop. Elements E, perpendicular to the Burgers vector are purely edge-type elements; elements S, parallel to the Burgers vector, are purely screw type.

Dislocation loops are probably of frequent occurrence, since dislocations cannot be terminated somewhere in the middle of a perfect crystal.
Along the dislocation there are invariably two zones which are displaced a definite amount with respect to each other. Where the dislocation is terminated, this relative displacement must disappear, and this can only happen on the surface of a crystal, at a crystal boundary, or at points where another dislocation is present. In the absence of any of these essentials, that is in the centre of a perfect crystal, only continuous, e.g. loop dislocations are possible.

Mechanical model of a dislocation

To study the characteristics of dislocations, it is usually sufficient in principle and without sacrifice of generality to confine our investigation to pure edge and screw dislocations. Using a mechanical model [11]) such as that shown in *fig. 8*, it is possible to illustrate the occurrence and movement of the dislocations.

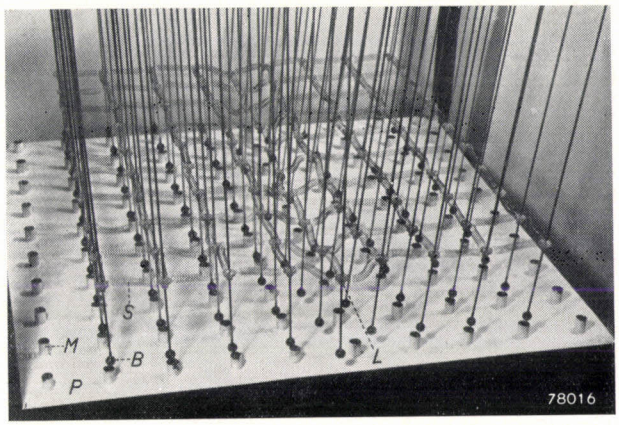


Fig. 8. Mechanical model for reproducing dislocations. Magnets N are arranged in a square pattern on a base plate P. A rack capable of movement to the left and right (not shown in the illustration) carries a number of pendula with iron balls B which, in the condition of equilibrium, hang just above the magnets. Springs S are fitted between the pendula. The illustration shows an edge dislocation with axis L. The extra atomic plane is clearly seen. Slight movement of the board carrying the pendula causes the dislocation to progress through the model in a direction perpendicular to its axis, i.e. to the right or left hand side, thus producing a displacement.

A large number of permanent magnets are arranged on a board in a square pattern, to represent a single lattice plane of the crystal (which for convenience is supposed to be simple-cubic). Above this board a rack is mounted which is movable in either of the square-directions and which carries a number of pendula with "bobs" of soft iron. These iron balls also represent a lattice plane. In the state of equilibrium, each pendulum hangs over a magnet.

Whereas the "atoms" in the lower plane have fixed positions, those in the upper plane are capable of movement. To represent the inter-atomic forces, all the pendula are interconnected by small helical springs which exert no tension in the state of equilibrium, but which somewhat impede the approach or separation of any two of the pendula.

When the rack of pendula is pushed to one side by hand, the magnetic forces between the magnets and iron balls is at first sufficient to resist the resultant "shear stress". When the stress reaches a certain appreciable value, first one row of balls will be seen to jump over a distance equal to one lattice spacing in the direction of the stress; the application of a very slightly larger stress causes neighbouring rows in succession to carry out the same movement until finally the whole system of

[11]) H. G. van Bueren, Brit. J. Appl. Phys **4**, 144-145, 1953.

pendula has moved up a distance equal to the lattice space. Thus a dislocation is seen to migrate through the lattice.

It is possible that the first row of balls to jump over will be perpendicular to the direction in which the stress is applied, as shown in fig. 8; in this case we have an edge dislocation. Again, the first row of "atoms" executing the movement may be parallel to the direction of the applied stress; in this case it is a screw dislocation which is seen to develop and move through the lattice (*fig. 9*). Which particular kind of dislocation will occur depends entirely on accidental factors.

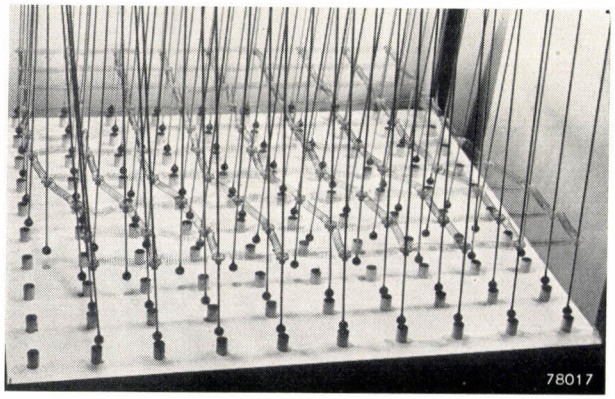


Fig. 9. As fig. 8, but showing a screw dislocation. Movement of the pendulum rack causes the dislocation to move backwards or forwards through the model.

Movement of dislocations

Although the model provides only a very rough approximation to the real conditions, it is possible to note at once two important properties of dislocations. Firstly, the stress needed to *initiate* a dislocation is very much greater than that required to *set it in motion*. Secondly, under the influence of the stress, the dislocations are indeed propagated in such a way that the deformation of the crystal by slip is *completed*. Edge dislocations move in the direction of the stress and screw dislocations perpendicular to it.

It can be said, then, that a force operates on the dislocation. The magnitude of this force can be computed in the following manner (vide Mott and Nabarro [12]).

Let us assume a dislocation of which the Burgers vector is b, running from one side of a cube-shaped crystal of dimension L to the other side (cf. fig. 5). When it reaches the opposite side, one part of the crystal will have been displaced in relation to the other through a distance b. Let the components of the shear stress in the slip direction be τ_g; the crystal area on which this operates is L^2, so that the work performed by the external force in producing the slip is:

$$W = \tau_g L^2 b .$$

The length of the dislocation line is L, as is also the distance travelled by the dislocation. If we now denote the force acting on a unit length of the dislocation by F, it may be said that the work done must be equal to:

$$W = FL^2 .$$

(This is actually the *definition* of a "force acting on a dislocation".)

The result of a shear stress whose component in the direction of the Burgers vector is τ_g (and it can be shown that this is universally valid) is therefore a force F per unit length of the dislocation equal to

$$F = \tau_g b , \quad . \ . \ . \ . \ . \ . \ . \ . \ (1)$$

which tends to move the dislocation in the sense of completing the slip movement.

Only a slight force is needed for the propagation of the dislocation along its slip plane. If the whole dislocation is to be displaced by atomic distance, it is only necessary for the atoms themselves to move a fraction of this distance (*fig. 10*). In doing this, half of the atoms move under the influence of an attracting force and the other half under a repelling force, so that, to a first approximation, the forces counterbalance each other and the dislocation moves easily. It is owing to the discontinuous structure of the crystal that a dislocation cannot move in its slip plane entirely without effort, as might be inferred by the above remarks. Different positions of a dislocation axis between two lattice planes are not exactly equivalent

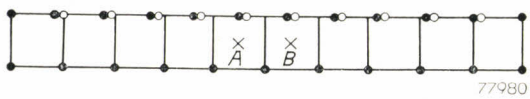


Fig. 10. Diagrammatic representation of the movement of an edge dislocation in its slip plane. The upper row contains one more atom than the lower. In the first instance, the dislocation axis lies at the cross marked A, and the atoms assume the positions shown by the dots. In order to displace the dislocation axis by one atomic spacing to the point B, the atoms themselves need undergo only slight displacement, viz from the dots to the adjacent circles. In doing so, the atoms in the upper row to the left of A move so as to approach more closely to the position of equilibrium, i.e. directly above the atoms in the lower row. These atoms move under the influence of attracting forces. The atoms to the right of A move further from the position of equilibrium against the action of repelling forces. As a first approximation, the effects of the two forces are counterbalanced and the dislocation can move freely in its slip plane. (A. H. Cottrell, Progress in Metal Physics **1**, 77-125, 1949).

[12]) N. F. Mott and F. R. N. Nabarro, Report on strength of solids, The Physical Soc. London, 1948, p. 1.

energetically. In metals of cubic and hexagonal structure, this results in only a slight reduction in the mobility of the dislocation in its slip plane. The mobility is in any case quite high and, as the atomic arrangement and hence the lattice vibrations in the crystal affect it but little, it is not very dependent on temperature. It is for this reason that dislocations play such an important part among lattice defects.

In the foregoing we have referred to motion within the slip plane of the dislocation. For a screw dislocation, every plane of movement can be considered as a slip plane. For dislocations not bearing the character of *purely* screw dislocations, however, we should also consider movements which do not meet this condition.

Dislocations with some edge-character are always associated, in the manner described for pure edge dislocations, with an extra half plane of atoms.

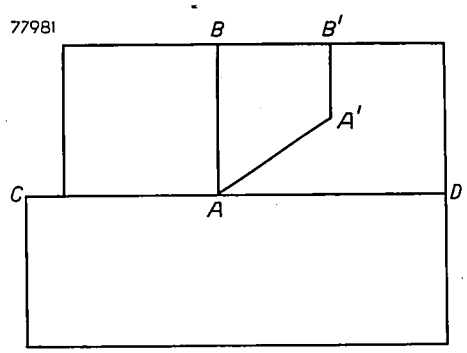


Fig. 11. Illustrating non-conservative movement of an edge dislocation (vide A. H. Cottrell. Dislocations and plastic flow in crystals. Oxford Univ. Press, 1953).

Fig. 11 illustrates this situation diagrammatically. The plane of the drawing represents a cross-section perpendicular to the dislocation axis A. The extra half plane is represented by the line AB, and the Burgers vector lies in the slip plane of which the line CAD is the cross-section. With movement out of the slip plane, for example towards the point A', the length of the extra half plane changes from AB to $A'B'$. This means that a number of extra atoms have been removed (or added), and this can happen only if accompanied by a simultaneous transport of atoms to or from other parts of the crystal, or, if this is impossible, at the expense of considerable local distortion. Both processes necessitate activation energy (e.g. for diffusion); so also, therefore, does the movement of a dislocation beyond its slip plane.

Accordingly, whereas movement within the slip plane needs very little energy and is hardly dependent on temperature (conservative movement), movement outside the slip plane (non-conservative movement) demands very much more energy, which cannot in general be derived from the externally applied forces; furthermore, owing to the fact that an activation energy is involved, the latter is highly dependent on the temperature.

Clearly, a screw dislocation performs only conservative movements, since each crystallographic plane containing its axis can serve as slip plane. Needless to say, screw dislocations in crystals as well as other kinds of dislocation can be propagated only over crystallographic planes. Chalmers and Martius [13]) have pointed out that it makes a difference *which* planes are involved. They showed that during its propagation, a dislocation will reveal a preference for planes containing the most closely packed layers of atoms. The slipping movement therefore occurs mainly along such planes.

Since the slip lines in a distorted crystal are the markings of the slip planes, this explains why, as mentioned above, the slip lines in many metal crystals run along the most closely packed crystal planes.

Origin of dislocations

Dislocations represent a considerable amount of energy by reason of the distortion in the crystal lattice that accompanies such lattice defects. The energy of a dislocation is defined by its stress field, the nature of which is rather intricate and would take us too far afield to consider here. It appears that the stresses decrease in inverse proportion to the distance from the dislocation axis, in consequence of which, in a continuous medium, the energy of a dislocation would be infinitely high. This is not the case with crystals because of their discrete atomic structure; it is found that the energy of a dislocation per atomic plane is several times Gb^3, i.e. in most metals, several times 10^{-12}erg.

A proper dislocation is at least some tens of times the atomic spacing in length. The energy of such dislocations, that is, the energy required to produce them, is thus always 10^{-10} erg or more.

The presence of a dislocation increases the entropy of the crystal. It can be shoron that this effect is negligible compared to the energy of formation, which means that dislocations cannot possibly be initiated thermally, or exist in thermal equilibrium with the crystal lattice. In a metal crystal plastically elongated about 1%, we must nevertheless accept the occurrence of dislocations

[13]) B. Chalmers and U. N. Martius, Nature **167**, 681, 1951.

with a total length in the crystal of at least 10^7 cm per cm^3 (i.e. a dislocation density of 10^7. cm^{-2}), as demonstrated by experiments which will be discussed in the second part of this article.

This apparent contradiction has been solved by adopting the view that dislocations occur as a result of the actual process under consideration, viz. the plastic deformation. In principle this can be explained in the following manner.

If inhomogeneities occur somewhere within the crystal or at its surface, e.g. small cracks or inclusions, applying a stress will result in local stress concentrations, i.e. zones in which the stresses are much higher than elsewhere in the crystal. It is plausible that the formation of one or more dislocations in such zones might lead to a reduction in energy. If this is so, such zones will be capable of acting as sources of dislocations.

Which particular kind of inhomogeneities are likely to act best as sources, is not yet known with certainty. It is very probable that the most likely source of dislocations will be found at the crystal boundaries.

Dislocations may also be formed during the growth of the crystal. According to a theory put forward by Frank [8]), screw dislocations can very appreciably accelerate the growth of a crystal. This point has already been mentioned in the article referred to in footnote [1]), so that there is no need to dwell further on this very remarkable and interesting point here. So far as we are concerned here, we are interested only in the fact that a crystal grown in accordance with the mechanism described by Frank will invariably contain dislocations.

Although the two formative mechanisms described throw some light on the origin of certain dislocations distributed at random throughout the crystal, they by no means explain the observed phenomenon that, in a given slip plane, slip often occurs over a number of atomic spacings simultaneously, whereas in other crystallographic planes no slip takes place at all. This phenomenon can be due only to the movement of large numbers of dislocations all lying in the same slip plane. A possible way out of the difficulty is offered by the hypothesis of a *dynamic multiplication* mechanism for the dislocations initiated in one of the ways outlined above. A condition for the occurrence of this mechanism (a *Frank-Read* source, so named after the originators of this conception [14]) is the presence of dislocation lines in the crystal, anchored at two points, so that they can deflect as a result of a shear stress without being able to move as a whole. Mott [15]) has shown the plausibility of the existence of such anchored dislocation elements in metallic crystals. He pointed out that the dislocations in non-distorted metals with cubic symmetry very probably form a space

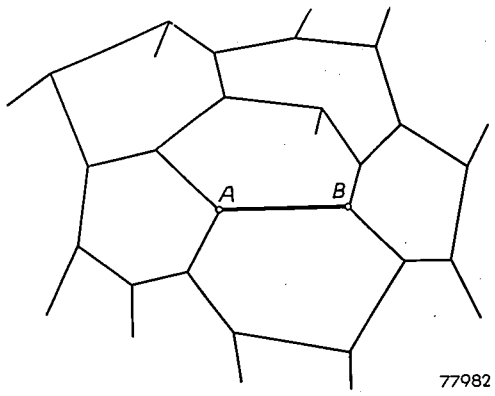


Fig. 12. Spatial network of dislocations in a crystal (vide Mott [15])).

network, containing many nodes in which three dislocation lines, not located in the same plane, meet (*fig. 12*). If the orientation of the applied stress is such that a considerable force acts on one of the three converging dislocations, this usually means that the other two dislocations are subjected to only a slight force, or even to a force in the opposite direction. The nodes (e.g. *A* and *B* in fig. 12) will then constitute more or less fixed points in the dislocation network, and hence act as anchoring points for those dislocations which would otherwise tend to move under the inflence of one force or another.

An anchored dislocation element (1) is illustrated in *fig. 13*. For convenience we will suppose that this

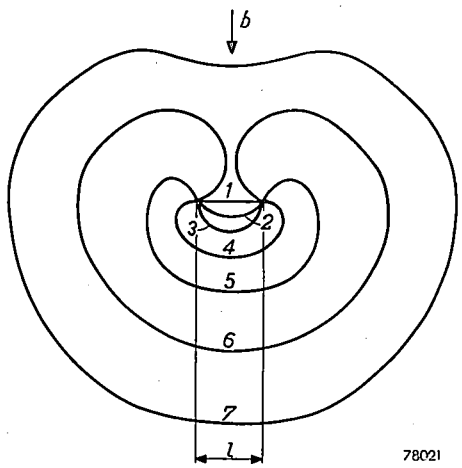


Fig. 13. Diagram demonstrating the multiplication mechanism of dislocations according to Frank and Read.

[14]) F. C. Frank and W. T. Read, Phys. Rev. 79, 722, 1950.

[15]) N. F. Mott, Phil. Mag. 43, 1151, 1952.

is a purely edge-type dislocation. If a shear stress τ is applied, the dislocation will be deflected; it thus becomes longer, and the dislocation energy accumulated within it increases. Deflection is continued until a position is reached (e.g. position 2 in fig. 13) whereby the gain in potential energy by the dislocation in the stress field τ is balanced by this increase in dislocation energy. If the stress increases, position *3* will ultimately be reached, when the dislocation will have assumed the form of a semi-circle. This is a critical form. It can be shown that if the stress is increased gradually from this point, the energy required to produce further deflection of the dislocation (and hence to increase its length) is always less than the gain in potential energy that accompanies this further deflection. The stress required to produce the critical semi-circular configuration can be computed, being approximately:

$$\tau_0 = \frac{Gb}{l}, \quad \ldots\ldots\ldots (2)$$

where b is the Burgers vector of the dislocation and l the length of the original straight anchored element. If the critical value τ_0 of the shear stress is exceeded, the dislocation expands *ad lib* without any further increase in the stress. Those parts of the semi-circle which are now practically perpendicular to the original edge dislocation now have the character of a screw dislocation (the Burgers vector, which remains constant, is at those points almost parallel to the dislocation axis). These screw-like parts of the dislocation thus move laterally outwards, perpendicular to the direction of the strain. Those parts which are practically parallel to the original dislocation, progress as an edge dislocation parallel to the direction of the strain. This being so, positions *4* and *5* are traversed, the strain being the same all the time, until position *6* is reached. Here, two parts of the screw of opposite sign face each other and, the movement being continued, will meet. They then cancel each other, since two dislocations of opposite sign coming together result once more in a perfect lattice. Finally position *7* is assumed, this being none other than the condition at *1*, plus a dislocation loop enclosed by the "source" and moving outwards under the influence of the applied stress.

In principle, an unlimited number of loops could emanate from the source, assuming that the applied stress is at least equal to the critical stress τ_0.

This would provide a simple explanation for the observed translations of several hundreds of times the atomic spacing along a slip plane.

It is an obvious step to identify (in order of magnitude) the observed critical shear stress with the critical stress τ_0. If this is done, it will be found from formula (2) that in most metals the length l of the source in fig. 13 is several times 10^{-4} cm. Of course, longer sources will start working at lower stress levels than short ones, but it must apparently be accepted that the length of most sources is about 10^{-4} cm; this is of the same order of magnitude as the estimated average length of the *network elements* in Mott's model.

Formation of vacancies and interstitial atoms

If the energy of formation of a lattice imperfection and the change in entropy of the crystal associated with it are known, it is possible to estimate the thermal equilibrium concentration of such imperfections as a function of temperature. This concentration is proportional to the Boltzmann factor

$$e^{-U/kT},$$

where U is the (free) energy of formation, T the absolute temperature and k Boltzmann's constant. At room temperature, the value of the product kT is 4×10^{-10} erg.

The formation energy of vacancies and interstitial atoms in metals is not known with any degree of accuracy. Huntingdon and Seitz [16]) have computed the energy for a vacancy in a metal such as copper to be roughly 10^{-12} erg. The formation energy of interstitial atoms will probably be several times higher. These values agree fairly well with the conclusions drawn from experimental work, and with their aid it can be computed that a concentration of vacancies or interstitial atoms of some appreciable magnitude can exist in equilibrium with the lattice only at temperatures near the melting point of a metal. If any indications are found of the formation of such lattice defects at normal temperatures, it must be concluded that their origin is not thermal. In fact, there *are* indications that during the process of deformation of a metal at ordinary temperatures, very large numbers of vacancies and interstitial atoms do occur, and we shall refer to this again in our next article.

The obvious course is to seek a mechanism that is responsible for the occurrence of lattice defects during deformation, i.e. during the formation and propagation of dislocations. Several mechanisms have been proposed, of which only one will be discussed here.

[16]) H. B. Huntingdon and F. Seitz, Phys. Rev. **61**, 315 and 325, 1942.

We have already seen above that the non-conservative movement of a dislocation (i.e. of an edge type of dislocation outside its slip plane) is necessarily accompanied by a variation in the length of the "extra half plane" of atoms involved in the dislocation. A number of extra atoms is withdrawn or added. If the temperature is high enough, or if the movement takes place very slowly, these atoms can be transferred to other parts of the crystal by diffusion. If these conditions are not fulfilled, however, a shortage, or surplus, of atoms occurs in the neighbourhood of the dislocation; in other words, there will be a number of vacancies or interstitial atoms. The question is, then: can non-conservative movements of dislocations occur during plastic deformation at low temperatures? Seitz and Mott [17]) have shown that such movements do indeed take place on a large scale.

Dislocations emanating from a Frank-Read source travel through the crystal, and it is inevitable that they encounter parts of the original and still existent network. The further movement, as a result of the applied stress, of a dislocation loop straight across the network element, results in the formation of discontinuities in both dislocation elements at the point where they intersect. Consider the simple case of two screw dislocations at right angles to each other (*fig. 14*). If dislocation *1* is intersected by dislocation *2*, the relative translation within the crystal produced by the latter will result in a displacement of the two parts of dislocation *1* on both sides of dislocation *2*. A "jog" is thus formed in dislocation *1* equal in length to the Burgers vector of *2*, and in the same direction as this vector. Conversely, a jog occurs in dislocation *2* in accordance with the Burgers vector of dislocation *1*. These jogs behave like small sections of dislocation, with the same Burgers vector as that of the original dislocation, but of course with a different orientation of axis. In this particular case, the jogs constitute small edge dislocations, since their Burgers vectors (those of the original screw dislocations) are perpendicular to their axes. Further movement of the screw dislocation is in effect non-conservative movement of these edge dislocations. The movement of the jogs would be conservative only if it took place in their own slip planes, that is, the planes through the jog and the axis of the screw dislocation. Actually, however, these jogs move together with the screw dislocation, i.e. they move in a plane perpendicular to that of the axis of the screw dislocation. Consequently, with further rapid movement as produced by the applied stress, the jogs leave behind them a row of vacancies or interstitial atoms.

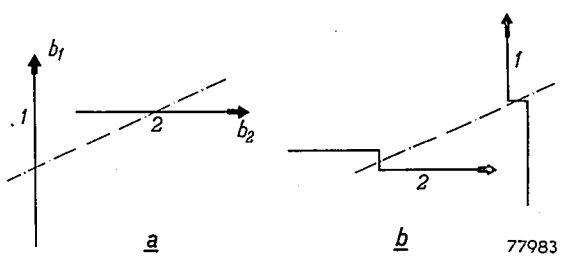


Fig. 14.*a*) Two mutually perpendicular screw dislocations 1 and 2 with Burgers vectors b_1 and b_2 approach each other along the dotted line. b) When they intersect each other, each dislocation reveals a kink equal in magnitude to the Burgers vector of the other dislocation. These kinks behave as edge dislocations, with a Burgers vector which is the same as that of the screw dislocation in which they occur. Further movement of the screw dislocations in the original sense results in non-conservative movement of the kinks.

[17]) N. F. Mott, Proc. Phys. Soc. **B 64**, 729, 1951; F. Seitz, Advances in Physics **1**, 43, 1952.

Not only will the perpendicular intersection of screw dislocations almost always produce a jog with a non-conservative movement impressed on it: in general, the intersection of any two dislocations will also do this. In this way it is possible for very large numbers of vacancies and interstitial atoms to occur during plastic deformation.

We have now discussed both the formation and the principal characteristics of dislocations and other lattice defects in metals. In the next article, the effect of such lattice defects on the deformation and other related phenomena will be investigated.

Summary. The theoretical necessity for the existence of lattice imperfections in crystals, such as vacancies, interstitial atoms and dislocations, is clarified in the light of a number of phenomena related to the plastic deformation of metals, e.g. the occurrence of slip lines, work hardening and variations in the electrical conductivity. This article is devoted mainly to a discussion of the more important characteristics of dislocations, and deals successively with the geometrical aspects of various kinds of dislocation, their behaviour under shear stresses in the crystal as illustrated by means of a model, and the movement of dislocations, taking into account the difference between conservative and non-conservative movements. Lastly, the formation of lattice defects of different kinds as the outcome of interaction between existing defects is examined. The effects of these imperfections on certain mechanical properties will be dealt with in a subsequent article.

AN IMPROVED ION-TRAP MAGNET

by W. F. NIKLAS.

621.385.832:621.397.62:537.533.7

The quality of television pictures has improved so much in recent years that minor improvements in the receiver, the effects of which would formerly not have been noticeable, are now of some significance. Although such improvements do not yield spectacular results individually, their combined effect is of considerable importance in further raising the quality of reception. This article refers to an example of one such improvement in the picture-tube.

Object and principle of the ion trap

The electron beam in a television tube comprises not only electrons, but also positive and negative ions. The positive ions migrate to the cathode and need not be further discussed, but negative ions are accelerated in the direction of the screen. Owing to the fact that these negative ions are deflected to a much smaller extent by magnetic fields than electrons, the focusing magnet and deflection coils have but little effect on their paths. Unless they are removed from the beam, they fall in a continuous stream on the screen, roughly in the centre, where they tend to render the screen inactive; the result is a dark spot in the picture, known as "ion burn". This increases in density the longer the tube is used, and thus considerably shortens the life of the tube.

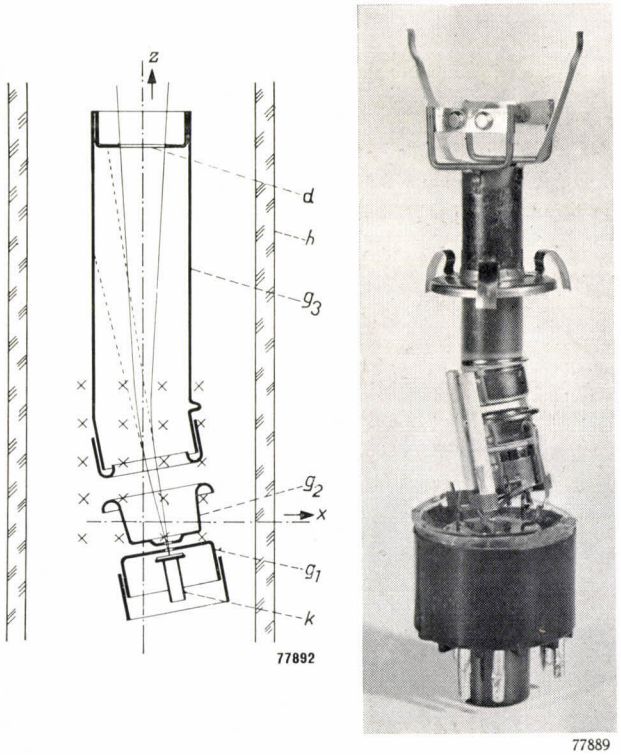


Fig. 1. *a*) Electron gun of a television tube with bent electrode g_3 to function as an ion trap. The electron beam (full lines) is deflected by a transverse magnetic field (indicated by crosses) and is directed through the aperture d in the electrode. The ions (dotted lines) undergo little or no deflection and are intercepted by the electrode g_3. k cathode, g_1 control grid, g_2 acceleration electrode, h neck of tube.
b) Photograph of electron gun of the type shown in (*a*).

It is necessary therefore to eliminate the ions from the beam, and the device by means of which this is done is known as an "ion trap". This may take a number of forms, but *fig. 1* depicts the more usual arrangement (with bent accelerating electrode). A feature common to all ion traps is that they employ a transverse magnetic field which deflects the electrons to a much greater extent than the ions. This difference in the degree to which ion and electron paths are affected by magnetic fields is therefore equally the cause of ion burn and the means of its elimination. In the arrangement shown, the electrons are deflected some 11° while the ions pass straight on; the latter accordingly follow the path shown by the dotted line, and are intercepted by the final accelerating electrode, whereas the electrons pass through the aperture. The magnetic field is supplied by a small exterior magnet. A diagram of a hitherto widely used form of ion trap magnet is seen in *fig. 2*; the magnet mounted on the tube is depicted in *fig. 3*.

The angle through which the electrons are deflected is greater the lower the speed at which they traverse the magnetic field, that is, the lower the voltage of the final electrode (the acceleration voltage). To ensure that the electrons will pass exactly through the aperture in this electrode, the field must be matched to the acceleration voltage, this being done by moving the ion trap magnet along the neck of the tube. The nearer the magnet is placed to the cap of the tube, the more of its field lies behind the cathode and is therefore ineffective; hence the lower the accelerating voltage, the nearer the magnet must be placed to the cap on the tube.

The distribution of the field, which is determined

largely by the shape of the pole-pieces on the ion-trap magnet, has some effect on the focusing of the beam on the screen of the tube. By careful design of the pole pieces it has proved possible to reduce this usually undesirable effect to a minimum, without increasing the cost of production. In the following paragraphs we shall set out the considerations that have led to the present design.

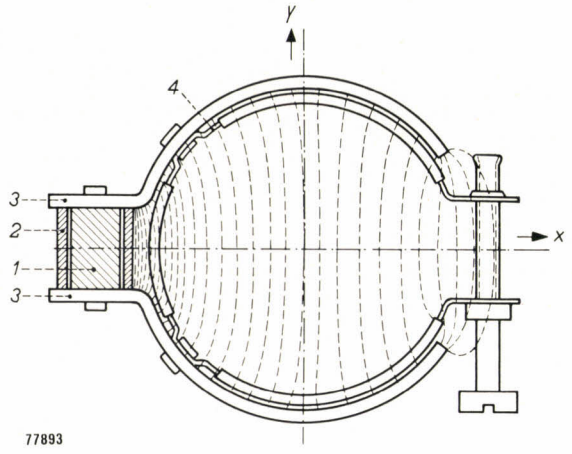


Fig. 2. Ion-trap magnet of the conventional type. *1* permanent magnet (Ticonal E). *2* bush against which the soft iron pole pieces *3* are clamped. *4* strap with screw, of non-magnetic material, for mounting the magnet on the neck of the television tube. This strap is fitted with pieces of soft material to avoid damage to the glass tube. The pattern of the lines of force between the pole pieces is shown diagrammatically.

Deflection in a magnetic field

It is probable that the negative ions originate partly in the cathode and partly in residual gases in the neighbourhood of the cathode [1]). Only a small error is involved in assuming that both electrons and negative ions leave the cathode with zero velocity. In a cross-section of the beam, where the potential with respect to the cathode may be denoted by V, a particle with a negative charge q will have acquired a kinetic energy of qV. Let the mass of the particle be M; its velocity will then be:

$$v = \sqrt{\frac{2qV}{M}}. \quad \ldots\ldots\ldots \quad (1)$$

Elementary mechanics tells us that the radius of curvature of the path of a particle under the influen-

[1]) An article is shortly to be published on this subject in Philips Research Reports. See also C. H. Bachmann, G. L. Hall and P. A. Silberg, J. appl. Phys. **24**, 427-433, 1953 (No. 4).

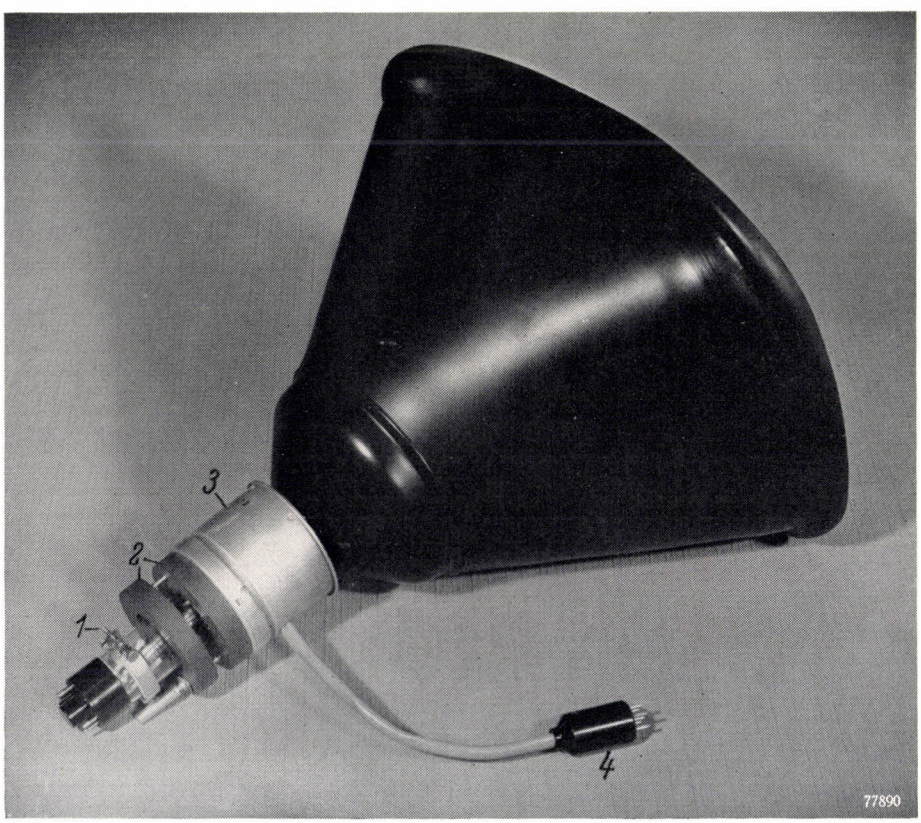


Fig. 3. Television picture-tube with ion-trap magnet *1*, of the type depicted in fig. 2, clamped to the neck of the tube. Also shown are the focusing magnet *2* (Ferroxdure permanent magnets) and the housing *3* containing the deflection coils; *4* plug for connecting deflection coils.

ce of a force K_n perpendicular to its path is given by

$$R = \frac{Mv^2}{K_n}.$$

In the field B of the ion-trap magnet, which is at right angles to the plane of the path, the perpendicular force K_n is the electromagnetic force Bvq. Substituting this and eq. (1) in the expression for the radius of curvature, we find that

$$R = \frac{1}{B}\sqrt{2V}\sqrt{\frac{M}{q}}. \quad \ldots\ldots \quad (2)$$

The accelerating electric field will generally also have a component perpendicular to the path of the particle, but the contribution that this makes towards the force K_n is not taken into consideration in formula (2).

At every point, then, the radius of curvature is proportional to $\sqrt{M/q}$. This quantity is much greater for ions than for electrons (43 times for hydrogen ions; 170 times for singly charged oxygen ions), so that the radius of curvature for the ions in the electron beam is much greater and they are not affected to such an extent by a given magnetic field.

The field of the ion trap magnet

The angle of 11° through which the electron beam is deflected can be obtained by means of magnetic fields of widely differing configurations. In order to get some idea of the kind of configuration that is desirable for an ion-trap magnet to give the least possible defocusing of the electron beam, we will take the following simple example.

Let us assume a system of co-ordinates as shown in figs. *1a* and 2 (the z-axis coincides with the centre line of the tube). We shall further assume that the field of the ion-trap magnet in the region of the z-axis is solely a function of z and that the same applies to the electric potential V. According to (2) the radius of curvature of the electron paths is then dependent only on z. A parallel electron beam, after passing through a field of this kind emerges parallel; thus if the path of one electron is given, that of all electrons in the beam can be obtained by displacement at right angles to the z-axis. From this it follows that for a beam bent as shown in *fig. 4a*, the cross-section perpendicular to the beam is slightly *larger* on leaving the deflecting field than on entering it; the difference is very small however. The cross-section changes by a factor determined by the cosine of the deflection angle; in the ion trap under discussion this factor is $1/\cos 11° = 1.02$.

Consider now the effect of a field that *is* dependent on x. If, for every value of z, the field decreases in the direction corresponding to the positive direction of x in fig. 4, the path *2* of electrons which are parallel on entry (fig. 4*b*) will

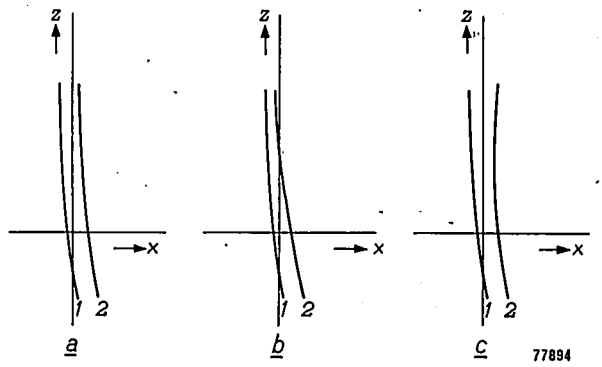


Fig. 4. *a*) If a beam of electrons whose velocities depend only on z, enters a magnetic field whose strength varies only in the z direction, all the paths in the beam can be obtained from a given electron path, merely by a displacement perpendicular to the z-axis. An initially parallel beam thus emerges parallel.
b) If the field decreases in strength in the positive direction of x, an electron *2* will at all times be in a weaker field than an electron *1*. The path of *2* is therefore less curved than that of *1*, and the beam is convergent in the x-z plane.
c) If the field increases in strength in the positive direction of x, electron *2* will at all times be in a stronger field than *1*, and the beam is then divergent in the x-z plane.

be slightly less curved than that of *1*, for the same value of z. The magnetic field thus functions as a convergent lens in the x-z plane. Conversely, if the field increases in the positive direction

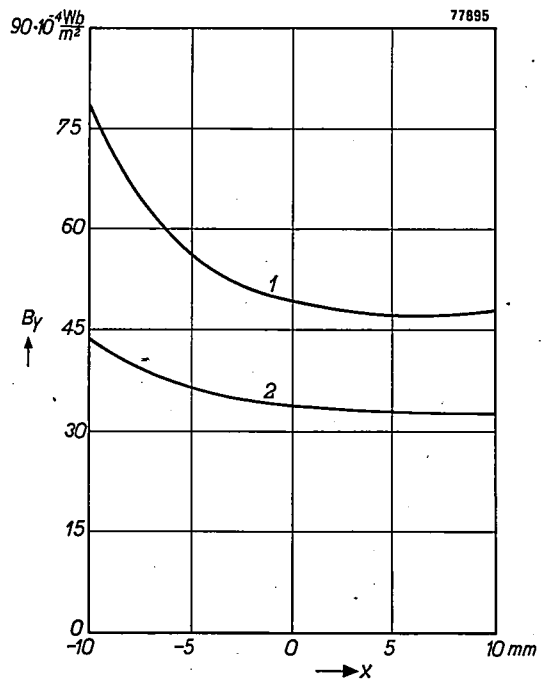


Fig. 5. Form of the field along the x-axis of an ion-trap magnet of the old type (curve *1*), and that of the improved type (curve *2*).

of x (fig. 4c), we have a divergent lens [2]). In both cases, therefore, the ion-trap magnet introduces astigmatism (different focus in the x-z plane from that in the y-z plane).

From these considerations it follows that the ideal field for the deflection of a parallel electron beam would be a field dependent on z only. It is plausible to expect that the same will apply for the deflection of the not exactly parallel beams which which occur in practice.

Curve *1* in *fig. 5* shows the form of the field along the x-axis of an ion-trap magnet of the conventional type (fig. 2). It is seen that the strength of the field is highly dependent on x. In the y-direction it is practically homogeneous in the region of the z-axis (curve *1*, *fig. 6*). This is not surprising in view of the

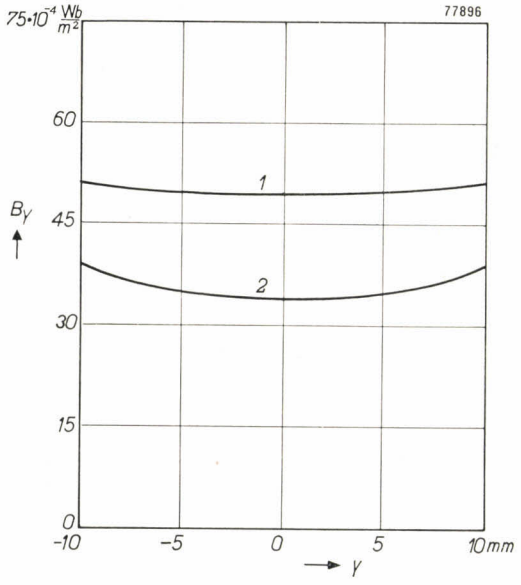


Fig. 6. Form of the field along the y-axis of an ion-trap magnet of the old type (curve *1*) and that of the improved type (curve *2*).

fact that the x-z plane is a plane of symmetry of the magnet. The electrons further from the z-axis are deflected rather more strongly than those in the centre of the beam because they move in a somewhat stronger part of the field. The result is a symmetrical distortion of the beam which, however, is so slight in practice as to be barely noticeable. Distortion due to variation in the x-direction of the field, on the other hand, is clearly perceptible.

The new ion-trap magnet

Clearly, an improvement in the ion-trap may be expected to result from a reduction in the dependence of the field of the magnet on x. As already mentioned, this has been achieved by modifying the shape of the pole pieces (*fig. 7*). It will be seen that the new pole pieces are first bent away sharply from the magnet, in contrast with the old type (fig. 2), and then gradually converge. At the point where the electron beam passes, the path length of

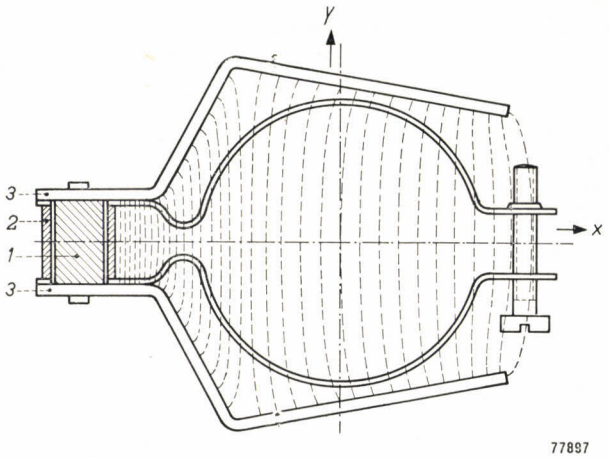


Fig. 7. Diagram of ion-trap magnet of the improved pattern showing lines of force between the pole pieces. The letters and numbers have the same significance as in fig. 2.

the lines of force now decreases slightly with increasing x; this results in a small increase in the field strength with x which compensates the reduction due to the longer path of the lines of force through the pole pieces. Curve *2* in fig. 5 demonstrates the improvement in the homogeneity of the field in the direction of the x-axis, as compared with the old type of magnet.

The distance from the permanent magnet to the centre of the electron beam is longer than in the old type, this being necessary to ensure that the bend in the pole pieces (where they begin to converge) shall be sufficiently remote from the path of the electron beam. As the lines of force emerge from the

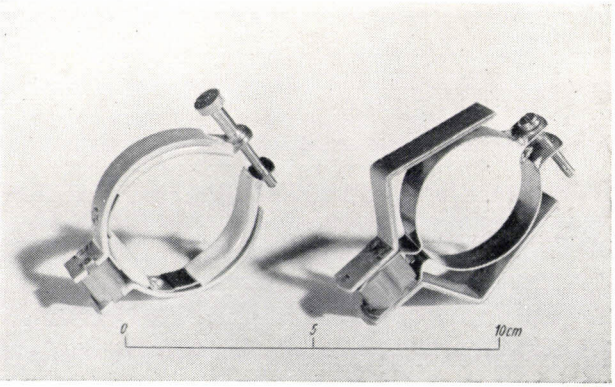


Fig. 8. Photograph showing the old type of ion-trap magnet (left) and the improved type (right).

[2]) A field that decreases with x can be changed to one that increases with x by rotating the ion-trap magnet through 180° about the y-axis.

faces of the pole pieces almost perpendicularly, it is to be expected that a concentration of the lines of force will occur on the x-axis opposite the bend in the pole pieces (fig. 7); at this point the field is not homogeneous. *Fig. 8* shows the old and the new types of ion trap side by side.

The field strength between the pole pieces is slightly less than that of the old magnet, viz, 35×10^{-4} in place of about 50×10^{-4} Wb/m², but it is found that the magnet produces an adequate deflection, even with an acceleration voltage of 20,000 V on the tube.

Results

The degree of astigmatism of the light spot on the screens of a number of television tubes fitted with the old and the new ion trap magnets was compared under otherwise similar conditions. Assessment of the quality of the light spot was made subjectively and expressed as a figure of merit of from 1 to 10. To reduce the subjectivity of the tests, they were carried out by a number of observers. An average figure of 7 was obtained for the old type of ion trap, whereas the new type received an average of 8.5, which is a considerable improvement. It should be remembered of course that the quality of the spot is dependent not only on the field of the ion trap, but also on any defects in fields of the electron gun and of the focusing and deflection coils.

In the development of the new ion-trap, valuable assistance was given by Mr. A. P. van Rooy and Mr. J. A. van Wijngaarden.

Summary. A general survey of the working of the ion-trap used in television picture-tubes reveals the fact that defocusing of the electron beam is minimized when the field of the ion-trap magnet is made as homogeneous as possible in directions perpendicular to the beam. A considerable lack of such homogeneity is found to exist in ion traps of the conventional form: a more uniform field can be obtained by modifying the shape of the pole pieces, and a significant improvement in the quality of the light spot on the screen is thus obtained.

BOOK REVIEW

Manual for the illuminating engineer on large size perfect diffusors, by H. Zijl; pp. 196, 120 figures, 49 charts — Philips Technical Library.

The direction taken by developments in illuminating engineering is such that the old-established approximation of a lighting system by "point sources" is gradually losing its significance. Approximate formulae for the illuminating of large-size light sources are now fairly well-known.

The author of this book, however, was interested in more exact results, preferably in the form of simple functions, to facilitate computations with the aid of logarithms etc. Such exact results can be obtained by working out a number of integrals, many of which can be solved only by lengthy numerical methods. Owing to exceptional circumstances, the author has been able to devote a great deal of time to this problem, with the result that his aim has largely been achieved. The results of this painstaking work are now offered in book form.

In order to limit the size of the book, the subject matter is restricted to the calculation of illumination at a given point, the average value along a given line, and the average value over a given rectangular surface. Moreover, only three idealized forms of light source are discussed: the ideal plane diffuser, the spherically symmetrical source and the cylindrical source. The luminance of these sources is assumed to be constant over their entire surface. The space to be illuminated is held to be a rectangular room, i.e. a parallelepiped, of height h and floor area $a \times b$, the light sources being assumed to be situated at a point, along a straight line, or over a rectangular surface.

To further limit the number of formulae to be dealt with, the possible configurations are reduced to a relatively small number of standardized cases (Chapter III). These standard configurations are: light sources at the corners of the ceiling, along the four edges of the ceiling, at the corners of the walls, in the ceiling itself and lastly, on one of the walls.

The average illumination is calculated for the floor area and for its edges and corners. The number of standard instances for plane sources is 18, for the spherical 11 and for the cylindrical 12.

Fifty charts (Chapter IX), together with concise instructions for their application (Chapter X), add considerably to the practical value of the work. The author demonstrates that for each set of standardized conditions, the average illumination E can be expressed as:

$$E = \frac{\Phi}{h^2 P(A, B)},$$

where Φ is the luminous flux of the light sources,

$A = a/h$, $B = b/h$, and P is a function of A and B characteristic of the standard conditions. In the charts, of which there is one for each standard case, lines of constant P have been drawn in the system of co-ordinates A, B.

The concluding chapters are partly devoted to such theoretical considerations as the "field" of the luminous flux, and the inverse square law. Practical data for the illuminating engineer are given in Chapter X, which includes a number of examples, Chapter XII, which contains a comprehensive table of utilization coefficients (calculated along the lines suggested in the book), and Chapter XIV on "Daylight".

The whole work is an interesting monograph on a subject that is likely to be of value to an increasingly wide circle of experts in the field of illumination.

A. A. KRUITHOF.

ABSTRACTS OF RECENT SCIENTIFIC PUBLICATIONS OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN

Reprints of these papers not marked with an arterisk * can be obtained free of charge upon application to the administration of the Philips Research Laboratory, Eindhoven, Netherlands.

2036: R. Loosjes and C. G. J. Jansen: Distribution anormale des vitesses des électrons émis par une cathode à oxydes en régime d'impulsions (Le Vide **26**, 1131-1135, 1952, No. 37). (Anomalous velocity distribution of electrons emitted by an oxide cathode under pulsed operation; in French).

Former experiments led the writers to the conclusion that electrons emitted from a pulsed oxide cathode must show a spread in energy of some tens to some hundreds of electronvolts. An investigation, with a tube constructed especially for the purpose, showed this conjecture to be true. Against expectation it was found that the velocity spectrum consists of only a few lines. An attempt is made to explain these phenomena. See Philips tech. Rev. **13**, 337-345, 1952 and No. **1910**.

2037: J. H. van Santen and J. S. van Wieringen: Some remarks on the ionic radii of iron-group elements. The influence of crystalline field (Rec. trav. chim. Pays-Bas **71**, 420-430, 1952, No. 3).

The ionic radii of iron group elements do not decrease monotonically from Ca towards Zn but show maxima at configurations d^4, d^5 and possibly at d^9, d^{10}. An explanation is offered based on a heteropolar model, where the influence of the other ions on the electronic distribution of $3d$-electrons is considered as a purely electrostatic influence of a preponderantly cubic crystalline field, both for negligible and strong inter-electronic interaction.

2038: M. C. Teves, T. Tol and W. J. Oosterkamp: Die Röntgen-Bildverstärkerröhre (Fortschr. Röntgenstrahlen Röntgenpraxis, Beiheft zu Bd **76** (Tagungsheft), 26-27, 1952). (The X-ray picture intensifier tube; in German).

Short communication on an X-ray intensifying tube, see Philips tech. Rev. **14**, 33-43, 1952, No. 2.

2039*: B. D. H. Tellegen: Theorie der electrische netwerken (Part III of: Theorie der wisselstromen, by G. J. Elias and B. D. H. Tellegen: P. Noordhoff, Groningen-Djakarta 1952). (Theory of electrical networks; in Dutch.)

Investigation of passive networks, characterized by linear equations with constant coefficients and composed of a finite number of elements. The properties of the networks and the methods of solution are considered, as much as possible from a physical point of view. Special applications, such as filters, are not dealt with. After an introductory chapter on network elements, the book is divided into two parts, dealing with network analysis and network synthesis respectively, at both constant and variable frequencies.

R 193: F. A. Kröger, A. Bril and J. A. M. Dikhoff: A single component white luminescent screen for television tubes (Philips Res. Rep. **7**, 241-250, 1952, No. 4).

A new luminophor (Zn, Cd)S-Ag-Au-Al has been developed that shows a strong white luminescence upon excitation by ultra-violet, cathode-rays or X-rays, and is suitable for use in direct-view television tubes. Efficiency and current saturation are similar to those of the sulphide mixtures normally used.

R 194: K. S. Knol and G. Diemer: High-frequency diode admittance with retarding direct-current field (Philips Res. Rep. 7, 251-258, 1952, No. 4).

A linear-field theory is given of the susceptance of a plane-parallel diode with negative anode voltages. The contributions due to returning electrons (total emission susceptance) and crossing electrons (exponential susceptance) are dealt with separately. The theory agrees qualitatively with known experimental results.

R 195: B. D. H. Tellegen: A general network theorem with applications (Philips Res. Rep. 7, 259-269, 1952, No. 4).

It is proved that in a network configuration, for branch currents i satisfying the node equations and branch voltages v satisfying the mesh equations, Σiv summed over all branches is zero. By this theorem it is possible to prove the energy theorem and the reciprocity relation of networks, and to show that if, at a given instant, arbitrarily varying voltages are applied to a $2n$-pole network, the difference between the electric and the magnetic energy will at any instant depend only on the admittance matrix of the $2n$-pole, and not on the particular network used for realizing it.

R 196: A. E. Pannenborg: On the scattering matrix of symmetrical waveguide junctions, III (Philips Res. Rep. 7, 270-302, 1952, No. 4).

Continuation of **R 188** and **R 190**. The theory for directional couplers obeying certain symmetry requirements is derived. Two instruments, viz. an attenuator and a standard matching transformer, both having a directional coupler as the basic unit, are described in detail. The last section deals with resonant directional couplers. (See **R 181**.)

R 197: A. H. Boerdijk: Some remarks concerning close-packing of equal spheres (Philips Res. Rep. 7, 303-313, 1952, No. 4).

For estimating the mean density in local regions of configurations of equal spheres three criteria are stated. Some configurations are described, which have in certain regions a local mean density exceeding that of close-packing. These regions may even have an infinite volume. Further it is proved that the maximum number of spheres simultaneously touching a sphere is twelve. A conjecture of Fejes concerning a fourteenth sphere added to this configuration is shown to be false.

BOOK NOTICES

Strain gauges: theory and application; by J. J. Koch and R. G. Boiten (T.N.O.), and A. L. Biermasz, G. P. Roszbach and G. W. van Santen (Philips), pp. 103, 66 figures. — Philips Technical Library.

This book is directed to the user of strain gauges for the investigation of mechanical stresses. In compact form, it gives a thorough description of the measuring techniques employed with this important new device. The first chapter describes the construction and properties of strain gauges, while the second deals with their associated measuring instruments. The practical problems of cementing and connecting the gauges are then considered, followed by a chapter on the interpretation of results. Chapter V is concerned with the theory of stresses and theories of failure. The final chapter discusses the use of strain gauges for the measurement of quantities which are reducible to mechanical strain, such as force, pressure, acceleration etc.

This book, based on many years experience, may be said to cover all the essentials which need to be understood for the succesful employment of strain gauges. Its publication is one of the fruits of co-operation between the Netherlands Industrial Organization for Applied Scientific Research (T.N.O.) (Stress and Vibration Research Group) and the Philips organization.

Remote control by radio: an amplitude-modulation and a pulse-modulation system; by A. H. Bruinsma; pp. viii + 96, 43 diagrams, 6 plates. Philips Technical Library, popular series.

Rather than giving a survey of all existing forms of remote control by radio, the author has restricted himself to the treatment of two systems which have been tried out in practice and have given complete satisfaction. The first is the more or less conventional method using amplitude modulation, and having two independent channels. A second, more versatile system is described, based on pulse modulation, which permits the simultaneous transmission of eight independent signals over one carrier-wave. Finally, two demonstration boats are described which are equipped with the systems discussed. Full details and characteristics of the valves employed are given in an appendix.

VOL. 15 No. 10, p.p. 265-296 APRIL 1954

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
THE PHILIPS INDUSTRIES

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, NETHERLANDS

PU 65/60 518

A COLD CATHODE GAS-DISCHARGE TUBE AS A SWITCHING ELEMENT IN AUTOMATIC TELEPHONY

by J. DOMBURG and W. SIX.

621.387:621.385.12:
621.318.57:621.395.34

*The idea of replacing electro-mechanical contact devices in automatic telephone exchanges by electronic ones, is anything but new. In patent literature throughout the last few decades, a number of such proposals can be found, involving the application of photocells, cathode-ray tubes or other special vacuum tubes, transistors, gas-discharge tubes, etc. Vacuum tubes specially designed for use as switching elements in telephony were described in this Review some years ago *). In the following article a gas-discharge tube for the same purpose will be discussed. By means of this tube it is possible to built up the relay, selector and other circuits used in automatic telephony.*

Electro-magnetic switches are used on a large scale in automatic telephony for establishing and breaking connections — both line connections between subscribers, and connections between a variety of auxiliary equipment (for selecting, ringing current, busy-tone, call registration, etc.). A certain mechanical delay is always associated with such switching devices (relays, selectors) due to their inertia. The same applies to the controlling devices, which determine the positions of the switches. In modern installations, not every selector has its own controlling devices, but these are available centrally and cover a number of lines. The smaller the inertia of these control devices, the sooner they are again available for another subscriber and the fewer of them are required. For a considerable time, attempts have been made to develop electronic devices — which are practically inertia-free — for control devices and for other purposes.

A second and fundamental advantage of electronic switching devices is that unlike relays, they require no maintenance. Present developments in electrical engineering offer, in principle, a number of ways of achieving electronic switching. In this article we shall confine ourselves to one of these possibilities, viz. the use of cold-cathode gas-discharge tubes, which have been developed in the Philips Research Laboratory at Eindhoven. After a description of the tubes, we shall discuss a number of circuits in which they can be applied. Several of these circuits have been incorporated in an experimental all-electronic telephone exchange [1]).

The cold cathode gas-discharge tube

The tube in question is made in two versions, type Z 500 T and Z 501 T (*fig. 1*). Each contains a cathode and an anode but the former contains, in addition, one trigger electrode, and the latter two trigger electrodes, which serve to initiate the discharge. The tubes are filled with argon.

The discharge is a glow-discharge which spreads over the cathode according to the current intensity. At a current of 6 mA the cathode is completely covered by the discharge.

The cathode consists of a small nickel plate coated with a layer of activated barium oxide. The barium oxide is activated by partly decomposing it, by passing current pulses through the tube (a *getter* binding the oxygen released), until the lowest possible working voltage is obtained. The tubular-shaped anode and the wire trigger electrode(s) are also made of nickel and are mounted parallel to the cathode. The tube has been so designed that the greatest possible difference is

*) J. L. H. Jonker and Z. van Gelder, Philips tech. Rev. **13**, 49-54 and 82-89, 1951.

[1]) See also W. Six, Communication News **14**, 58-69, 1954 (No. 2).

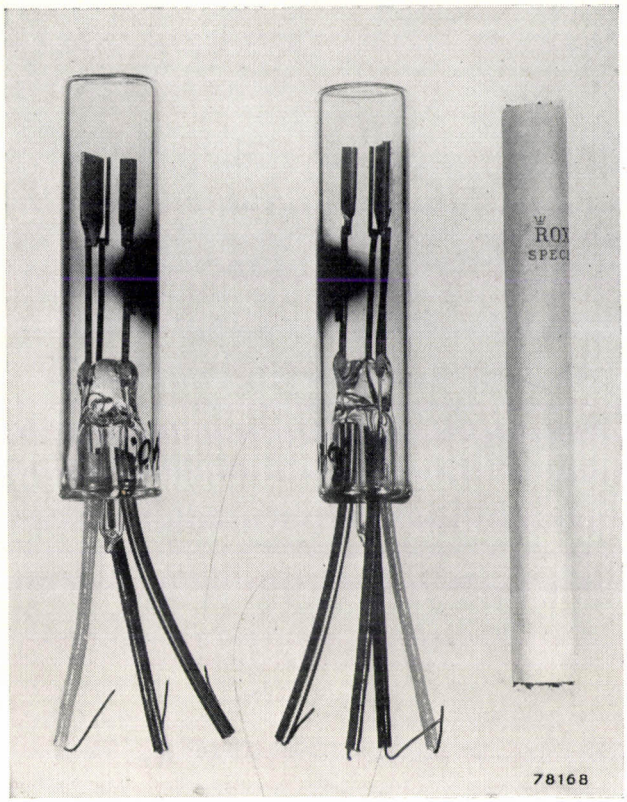


Fig. 1. Cold-cathode gas-discharge tubes with oxide-coated cathodes. Left, type Z 500 T with one trigger electrode. Right, type Z 501 T with two trigger electrodes. The tubes are without bases, and connections are soldered direct to the lead-in wires. Diameter 12 mm, length approx. 55 mm.

obtained between the breakdown voltage and the cathode-anode working voltage; thus the permissible tolerances of supply voltages, resistance values, etc. are fairly wide.

For a given gas, the breakdown voltage is determined by the product of the density p of the gas and the electrode distance d, according to the Paschen curve (*fig. 2*). The working voltage, on the other hand, depends very little on pd, as long as pd does not exceed a value at which *anode fall* occurs — a state which is undesirable anyway, due to the risk of relaxation oscillations. The values p and d have therefore been so chosen that pd is situated a good deal to the right of the minimum in the Paschen curve, though not so far that an anode fall can occur. The breakdown voltage ($V_{\text{ign}\,a}$) is then 175 V or higher. Owing to the oxide-cathode, the running voltage is kept down to approximately 60 V.

The trigger electrode is at a distance $d' < d$ from the cathode, such that pd' corresponds to the minimum in the Paschen curve. The ignition voltage of the trigger electrode has thus the lowest possible value, viz. 65-78 V.

Once a discharge has been initiated via the trigger electrode, the presence of ions and electrons in the tube reduces the breakdown voltage $V_{\text{ign}\,a}$ of the anode by an amount depending on the current i_a' flowing through the trigger electrode (*fig. 3*).

Fig. 3 shows that $V_{\text{ign}\,a} = 160$ V at $i_a' = 30$ μA. Thus, if a direct voltage of 160 V is applied between anode and cathode, the tube will ignite only if the trigger electrode current is at least 30 μA. A current pulse of only short duration is sufficient for this, since the main discharge can be initiated within an interval of 10^{-4}-10^{-5} sec. Once initiated, the discharge continues, even if the current i_a' drops to zero. It is therefore possible to supply i_a' from a capacitor, as shown in *fig. 4*: as soon as any current passes through the trigger electrode, this current will temporarily assume a high value, since the capacitor is then discharged via a circuit having only a low resistance. If a positive bias voltage is employed on the trigger electrode, only a low ignition voltage is required to fire the tube.

As mentioned above, the tube type Z 501 T is provided with two trigger electrodes. It will later be demonstrated that this is of special advantage for tubes intended for service as selector switches.

Since the tube has no filament, it generates heat only during actual operation, and even then at the rate of no more than a fraction of a watt and, as a rule, for short periods only. There are, therefore, two advantages compared to thermionic-emission tubes: low power consumption, and the absence of any cooling problems (particularly important in compact installations with a large number of tubes).

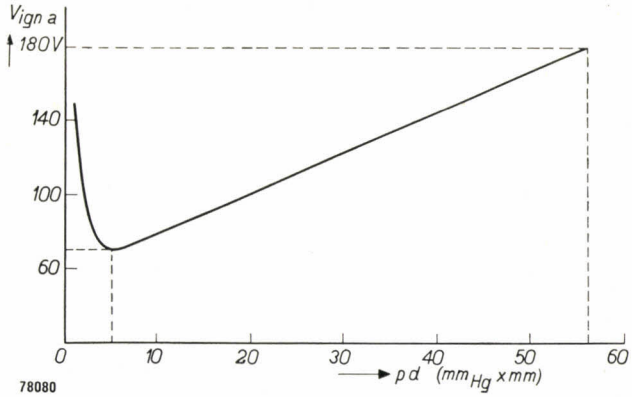


Fig. 2. The Paschen curve (ignition voltage $V_{\text{ign}\,a}$ as a function of the product of gas density p and electrode distance d), for argon, with an oxide-coated cathode.

As a third advantage we might add that these tubes have a very long life. A working life ("nett" life) averaging 12 000 hours at a current of 6 mA can be attained, and, according to prevailing standards, a life of 6000 hours can be expected for normal production-line tubes. Since the majority of the tubes in the circuits to be discussed presently,

are only operating for a fraction of the time, this means that these tubes will have a "gross" life of centuries. At currents lower than 6 mA, the life of these tubes is inversely proportional to the

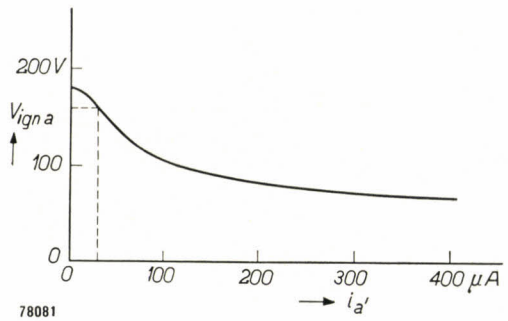


Fig. 3. Ignition voltage $V_{ign\,a}$ of the tube Z 500 T as a function of the trigger electrode current i'_a.

current; at currents higher than 6 mA, their life is inversely proportional to rather more than the cube of thecurrent [2]).

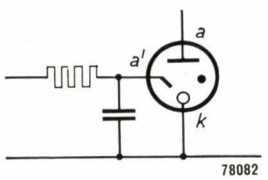


Fig. 4. A circuit for igniting tube Z 500 T (with anode a, cold cathode k and trigger electrode a'; the dot indicates the gas filling). As soon as the capacitor is charged up to the ignition voltage of a' (65 to 78 V), it discharges across k-a' with a fairly large current, which initiates the main anode current k-a.

Circuits for the cold-cathode tube

On ignition, the voltage between anode and cathode drops from the initial value V_{ak} — which, in order to avoid spontaneous ignition, should be below 175 V — to the working voltage (approx. 60 V). The difference V_{ak} — 60 V is then developed across a resistor in series with the tube and can, in turn, be used for firing another tube.

It has already been mentioned that the trigger electrode cannot serve for extinguishing a tube already ignited. The question is, how can this be accomplished? In some cases there is no objection against a short interruption of the anode circuit. In other cases the supply voltage can be periodically reduced to below the working voltage, so that the discharge is terminated; this may be effected by a supply voltage composed of a direct voltage of 100 V, upon which an alternating voltage of amplitude 60 V has been superimposed. The discharge can, furthermore, be extinguished by reducing the tube voltage to below the working voltage by means of a pulse: either a negative pulse to the anode, or a positive pulse to the cathode. Still another possibility is to arrange the circuit in such a way that it functions as an oscillator as long as the potential of the trigger electrode is high enough; when the potential applied to the trigger electrode is brought down, the oscillation stops.

Examples of these various possibilities will be encountered in the circuits to be dealt with below. As a practical test and also for demonstration purposes, several of these circuits have been applied in an experimental automatic telephone exchange for ten lines (*fig. 5*), which has already been functioning for some years in the Philips Research Laboratory in Eindhoven. Neither mechanical switches nor relays are to be found in this fully electronic exchange; the tubes employed, apart from three thermionic amplifying tubes (ECC 81), are exclusively cold-cathode tubes of the types Z 500 T and Z 501 T.

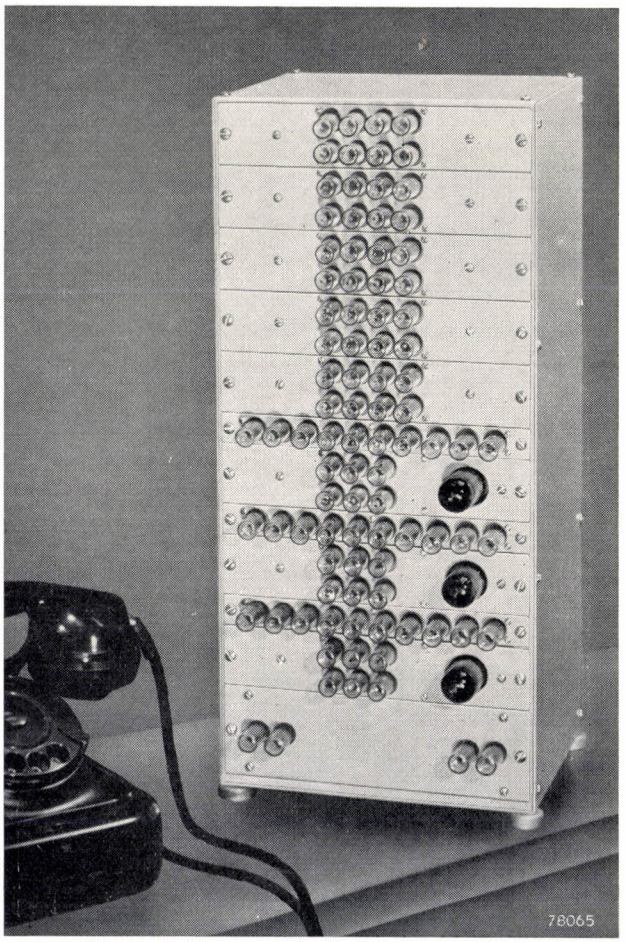


Fig. 5. Fully-electronic telephone exchange for 10 lines, in use in the Philips Research Laboratory in Eindhoven. This exchange contains 62 cold-cathode tubes Z 500 T, 30 cold-cathode tubes Z 501 T, and 3 double-triodes ECC 81. The power is supplied by two D.C. power packs having potentials of +160 V and −60 V with respect to earth.

[2]) G. H. Rockwood, Current rating and life of cold cathode tubes, Trans. Amer. Inst. El. Engrs. **60**, 901-903, 1941.

Relay circuits

Fig. 6a shows a cold-cathode tube, connected in series with an anode resistor R_a and a cathode resistor R_k, to a 160 V source of direct voltage. A voltage divider r_1-r_2 is conneted between the positive terminal and the cathode. It is assumed that $R_a = R_k$, and that $r_1 + r_2 \gg R_k$, so that, if no current passes through the tube, the cathode potential is practically zero. The ratio r_1:r_2 has been so chosen that the tapping point *1* has a potential of +60 V.

When the tube is burning, the voltage between anode and cathode drops from 160 V to 60 V, and the difference of 100 V is equally divided between R_a and R_k. It can be readily verified that the point *1* will then assume a potential of approximately +91 V (fig. 6*b*). The voltage surge at *1*, from 60 V to 91 V, can be used for igniting or extinguishing another tube.

This voltage surge can be somewhat increased by using for r_2 a "varistor" (fig. 6*c*), i.e. a resistor whose resistance increases at decreasing voltage. In this way point *1* can be given a potential of +102 V after the tube has been ignited.

These principles have been applied in the relay circuit shown in *fig. 7a*. As long as the tube does not pass a current, point *1* has a potential of +60 V and point *2* a potential of +102 V; after the tube has ignited, point *1* has a potential of +102 V and point *2* a potentials of +60 V. Ignition thus result in the potentials of points *1* and *2* being reversed; in this respect the circuit is equivalent to a relay with a change-over contact (fig. 7*c*). Automatic quenching of the discharge can be effected by supplying the circuit with a pulsating direct voltage dropping periodically below the working voltage (fig. 7*b*) as described above.

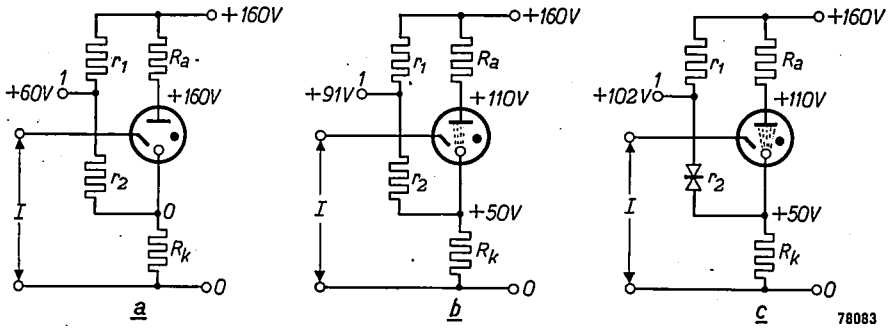


Fig. 6. The cold-cathode tube used as a relay.
a) The tube is not conducting. Point *1* on the voltage divider r_1-r_2 is at a potential of 60 V.
b) The tube has been ignited via the input terminals *I*. The potential at *1* increases to 91 V.
c) By using a varistor for r_2, the potential of *1* can be raised even higher, up to 102 V.
The resistances of r_1 and r_2 are assumed to be high compared to the values of R_a and R_k (= R_a).

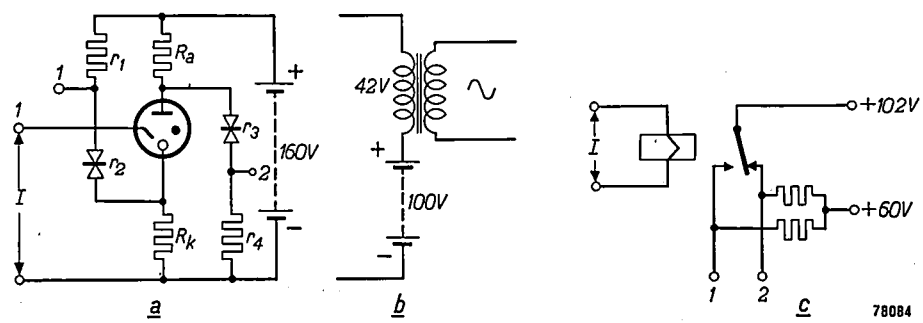


Fig. 7. *a*) Extension of the circuit of fig. 6*c* by a branch r_3 (varistor) - r_4. Tube non-conducting: point *1* at 60 V, point *2* at 102 V. Tube ignited: point *1* at 102 V, point *2* at 60 V.
b) Automatic extinguishing of the tube can be effected by using a supply of 42 V r.m.s. alternating voltage superimposed on a 100 V direct voltage.
c) Relay with a change-over contact *1-2*, of which (*a*) is the electronic equivalent.

It has been demonstrated by Ingram [3]) that by

[3]) S. B. Ingram, Cold-cathode gas-filled tubes as circuit elements, Trans. Amer. Inst. El. Engrs. 58, 342-346, 1939.

combining gas-discharge tubes and rectifier elements, an equivalent for every relay circuit can be found. *Fig. 8* shows some examples of equivalent switching circuits: *a*) two make-contacts in series, *b*) two make-contacts in parallel, *c*) a make-contact and a break-contact in series, *d*) a make-contact and a break-contact in parallel and a make-contact in series with the two. Explanations of these circuits are given below the figure.

Counting circuit

Decade pulse-counting circuits can also be built up with cold-cathode tubes. It will later be demonstrated how such counting circuits can be applied in automatic telephony, but for the time being, only counting in a general sense will be considered. A counting circuit is shown in *fig. 9*. The pulses to be counted, which must be of approx. 30 V amplitude, are applied, with positive polarity, via the input terminal *I* and the capacitors C_1, to the trigger electrode of the ten tubes *1*, *2*, ... *0* of the first decade (D_1). In the initial position, only the trigger electrode of tube *1* has a bias (+60 V), derived from a voltage divider R_3-R_4 across the 160 V supply voltage. As a consequence, the first pulse will ignite only tube *1*, whose cathode potential will then rise from 0 to approx. 100 V, bringing the potential of the trigger electrode of tube *2* to the same

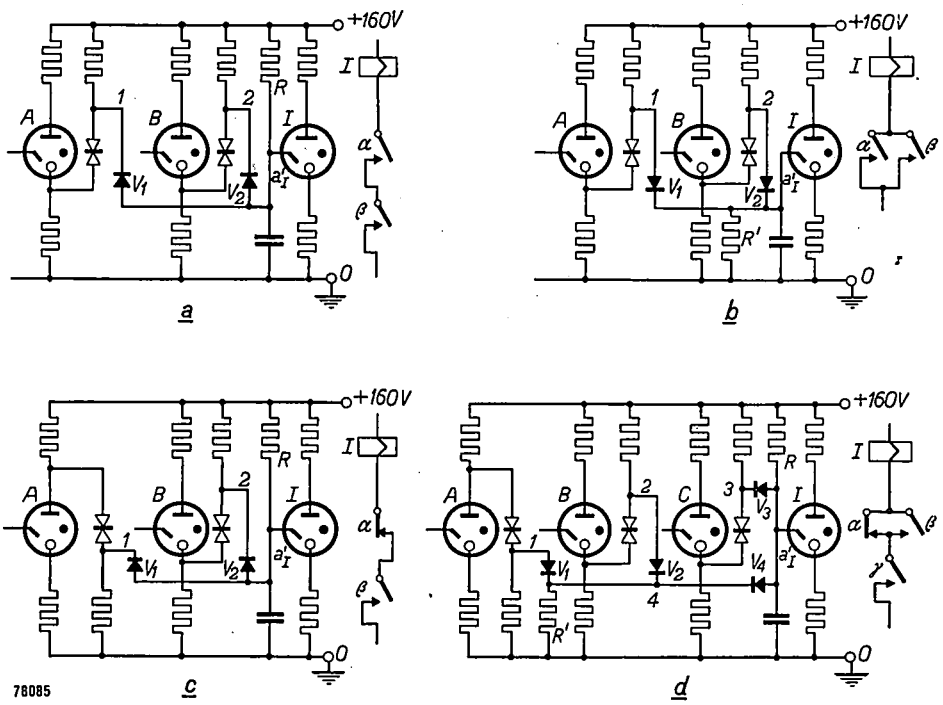


Fig. 8. Various relay circuits with cold-cathode tubes (derivatives of fig. 7*a*). Whether tube *I* ignites depends on the condition of tubes *A* and *B* and *C*, in a manner analogous to the energizing of the adjacent relay *I*, which depends on the position of the contacts α and β and γ.

The resistance of *R* is large compared to the reverse resistance of the rectifier elements, and that of *R'* (in *b* and *d*) is low compared to *R*. The capacitor between the trigger electrode of *I* and earth, is the trigger condenser mentioned in fig. 4.

Further particulars can be found in the adjacent table (except that for circuit *d*, not all possibilities are given). 0 signifies non-conduction, and × conduction, of a tube or rectifier.

Circuit	Tubes A	B	C	Potential at point 1 (volts)	2	3	4	Rectifier elements V_1	V_2	V_3	V_4	Potential a'_I (volts)	Tube *I*
a	0	0		60	60			×	×			60	0
	×	0		100	60			0	×			60	0
	0	×		60	100			×	0			60	0
	×	×		100	100			×	×			100	×
b	0	0		60	60			×	×			60	0
	×	0		100	60			×	0			100	×
	0	×		60	100			0	×			100	×
	×	×		100	100			×	×			100	×
c	0	0		100	60			0	×			60	0
	×	0		60	60			×	×			60	0
	0	×		100	100			×	×			100	×
	×	×		60	100			×	0			60	0
d	0	0	0	100	60	60	100	×	0	×	0	60	0
	0	0	×	100	60	100	100	×	0	×	×	100	×
	0	×	×	100	100	100	100	×	×	×	×	100	×
	×	0	×	60	60	100	60	×	×	0	×	60	0
	×	×	×	60	100	100	100	0	×	×	×	100	×
	etc.												

potential. After the ignition of tube *1*, a current passes through the cathode resistor R_1 of tube *1* and through the resistor R_2, common to all the ten tubes. This results in a voltage drop of 40 V across R_2, so that the cathode potential of the nine non-conducting tubes becomes 40 V. The trigger electrode of tube *2* thus assumes a bias of 100 — 40 = 60 V with respect to the cathode. The second pulse will therefore ignite tube *2*. The cathode potential of this tube then jumps from 40 to 100 V, which causes a charging current through the capacitor C_2 connecting the cathodes of tubes *1* and *2*. Part of this charging current flows through the cathode resistor R_1 of tube *1*, thus temporarily raising the latter's cathode potential to such an amount that this tube is extinguished (an example of quenching by a positive pulse applied to the cathode).

After the second pulse, the state of the circuit is as follows: only tube *2* is burning; its cathode potential is +100 V, while that of all other tubes is +40 V; the trigger electrode of tube *1* has a bias of 60 — 40 = 20 V (which is insufficient to allow firing of tube *1* by a subsequent pulse); the trigger electrode of tube *3*, however, has a bias of 100 — 40 = 60 V, while the remaining trigger electrodes have no bias at all. The third pulse will therefore ignite tube *3* only (and in doing so extinguish tube *2*), and so the process will continue.

For counting beyond 10, so that the eleventh pulse will again ignite tube *1*, etc., a connection similar to those between the other tubes is required between tube *0* and tube *1*. For this, however, a rectifier element (V) is required, connected so that its conduction direction is from point *b* to point *a*. This is essential in order to prevent point *a* from assuming the potential of point *b* (zero) when all tubes are non-conducting (the resistances of R_1 and R_2 are considerably smaller than those of R_3 and R_4); if this happened the trigger electrode of tube *1* would lose its bias. On firing, the cathode potential of tube *0* jumps from 40 to 100 V, and part of this voltage surge is fed to point *a* via the rectifier element. From here, it can be fed to a second decade (D_{10}) for counting the tens. At the tenth pulse therefore, not only tube *0* of D_1, but also tube *1* of D_{10} will fire. A third decade (D_{100}) can be connected for counting the hundreds, and so on.

Since the discharge in the tubes is visible as a violet light, the number of pulses counted can be directly read if the tubes are provided with figures.

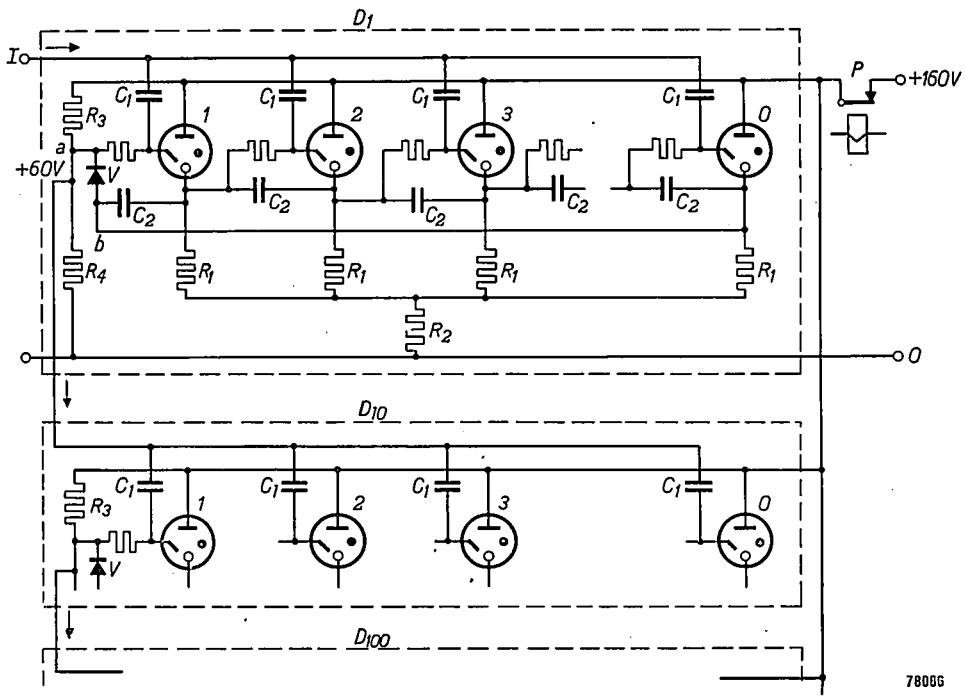


Fig. 9. Decade-counter employing cold-cathode tubes. D_1 units decade, D_{10} tens decade, D_{100} hundreds decade, each containing 10 tubes Z 500 T (*1*, *2* ... *0*). *I* input terminal. C_1 input capacitors. R_1 cathode resistors. R_2 common cathode resistor. C_2 quenching capacitors. R_3-R_4 voltage divider, from which the +60 V trigger electrode bias of the first tube is derived. V rectifier element through which the tubes of the next decade are ignited. *P* re-set contact, which is opened momentarily after counting, to extinguish all tubes.

A description of a special decade counting tube (type E1T) and the counting circuit designed for it, was published recently in this Review [4]). It may be of interest to draw a brief comparison between this tube and the tube Z 500 T.

Some characteristic advantages of the E1T are:

1) The maximum counting speed is far greater.
2) The tube E1T in itself constitutes a decade, so that one tube E1T (together with an auxiliary valve E 90 CC) takes the place of 10 tubes Z 500 T.

As regards item 1), the counting circuit of the E1T is capable of counting speeds of at least 30,000 pulses per second; far greater speeds can be attained with special circuit adjustments to suit individual tubes, and with the aid of certain auxiliary tubes. For the gas-discharge tube, the counting speed is, in principle, limited, firstly by a delay depending on the number of electrons (mainly photo-electrons) fortuitously present in the tube just before ignition, and furthermore by the period required for building up the starting discharge and the main discharge, and finally by the de-ionization (quenching) time. The maximum counting rate lies somewhere between 1500 and 2500 pulses per second (which is considerably more than is required for telephony purposes).

On the other hand, the cold cathode tube Z 500 T shows the following advantages:

1) no filament current is required;
2) the "gross" life is longer;
3) less heat is generated;
4) the supply voltage is lower (160 V compared to 300 V);
5) a counting circuit with tubes Z 500 T not only gives a visual indication of the counted number, but also gives an indication in terms of voltages: the anode or cathode potential of ignited tubes is different from those of non-ignited tubes; these marking voltages (see below) can be used, for example, for the ignition of other tubes.

The above considerations show that the existence of both types of tubes is justified.

To reset the circuit after a count, the tubes can be extinguished by a short interruption of the supply voltage, for example, by a relay-contact (P, fig. 9).

Register circuit

With certain automatic telephone systems, the numbers dialled by a subscriber are temporarily recorded by means of switches or relays in "registers"; after the whole number has been dialled, the registers determine the positions to be taken up by the selectors in order to establish the connection with the called subscriber. All the electro-mechanical devices usually employed for this purpose can be replaced by cold-cathode tube circuits. One decade of the counting circuit described above then serves for both counting and marking. The group of registers comprises a number of decade registers, one for each digit in the telephone number. The counting-and-marking circuit counts the number of pulses generated each time a digit is dialled, marks the corresponding tube of the relevant register, so that this tube is also ignited, and then returns to its initial position in order to count the pulses of the next digit. This counting-and-marking circuit will from now on be referred to, for short, as a "counter".

Consider the example of a telephone system with three-digit numbers. A subscriber wishes to be connected to number 382. The number 3 is first dialled. The returning dial produces three pulses, so that tube *3* of the counter will be left ignited. This causes tube *3* of register *I* to ignite, whereupon tube *3* of the counter extinguishes. (The manner of doing this will be described later.) The subscriber now dials 8: tube *8* of the counter ignites, tube *8* of register *II* ignites, and tube *8* of the counter is extinguished. A similar procedure takes place with the third digit. As a result, the selected number is recorded by the burning of tube *3* in register *I*, tube *8* in register *II* and tube *2* in register *III*.

All this can be effected by the circuit represented in *fig. 10* (for simplicity, only two registers are shown; more registers, however, can simply be added to this diagram). The counter (C) is similar to the decade circuit shown in fig. 9, with the exception that there is now no connection between tube *0* and tube *1*, since the counter is not required to count more than 10 pulses. If three pulses are applied to the counter, then, as described in the previous section, tube *1* is ignited by the first pulse, tube *2* by the second pulse (whereupon tube *1* extinguishes) and tube *3* by the third pulse (whereupon tube *2* extinguishes); the ignition of a tube means that its cathode reaches a potential of approx. 100 V, the marking voltage. The cathode of each counter tube is connected to the trigger electrode of the corresponding tube in register *I*, so steps must be taken to ensure that while the prolonged burning of tube *3* (in this example) causes registration in *I*, the transient burning of the counter tubes *1* and *2* are *not* recorded in register *I*. The distinction between these burning periods is made with the aid of the time-delay networks R_5-C_3. The effective delay is greater than the interval between two consecutive pulses, so that the trigger electrodes of the register tubes *1* and *2* cannot attain the ignition voltage (65-78 V); the trigger electrode of tube *3*, on the other hand, has ample time in which to do so.

The counter must now be returned to its initial state in order to be ready to deal with the next number dialled, i.e. the counter tube *3*, which is still burning, has to be extinguished. This is done by the action of an auxiliary tube (E). When a tube

4) A. J. W. M. van Overbeek, J. L. H. Jonker and K. Rodenhuis, A decade counter tube for high counting rates, Philips tech. Rev. **14**, 313-326, 1953 (No. 11).

of register I is ignited, a current passes through resistor R_7, which raises the potential of point p from 0 to 50 V. This voltage surge passes through the capacitor C_5 and ignites the tube E, the trigger electrode of which already had a 60 V bias, derived from a voltage divider R_8-R_9. Incorporated in the cathode circuit of tube E is a resonant circuit (L_0-C_4, with a resistor in parallel to C_4). On ignition of tube E, the condenser C_4 is charged through L_0. (i.e. tube *3* in register I is burning), while the counter and the auxiliary tube E have returned to their initial condition.

The subsequent digit has to be recorded in register II. This is effected in the following manner. The burning tube of register I produces a 50 V voltage drop across resistor R_7, and this is used to render register II sensitive and register I insensitive to whatever marking voltage occurs in the counter

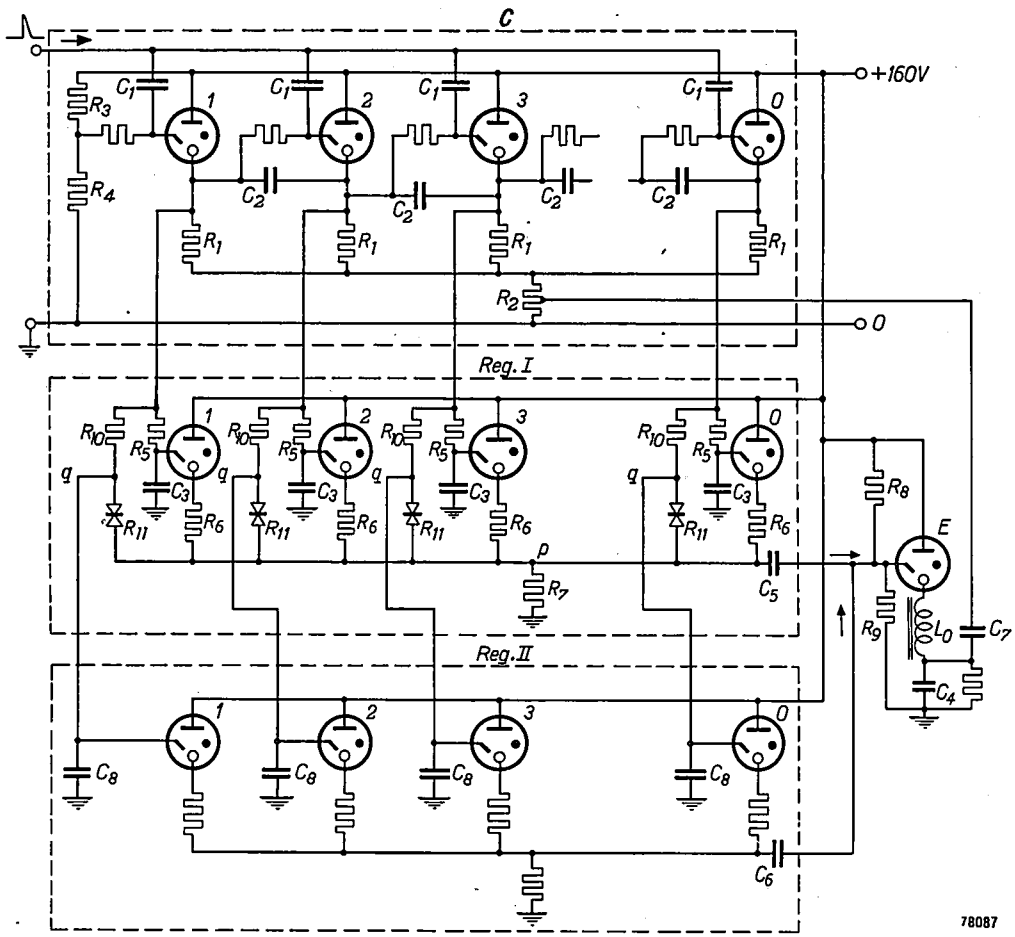


Fig. 10. Counting and marking circuit (C) with two registers (*Reg. I* and *Reg. II*). The circuit of C is almost the same as that of decade D_1 in fig. 9. E quencher tube, which, after a digit has been recorded in a register, extinguishes the tube burning in C. R_5-C_3 delay networks, preventing the transient burning of the tube in C influencing *Reg. I*. R_{10}, R_{11} and C_8 constitute similar delay networks for *Reg. II*.

The voltage across C_4 is applied through a coupling capacitor (C_7) to a tapping on the resistor R_2; this voltage increases sufficiently to extinguish the burning counter tube. Tube E itself is extinguished when the voltage across it drops below the sustaining voltage. This happens because the inductance L_0 prolongs the flow of the condenser charging current after the voltage across the tube has dropped to its sustaining voltage. The net result is thus that the first digit, 3, is recorded in register I during the dialling of the second digit. In fig. 10 it is seen that the trigger electrodes of register II are connected to the tappings q of the voltage dividers R_{10}-R_{11}, the lower part of which (R_{11}) are varistors. These voltage dividers have been so proportioned that, at a marking voltage of 100 V, point q assumes a potential of 60 V or 90 V, according to whether the potential of the point p is 0 (initial state) or 50 V (already registering) respectively.

Now consider again the counting of the pulses

of the first digit by C (fig. 10). The potential of point p is then still zero, and the potential of 60 V assumed by the marked point q is not sufficient to ignite the corresponding tube of register *II*, but the full marking voltage is available to ignite the corresponding tube of register *I*. The first digit is thus recorded in register *I*, but not in register *II*.

When the second digit is being dialled, the point p, and therefore also the cathode of the non-burning tubes of register *I*, are already at a voltage of 50 V. The result of this is that this register is blocked for the second digit. The voltage at the marked point q is now not 60 V, but 90 V, so that the second digit will be recorded in the second register. The auxiliary tube E is then again fired by a pulse (via C_6), and operates as before to return the counter to its initial condition.

The capacitors (C_8) between the trigger electrodes and earth ensure that the transient burning of the counter tubes cannot influence the register.

The cold-cathode tube in speech-current circuits

A speech current may be superimposed on the direct current of a few milliamps flowing through the cold cathode tubes. The gas discharge constitutes a certain impedance to this alternating current. This impedance, by analogy with the contact resistance of a relay contact, should be kept as low as possible. The impedance depends on the frequency, and on the intensity of the direct current and, with regard to the tube, on the gas filling, the geometry of the electrodes and the work function of the cathode material [5]). The oxide cathode results in a low value of the work function, which gives a low impedance. At a current of 6 mA and frequencies between 300 and 3000 c/s, the tubes have an impedance of approximately 500 Ω. Such a value may be considered low for a gas discharge, but is still fairly high when compared with the contact resistance of a relay contact. In order to keep the attenuation caused by the discharge impedance as low as possible, it is necessary that the impedance of the circuit in which the tube functions, is increased as much as possible without giving rise to "crosstalk". If the impedance of the circuit can be stepped up in this way to say, 2400 Ω, the attenuation caused by the tube will only be 1 dB.

Engaged-test and blocking

In automatic telephony it is a common occurrence that a free line has to be selected from a bundle of lines. Two conditions as regards "blocking" then have to be satisfied:

1) Once a free line has been found, no subsequent connection with a second free line must be made.
2) Only one connection must be possible with the free line found.

Using electro-mechanical devices, this selection generally proceeds by a switch scanning the lines one at a time, and stopping upon finding the first free line. Electronic devices, on the other hand, are capable of a *simultaneous* examination of all lines, whilst at the same time meeting the blocking conditions. The time saved by this simultaneous examination may result in a saving in equipment.

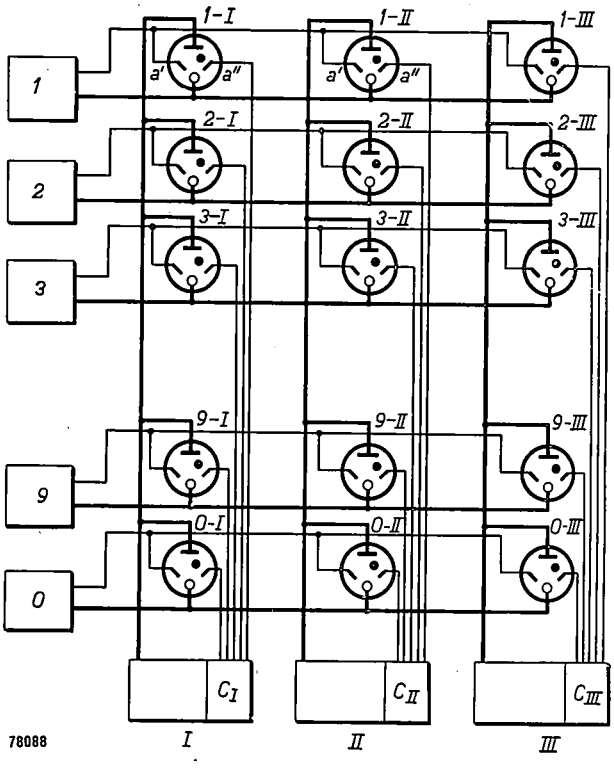


Fig. 11. Block diagram of an electronic telephone exchange for 10 lines. *I*, *II*, *III* connecting-circuits, each with a counting and marking circuit C_I, C_{II}, C_{III} (like C in fig. 10). *1-I*, *1-II*, ... *0-III* are 30 cold-cathode tubes Z 501 T.

Fig. 11 shows the principle of an electronic circuit having the above-mentioned properties, for a small, ten-line exchange. To the left are the subscribers' sets, numbered *1*, *2*, ... *0*. Within the exchange, the connections between the subscribers can be established via connecting-circuits — in this case there are three (*I*, *II*. and *III*) — by means of $3 \times 10 = 30$ cold-cathode tubes, placed at the junctions of the subscribers' lines and the connecting-circuits. Here the tubes with two trigger electrodes (type Z 501 T) can be used with advantage.

[5]) See C. van Geel, Untersuchungen von Gasentladungen mit Rücksicht auf ihre dynamischen Eigenschaften und ihre Stabilität, Physica **6**, 806-816, 1939.

As an example, suppose subscriber *2* wishes to telephone subscriber *9*. When the receiver is lifted from its cradle, one of the speech current tubes *2-I*, *2-II* or *2-III* is ignited, and the subscriber *2* becomes connected to one (and only one) of the connecting circuits. Once this connection has been established, this connecting circuit must be blocked for other calls. Should all the connecting circuits be in use at the moment the receiver was lifted off the cradle, then the subscriber should hear the busy-signal.

From the above description it will be clear that first a horizontal row of tubes (row *2*) has served as a pre-selector and then a vertical column (column *II*) as a final selector. Any of the thirty tubes must therefore be suitable for both functions. In order that the tubes may play both roles, tubes with two trigger electrodes have been employed; the trigger electrodes a' function if the tube is used as a pre-selector, and the trigger electrodes a'' function if the tube is used as a final selector. (This is not absolutely necessary; tubes having one trigger electrode can also be used, but in that case each tube requires, among other things, a rectifier element whose resistance in both conduction directions has to be accurate within fairly narrow limits. Consequently, tubes having two trigger electrodes are preferred here.)

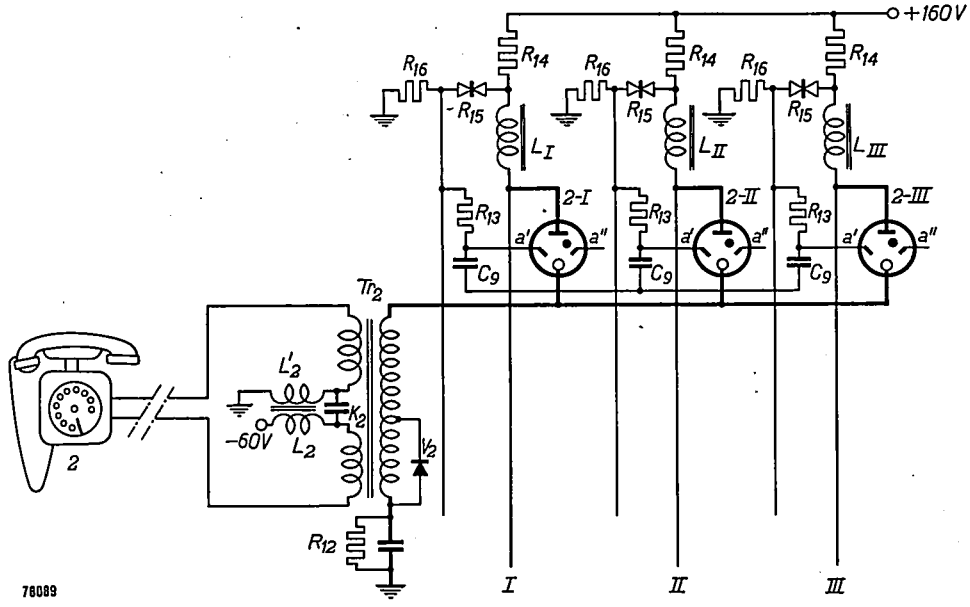


Fig. 12. Part of the circuit of fig. 11, further developed and slightly modified. Tr_2 transformer coupling the (symmetrical) line of subscriber *2* to the (asymmetrical) connecting circuits *I*, *II* and *III*. The choke coils L_I, L_{II} and L_{III} prevent the 160 V battery presenting a short-circuit to the speech current. The voltage dividers R_{15} (varistor)-R_{16} apply a +60 V bias to the trigger electrodes a'. If the subscriber is connected up to one of the free connecting-circuits, a voltage drop occurs in the resistor R_{14} of the active circuit and in the common resistor R_{12}. This voltage drop is such that the remaining circuits are blocked to him, and also the active connecting-circuit becomes inaccessible to the other subscribers, whilst his own line is marked as being engaged. The capacitors C_9 ensure that the ignition of the auxiliary discharge via a' is sufficient for temporarily preventing more than one tube igniting at the same time. The coils L_2-L_2' keep the speech current out of the 60 V battery and constitute a step-up autotransformer of ratio 1 : 2 for the ringing voltage (see under *Ringing current*). The capacitor K_2 and the one in parallel with R_{12} provide paths for the speech current. The rectifier element V_2 limits the negative voltage pulse which occurs when the subscriber lifts the receiver.

After the connection to one of the connecting circuits, e.g. to circuit *II*, has been established, subscriber *2* dials number 9. The pulses thus produced are counted by counter C_{II} of connecting circuit *II*, so that tube *9* in this counter ignites. This causes tube *9-II* to be ignited, which serves as a selector contact and establishes the connection between subscriber *2* and subscriber *9*. If subscriber *9* should be engaged, then subscriber *2* must hear the busy-signal; otherwise the bell of set *9* must ring. The production and circuiting of the busy-signal and the ringing current will be dealt with later on in this article.

Fig. 12 demonstrates in greater detail the procedure of selecting a free connecting circuit. All 30 tubes — only three are shown — are supplied with

a direct voltage of 160 V, and the trigger electrodes a' are given a bias of 60 V by the voltage dividers R_{15}-R_{16}. The coupling between the connecting circuits, which are asymmetrical with respect to earth, and the symmetrical subscribers' lines, is effected by means of transformers, such as Tr_2, which also step up the voltage so as to minimize the attenuation caused by the cold-cathode tubes. The microphone current is supplied by a 60 V battery, via the transformer L_2-L_2'. The speech current passes through the capacitor K_2 and is kept out of the battery. The circuit of the microphone is closed only when the calling subscriber lifts the receiver from the cradle. The closing of this circuit causes a voltage pulse in the secondary coil of the transformer. At the cathode side of the tubes of row *2* this pulse has a negative polarity. A rectifier element (V_2), which bridges part of the secondary of Tr_2, limits the amplitude of the pulse to approximately 25 V, which makes the cathodes sufficiently negative with respect to the trigger electrodes a' to permit ignition.

Now assume that all connecting-circuits are free. One of the tubes of row *2*, e.g. tube *2-II*, is thus ignited, so that subscriber *2* is connected up with a connecting-circuit. The current then flowing through the tube causes a voltage drop of 60 V across resistor R_{12}, which raises the cathode potential of the remaining tubes of row *2* to such a value that these tubes can no longer be ignited.

There could, however, still be a remote chance that the pulse will ignite more than one tube. During the short interval of 10^{-4}-10^{-5} sec, in which the auxiliary discharge is established but the main discharge has not yet started, the voltage produced across R_{12} is not yet high enough to block the other two tubes. In order to prevent another tube igniting within the above interval, the three trigger electrodes a' are coupled to each other by capacitors (C_9). If now an auxiliary discharge is initiated in one of the three tubes, then the potential of the relevant trigger electrode will drop, due to the voltage drop across resistor R_{13}, and the capacitors will temporarily transfer this drop to the trigger electrodes of the other two tubes. The latter two are thus blocked — even though the *main* discharge in the first tube has not yet been initiated — thus precluding the possibility that a subscriber would be simultaneously connected with more than one connecting-circuit.

The second condition for blocking, viz. that an engaged connecting-circuit cannot possibly be reached by the other subscribers, is satisfied in the following manner: If, as in our example, subscriber *2* is connected, via tube *2-II*, to connecting-circuit *II*, then a voltage drop of 40 V will occur across resistor R_{14} of this circuit, so that the voltage across $R_{15}+R_{16}$ is also reduced by this amount. This causes the voltage at the tapping on R_{15}-R_{16} to be reduced by nearly the same amount (30 V), owing to the fact that R_{15} is a varistor. The potential at the trigger electrodes a' of the other tubes which might give access to connecting-circuit *II*, is thereby lowered so far that the engaged connecting-circuit can no longer be reached. If, at the moment that subscriber *2* lifted his receiver, one or two connecting-circuits were already engaged, then the relevant tubes, due to insufficient bias of their trigger electrodes a', could not ignite. It is clear that with this electronic circuit, the test for a free connecting-circuit, in fact takes place simultaneously.

We have seen how the subscriber, by lifting his receiver, has been connected to a free connecting-circuit. It now remains to consider how dialling establishes a connection with the called subscriber, how the latter is given ringing current, or how the former receives the busy-signal, and finally how, on replacing the receiver on the cradle, the connection is broken.

Final selector

As we have assumed, tube *2-II* has been ignited by the raising of the receiver *2*. When now a number is dialled, the return of the dial produces square current pulses (*fig. 13a*). This induces in the coils of transformer Tr_2 a voltage proportional to the

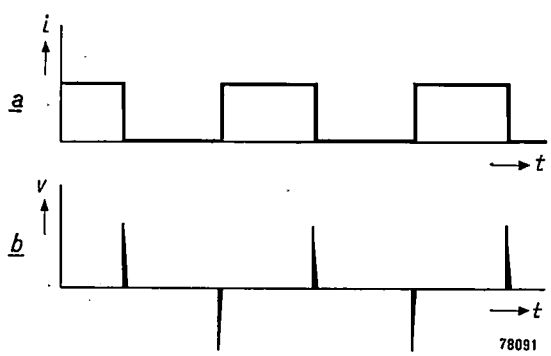


Fig. 13. After dialling a digit, the returning dial periodically breaks and closes a contact in the microphone-current circuit. The square pulses so generated (*a*) induce positive and negative voltage pulses (*b*) in the secondary winding of transformer Tr_2 (fig. 12).

derivative of the current, i.e. consisting of positive and negative pulses (fig. 13*b*). These pulses are also present across the choke coil L_{II} (*fig. 14*, which shows only the calling set *2* and the called set *9*, and only circuit *II* of the connecting-circuits), which prevents the 160 V battery from constituting a short-circuit for the speech currents. The ampli-

tude of the pulses applied to L_{II} depends on the length of the subscribers' line, but counter C_{II}, which has to count the negative dialling pulses, requires positive pulses with an amplitude of approximately 30 V (as explained under *Counting circuit*). For this reason, an amplifier and a pulse-shaping circuit have been incorporated between the choke coil L_{II} and the input of counter C_{II} (fig. 14). The amplifier (A_1) consists of one half of a double-triode ECC 81, which is so biassed that only negative pulses are amplified (*fig. 15*); the amplifier output thus consists of positive pulses only. The pulse-shaping circuit (*PS*) contains a cold-cathode tube Z 500 T (*fig. 16*), which, owing to a trigger electrode bias of +60 V, reacts to the positive pulses from A_1, and supplies positive pulses of constant amplitude to the counter.

After 9 has been dialled, tube *9* of the counter ignites and, as explained in the description of the register circuit, tube *9-II* is thus ignited (by the trigger electrode *a''*), provided, at least, that line *9* is not engaged at that moment.

The burning counter tube has subsequently to be extinguished. This is done by using the sudden and substantial voltage drop occurring at the anode of tube *9-II* during ignition of the latter (voltage drop across R_{14}). A_1 converts this pulse into an amplified pulse of opposite polarity, which, with the aid of a quenching device (*EC*, fig. 14), extinguishes the burning counter tube. The latter process occurs

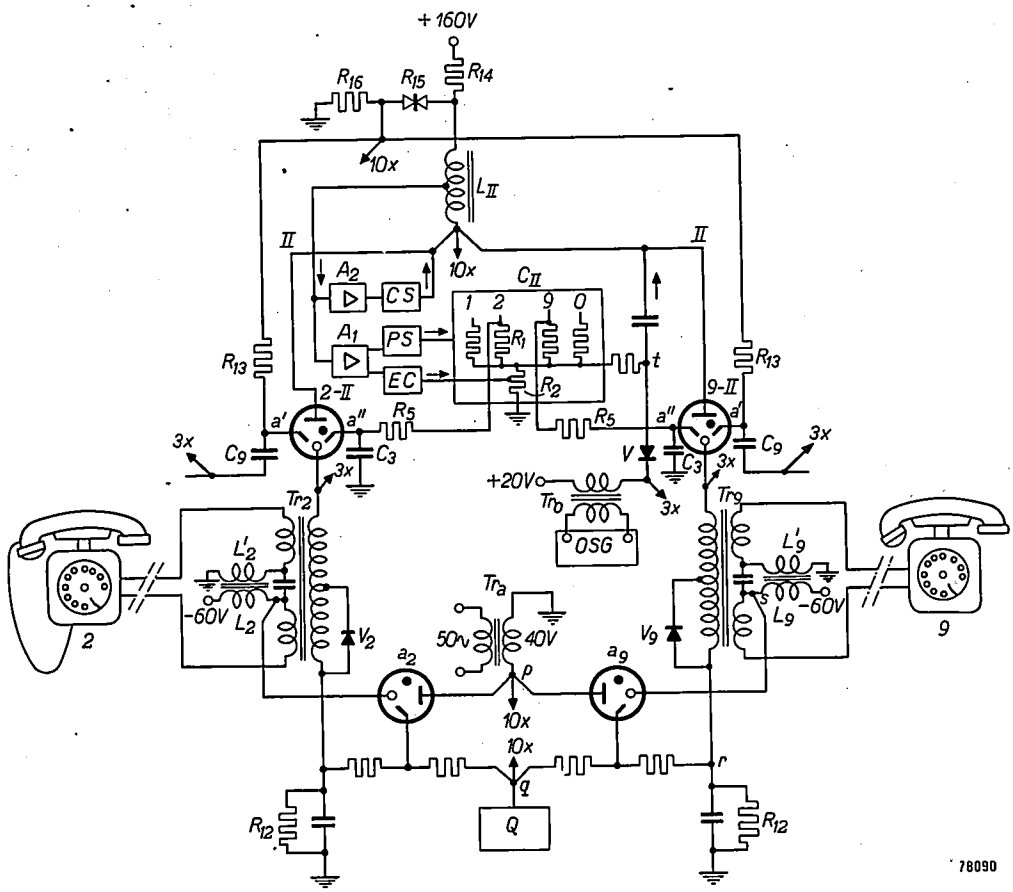


Fig. 14. Diagram showing the signalling and the speech connections between two subscribers (*2* and *9*), in an electronic telephone exchange. Of the three connecting-circuits and the 30 speech current tubes (cf. fig. 11) only the connecting-circuit *II* and the tubes *2-II* and *9-II* are shown (the arrows marked 3× and 10× indicate that in the actual exchange a threefold or tenfold connection is present).

A_1, A_2 amplifiers. *PS* pulse-shaping circuit. C_{II} counting and marking circuit (*C*, fig. 10) of the connecting-circuit *II*. *EC* quenching device of C_{II}. *CS* quenching device for the speech current tubes. Tr_a ringing-current transformer (50 c/s). a_2, a_9 tubes transmitting the ringing current to subscribers *2* and *9* respectively. Q= multivibrator determining the ringing "pattern" (1 sec on, 5 sec off). *OSG* busy-signal generator, with output transformer Tr_0. *V* rectifier element normally blocked, which becomes conducting, however, if an engaged subscriber is called, and passes the busy-signal. R_5-C_3 delay networks as in fig. 10. The remaining letters and symbols have the same meanings as in fig. 12.

exactly analogously to that in the register circuit (tube E in fig. 10). (The cold-cathode tube in EC has a trigger electrode bias of only +35 V and is thus less easily triggered than the tube of the pulse-shaping circuit PS: consequently it reacts to the ignition of tube *9-II*, but not to the dialling pulses.)

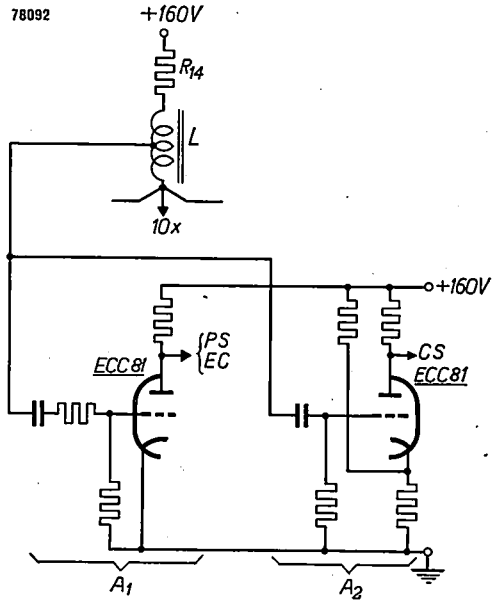


Fig. 15. Diagram of the amplifiers A_1 and A_2 (fig. 14), each using half of a double-triode ECC 81. A_1 is given no grid-bias and consequently reacts to negative pulses only, and only positive pulses occur at the anode, which are applied to the pulse-shaping circuit PS and to the quenching device EC. A_2 is cut-off by a bias of −8 V, which renders all negative and also small (unwanted) positive pulses ineffective, whereas large positive pulses are amplified into negative pulses at the anode, which are applied to CS.

Ringing current

After the tubes *2-II* and *9-II* have been ignited, the ringing current has to be circuited to the called subscriber *9*. The signal generally used for ringing is an alternating voltage with a frequency of 17-25 c/s, generated by an alternator which is often driven independently of the mains. In our experimental exchange, however, for reasons of convenience, a mains transformer (Tr_a, fig. 14) has been employed giving an output of 40 V, 50 c/s. Point p of the secondary winding is connected to the anodes of 10 more cold-cathode tubes (one for each subscriber's line), of which only tubes a_2 and a_9 are shown in fig. 14. In order to ring subscriber *9*, tube a_9 has to be ignited.

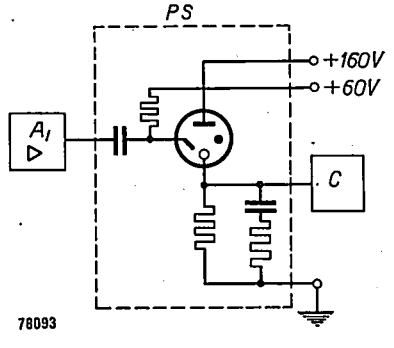


Fig. 16. Diagram of the pulse-shaping circuit (PS, fig. 14). A negative pulse of variable amplitude at the input of the amplifier A_1 (fig. 15) creates a positive pulse at the input of PS, which ignites the cold-cathode tube. The capacitor in parallel with the cathode resistor (in series with a low-value resistor for limiting the charging current) is charged and thus causes the tube to extinguish. A voltage pulse of constant amplitude is produced at the cathode, which is applied to the counting and marking circuit C.

The trigger electrode of this tube is connected to a tapping on a voltage divider connected between points q and r. For the time being, we shall assume the potential of q to be zero, and as long as tube *9-II* is non-conducting, the potential of point r and that of the trigger electrode of a_9 are also zero. Since the cathode of tube a_9, connected to point s, has the potential −60 V, the tube cannot be ignited in this condition.

Once tube *9-II* is ignited, the potential of point r rises from 0 to 45 V owing to the voltage drop across resistor R_{12} between r and earth, and the trigger electrode of a_9 is given a voltage sufficiently high for ignition to take place. This tube then transmits the 50 c/s ringing pulses: across the anode circuit is a direct voltage of 60 V, and an alternating voltage of 40 V (r.m.s.), 50 c/s, so that the voltage across the tube fluctuates between approx. 4 and 116 V. Hence the tube is extinguished each cycle. The telephone set *9* now rings (the coils L_9-L_9' functioning as an autotransformer, stepping the voltage up in the ratio 1 : 2), until the receiver is lifted from the cradle. At that moment microphone current commences to flow through coil L_9; owing to the resulting voltage drop, the potential of point s (and hence of the cathode of a_9) will be raised from — 60 V to — 30 V, so that a_9 cannot re-ignite. For the same reason subscriber *9* cannot receive a ringing current while being engaged in a call with another subscriber.

This likewise explains why tube a_2 was not ignited the moment subscriber *2* lifted his receiver, which, of course, results in a current through tube *2-II* and its associated resistor R_{12}. Although this current causes the trigger electrode potential of tube a_2 to rise, the microphone current simultaneously increases the cathode potential from −60 V to −30 V.

Fig. 17 shows one of the sliding panels of the experimental electronic exchange, containing the components particular to two subscribers, e.g., two subscriber's line transformers, 2 × 3 speech-current tubes and two ringing-current tubes.

The ringing current commonly used is not a continuous, but a periodically interrupted current.

This has been obtained here by making the potential of point q alternately zero and —45 V, viz. zero for one second and —45 V for five seconds.

engaged, and 2) if all connecting-circuits are engaged the moment the calling subscriber lifts his receiver off the cradle.

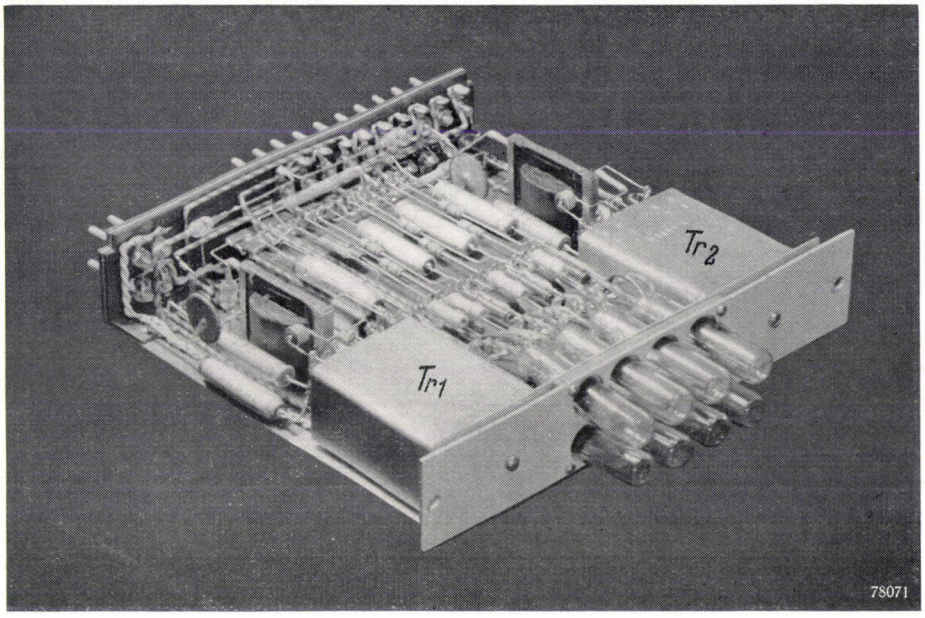


Fig. 17. One of the upper five sliding panels of the exchange in fig. 5. Tr_1 and Tr_2 are two subscriber's line transformers. In between them there are two ringing-current tubes and 2 × 3 speech current tubes.

While a potential of —45 V prevails at the point q, tube a_9 cannot ignite, even if point r has a potential of +45 V. The potential variation of point q is effected by a multivibrator Q. *Fig. 18* shows a multi-vibrator circuit with cold-cathode tubes, an explanation of which is given in the caption.

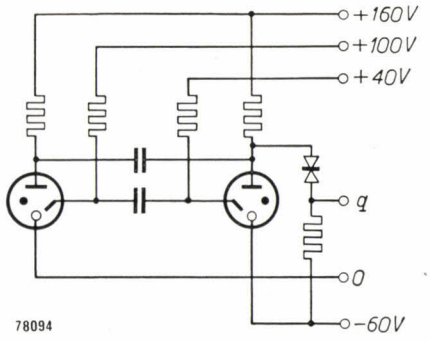


Fig. 18. Diagram of a multivibrator employing cold-cathode tubes (Q in fig. 14). The tubes are alternately conducting; owing to the different resistances in the trigger electrode circuits, the right-hand tube burns for 5 sec and the left-hand tube for 1 sec. The potential of point q (fig. 13) is therefore alternately —45 V for 5 sec, and zero for 1 sec, which determines the ringing "pattern".

Busy signal

In two cases the calling subscriber will hear the busy signal : 1) if the called subscriber's line is

Case 1 : Called subscriber's line engaged. The equipment required in this event is illustrated in fig. 14. *OSG* is the busy signal generator, supplying a periodically interrupted voltage of e.g. 800 c/s. This signal is fed through a transformer Tr_o. Normally, however, the passage of this signal is blocked by a rectifier element (V), the cathode of which, apart from a small alternating voltage, has a permanent potential of +20 V, whereas its anode (point t) is normally at zero potential.

If subscriber *2* dials *9*, then tube *9* will ignite in counter C_{II}. When subscriber *9* is engaged, however, via connecting-circuit *I* or *III*, tube *9-II* does not ignite, since the current flowing through the resistor R_{12} of set *9* gives the cathode of tube *9-II* a potential of +45 V, which prevents ignition. As tube *9-II* does not ignite, the quenching device (via A_1-*EC*, see above) for tube *9* of C_{II} does not come into action. Counter tube *9* thus remains burning (and is not extinguished until subscriber *2* replaces his receiver on the cradle), the voltage at point t increases and rectifier element V conducts, so that the busy signal now passes to the connecting-circuit *II* and is heard in the receiver of subscriber *2*.

Case 2 : All connecting-circuits engaged. For this case we refer to *fig. 19*, which shows again the busy-signal generator *OSG* and the transformer Tr_o. The

latter has a third winding connected to the anode side (g) of a rectifier element V'. If the direction of the steady potential across the element is such that it is conducting, then the third winding supplies an alternating current, via a capacitor C_0, to the secondary of transformer Tr_2, and the subscriber hears the busy-signal. This steady potential is present across the rectifier element only when

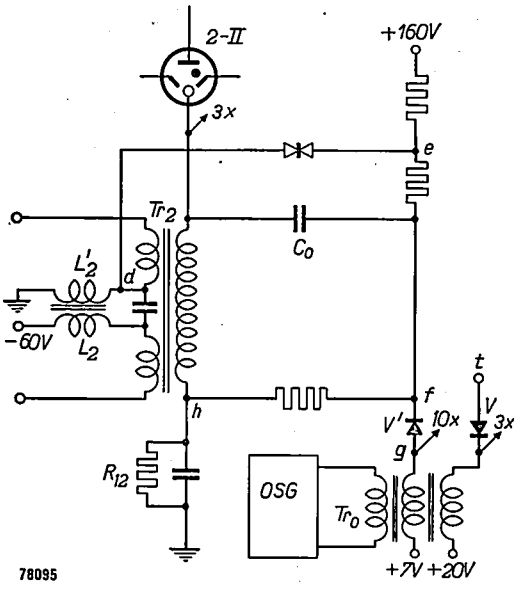


Fig. 19. If all connecting-circuits are engaged, the subscriber, on lifting his receiver, must hear the busy-signal. In that case the rectifier element V', which is normally blocked, becomes conducting, and the busy-signal from generator *OSG* passes through the transformer Tr_o, the rectifier V' and the capacitor C_o, to the transformer Tr_2 and thence to the subscriber's line.

the subscriber has lifted his receiver from the cradle whilst all connecting circuits are engaged.

This is seen as follows. The anode side of the rectifier element, point g, has (apart from a small alternating voltage) a permanent potential of +7 V. The element is therefore not conducting if the potential V_f of the cathode side f is in excess of 7 V, but it does conduct if $V_f \leqq 7$ V. The potential V_f is in fact the mean value between the potentials V_e of point e and V_h of point h. V_e depends on potential V_d of point d, which in turn depends on whether the receiver is on its cradle or not:

receiver on cradle . . . . $V_d = 0$, $V_e = 25$ V;
receiver lifted . . . . . $V_d = -30$ V, $V_e = 0$.

V_h depends on the current through resistor R_{12}. If all connecting circuits are engaged, then tubes *2-I*, *2-II* and *2-III* cannot ignite, hence $V_h = 0$; if, on the other hand, the subscriber does become connected to a connecting-circuit, then V_h will assume the value of +60 V if one tube is burning (e.g. *2-II*), and of +45 V if two tubes are burning (e.g. *2-II* and *9-II*).

There are thus 2 × 3 possible combinations, but in only one of these — the case where the receiver is lifted and all connecting circuits are engaged — can V_f become sufficiently low for the rectifier V' to conduct and thus pass the busy-signal. In all other cases, V_f remains above 7 V.

Fig. 20 shows the circuit of a busy-signal generator incorporating two cold-cathode tubes. The left-hand tube generates a non-sinusoidal voltage at a fundamental frequency of 800 c/s. This oscillation is interrupted for a short moment by the right-hand tube at intervals of 1 sec.

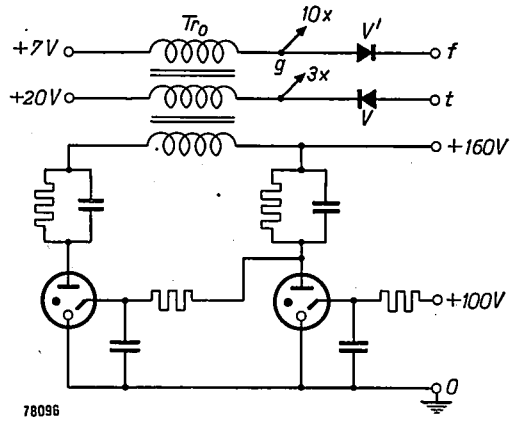


Fig. 20. Diagram of a busy-signal generator with cold-cathode tubes. The left-hand tube operates as a relaxation oscillator with a fundamental frequency of 800 c/s, but is interrupted for a fraction of each second by the right-hand tube which operates similarly at a fundamental frequency of 1 c/s. Tr_o is the output transformer (see figs 14 and 19).

Terminating the call

After completion of the call, the tubes *2-II* and *9-II* have to be extinguished. This is effected by either of the subscribers replacing his receiver on the cradle. By doing this he breaks his microphone current; this results in a positive pulse being applied to the occupied connecting-circuit. This pulse is amplified by an amplifier (A_2 in fig. 14) and its polarity is reversed (A_2 is the other half of the tube ECC 81 of A_1, and is so biassed that only positive pulses are amplified, see fig. 15). The negative pulse at the output of A_2 activates a quenching device (*CS*), which extinguishes tubes *2-II* and *9-II*.

In the event of subscriber *9* having been engaged, the counter tube *9* is, furthermore, extinguished.

Larger electronic telephone exchanges

The electronic exchange (fig. 5) dealt with in this article has been intended for purposes of demonstration and testing in practice. For as many functions as possible, cold-cathode tubes have been employed. Meanwhile, designs have been made for installations on a larger scale, such as exchanges for 1000 and 10 000 subscribers. Only then can the saving in material which is potentially possible by the application of inertia-free controlling devices, be shown to full advantage. Electronic registers and selectors require fewer controlling devices than slower-acting equipment.

It would be pedantic, however, to insist that all or nearly all tubes for these designs should necessarily be of the cold-cathode type. On the contrary, all possibilities should be carefully considered and compared on their merits. This exhaustive task is still proceeding. It is certain, however, that in automatic telephony and allied fields, these new tubes are destined for a wide range of applications.

Summary. This article deals with the application of oxide-coated cold-cathode gas-discharge tubes, fitted with one or two trigger electrodes (types Z 500 T and Z 501 T, respectively), for switching operations in automatic telephone exchanges and related fields. The tube is argon-filled and operates by means of a glow-discharge. With the trigger electrode(s) unbiassed, the breakdown voltage of the tube is higher than 175 V between anode and cathode; when a discharge between cathode and trigger electrode exists, the breakdown voltage lies between 175 V and approx. 70 V (depending on the trigger electrode current). The ignition voltage of the trigger electrode is between 65 and 78 V. At an anode current of 6 mA, the tube presents an impedance of approx. 500 Ω to speech currents. The life expectancy of this type of tube is over 6000 hours (actual burning time) and may be considerably in excess of this. Numerous circuits for this type of tube are reviewed, demonstrating that it can serve as a relay (with all kinds of contact combinations), in counters, marking circuits, registers, pre-selectors, final selectors, signalling circuits, multivibrators, etc. Most of the circuits described are part of an experimental, fully-electronic telephone exchange, which has already been functioning for some years in the Philips Research Laboratory in Eindhoven.

RADIO-CONTROLLED MODELS

by A. H. BRUINSMA. 621.398

It is perhaps a deplorable fact, that toys often take the form of models of military weapons and machines. In the war of 1940-1945 the position was reversed: a "toy", viz. remote control of models by radio, was developed for the remote control of weapons. Since the war the development of the technique of radio control has led to still more advanced models, both for toys and for serious scientific investigations.

The idea of controlling equipment from a distance by means of electromagnetic waves is roughly as old as radio broadcasting itself. Wireless-controlled boats were made some 25 years ago, but development proceeded slowly for various reasons; one of the greatest problems was that of the necessary supply voltages for the uneconomical valves then available. It was therefore usual to employ super-regenerative receivers which required only a small number of valves. The various "instructions" were transmitted by means of a rotary "escapement" switch, which meant that the instructions could be executed only one at a time, resulting in a certain lack of flexibility in control. The technique did not emerge from the amateur stage until its war-time possibilities were exploited during and after the second world war. These advances owed much to the development of high frequency techniques which greatly simplified the problem of the aerial. Radio-controlled aircraft, projectiles, tanks and boats were made, in which this method of control reached new levels of refinement and reliability.

For peace time purposes, too, radio-control is becoming more and more important. Aircraft designers, for example, can study radio-controlled prototype models, which permit enormous economies in the early stages of new designs. Human life, time and money can be saved in the early training of pilots by first giving them practice with radio-controlled pilotless aircraft. Scientific research in the ionosphere by means of controlled rockets is yet another instance. Radio control of toy models — the oldest member of the family — is now also a growing field of application. The possibilities, particularly in the last-mentioned field, are now considerably widened as a result of the introduction of miniature battery valves of low current consumption. In order to investigate the possibilities, two circuits have been designed for radio-controlled models which, for practical reasons, are in the form of model boats.

Design of the models

Where precise manoeuvrability of a model is of importance, one of the first requirements is the execution of the control orders with the minimum delay, or even simultaneously. This, of course, is not possible when an escapement switch is used, but it can be done if the transmitting and receiving equipment operate through several independent channels. This is a well-known and widely used technique in modern telecommunication systems.

The two circuits referred to above are as follows:

a) a simple system based on amplitude modulation, with two separate channels;

b) a system with eight channels employing pulse modulation. Combined with a system of amplitude modulation, this could easily be extended to three times the number of channels; this has not been done, however, as eight channels is ample for the present purpose.

Fig. 1 shows the two radio-controlled model vessels during a demonstration. The vessel on the left is the smaller one with the simple amplitude-modulation system, whilst the other is equipped with the pulse-modulation system. The frequency of both transmitters is 100 Mc/s; this frequency represents the highest at which the valves used give sufficient amplification, whilst at the same time it permits the use of conveniently small aerials.

The transmitters, the current consumption of which is not an important factor, are mains-operated; they radiate about 1/4 watt, this being quite enough for the object in view.

The two model vessels and their respective transmitters will now be briefly described [1]).

The actual control mechanism of the models, that is, the means by which the radio (or other) signals are translated into the various manoeuvres, will be referred to here only incidentally, to clarify the functioning of the radio sections

[1]) For a more detailed description the reader is referred to the booklet Remote Control by Radio, by A. H. Bruinsma, published in the Popular Series of the Philips' Technical Library.

of the equipment. In the last decade, many advances have been made in the field of electronic and electro-mechanical controls, for example, remote-control equipment used for the mass-production of radio-active substances, and electronic controls for machine tools such as lathes and milling machines [2]). In this article, however, we shall confine ourselves to a brief description of the radio side of the equipment used for the two model boats.

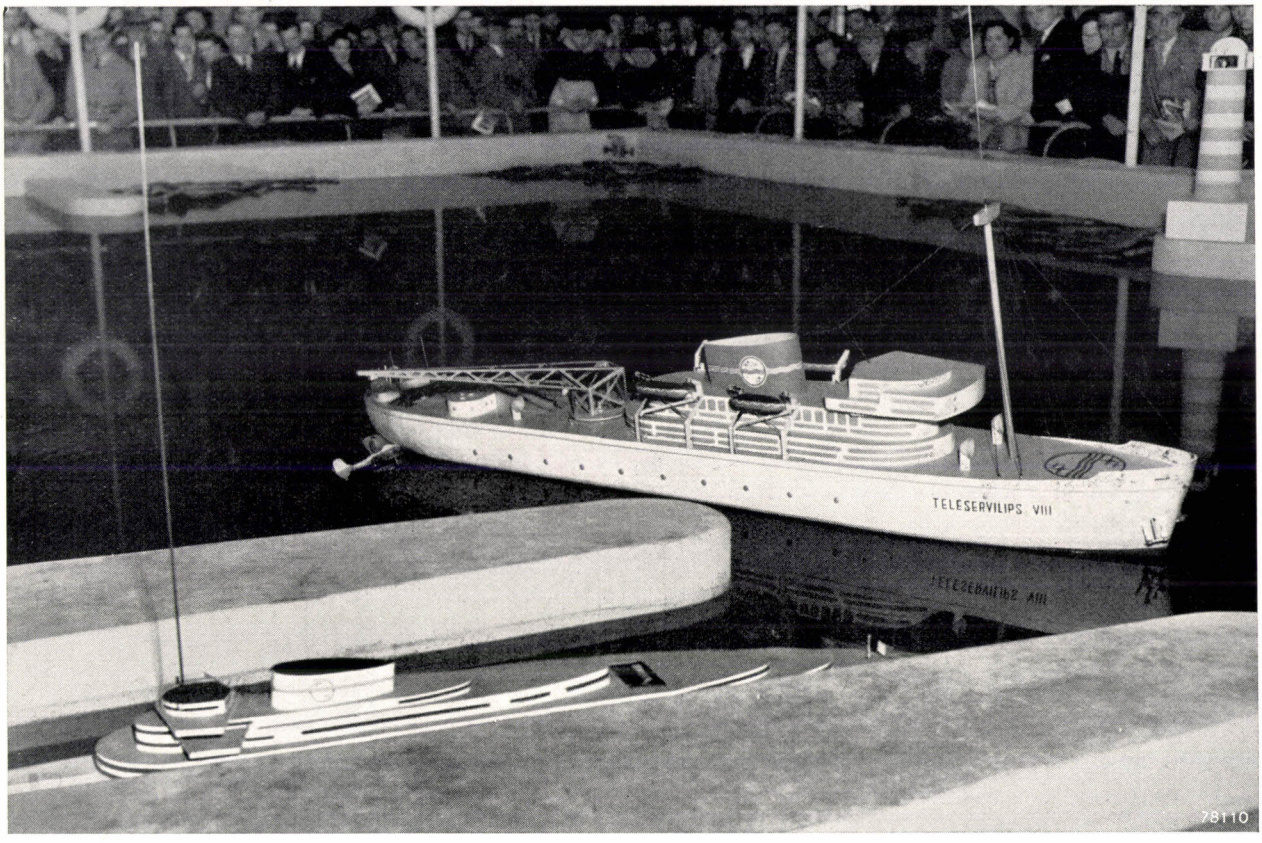


Fig. 1. The two models during a demonstration. Left, in the "harbour", is the boat equipped with the 2-channel system; right, the 8-channel model (the aircraft is just being hoisted on board).

The 2-channel A.M. system

The vessel used for demonstrating the 2-channel system of radio control is equipped with two motors, one to drive the propeller, and another for the steering.

The propulsion motor runs so long as the carrier wave is being transmitted; immediately the transmitter oscillator is cut off, it stops. Provision is made for modulating the carrier with a frequency of 50 c/s, this being effected in the output stage; the presence or absence of this modulation determines the direction of rotation of the propulsion motor, and hence the movement of the vessel — ahead or astern. In *fig. 2*, which is a block diagram of the transmitter, S_1 switches the transmitter on or off, and S_2 switches the 50 c/s modulation in or out.

In addition to the 50 c/s modulation, the carrier is modulated at 2000 c/s by the oscillator O_2, via an amplifier A. This modulation cannot be switched off at will, but is interrupted periodically by a multivibrator M at a frequency of roughly 5 c/s;

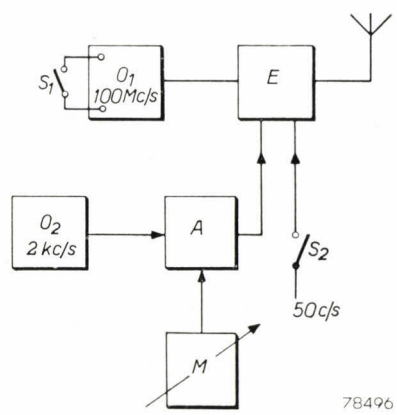


Fig. 2. Block diagram of the transmitter for the 2-channel system, employing amplitude modulation. O_1 and O_2 oscillators; E output stage; A amplifier; M multivibrator. The various "orders" are transmitted with the aid of a switch S_1 (motors on or off), the multivibrator M (steering) and the switch S_2 (direction of rotation of the screw).

[2]) See Electronics **25**, Nov. 1952, p. 172.

the timing of this modulation is variable and controls the rudder movements.

When the modulation pulse length is equal to the intervals between pulses, the steering motor, and therefore also the rudder, are stationary. Adjustment of the multivibrator to deliver modulation voltage for a period longer than the interval between pulses, causes the motor to rotate in one direction, and in the reverse direction when the modulation pulse length is shorter than the intervals between pulses. When the transmitter oscillator is switched off, thus stopping the main motor, the steering motor is also stopped. The position of the helm can be seen by a flag mounted on the stern and coupled to the rudder post.

The superheterodyne receiver in the vessel comprises seven valves (types DK 40, 2 × DL 41, DAF 91, DF 91, 2 × DL 92). In all, the filament consumption is 100 mA at 8 V; the anodes take a maximum of 25 mA at 100 V, and the total power consumed by the equipment on the vessel is therefore 3.3 W.

For the propulsion and steering, small motors of the kind fitted in "Philishave" electric razors are used, the armature and field coil being connected in series or parallel dependent on the direction of rotation [3]). With the series arrangement, these motors will work quite satisfactorily on 60 V D.C. taking a starting current of 75 mA; for the parallel connection 30 V with a starting current of 130 mA is necessary. Dependent on the load, the working current is about half to three quarters of these starting values.

The receiver and also the two motors are fed from 25 pocket-lamp batteries ($4^1/_2$ V), stowed within the vessel; this is sufficient to run the equipment for several hours without replacement of the batteries.

Fig. 3 is a photograph of the vessel with the top section removed to show the receiver, batteries and motors with their associated relays. The transmitter with its three control knobs is also shown (left, foreground). The model is 1.45 m in length and it weighs 12 kg. The receiver, the two motors and relays weigh only 1.7 kg, but the chassis and batteries account for 4 kg; it will therefore be appreciated that very much smaller dimensions would have sufficed if batteries of smaller capacity had been used.

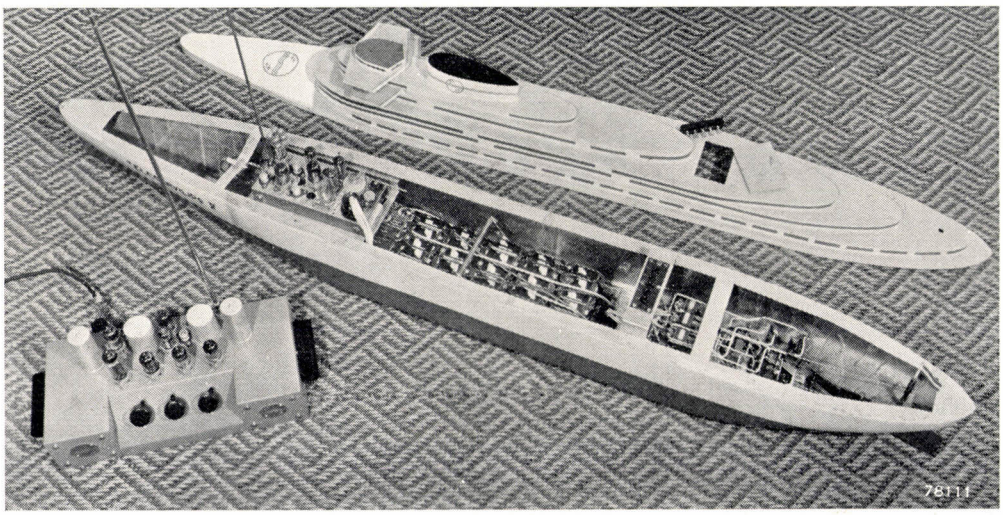


Fig. 3. Demonstration vessel for the 2-channel system with top section removed. The associated transmitter is seen in the foreground.

The 8-channel pulse modulation system

The vessel used for demonstrating the 8-channel system is a twin-screw model; in conjunction with the helm this ensures a high degree of manoeuvrability even at low speeds. Catapult equipment is included for launching a model aircraft which can subsequently be hoisted on board again with a crane. In addition, four lifeboats can be lowered and brought in again, and a loudspeaker on the foredeck provides sound in imitation of a ship's siren, or alternatively, music. The sound can be transmitted from "on shore", i.e. over one of the channels (that particular channel cannot then be included in the radio control proper).

The functions of the other 7 channels are as follows:

2. Control of port propeller

[3]) The "Philishave" motor is actually a series-wound motor.

3. Control of starboard propeller
4. Control of rudder
5. Lowering and hauling up the lifeboats
6. Rotation of the crane
7. Operation of the crane cable
8. Operation of the catapult.

The principle of pulse modulation and the method of transmitting simultaneously independent information by this means, have already been described in this Review [4]). It is sufficient to say here that modulation of the pulse *amplitude* is employed. In the sound channel the amplitude of the pulses is continuously variable. In the other channels, which are merely switching circuits, it is only necessary for the pulse amplitude to vary between two values, viz. zero and a certain maximum, corresponding to "on" and "off". The pulseduration is 11.1 μsec and the repetition frequency is 10 kc/s.

Fig. 4 is a block diagram of the transmitter, in which a — h denote the eight pulse generators, and i the synchronizing pulse generator. Modulation

An A.F. voltage is applied to the channel gating valve I (sound channel) through amplifier A. Control of the motors for the propulsion and steering is effected in the same way as in the amplitude modulation system described above. For this purpose, the channel gating valves II, III and IV are controlled by the multivibrators M_1, M_2 and M_3 which interrupt the train of pulses with a frequency of about 5 c/s. The direction of rotation of the motors is governed by the ratio of the pulse width to the pulse interval in the particular channel, the value of which depends on the setting of the corresponding multivibrator.

The motors controlled by the other four channels need only to be started and stopped [5]), or reversed, this being achieved by switching on and off the channel gating valves V-$VIII$ (switches S_1-S_4). Two channels are needed for the crane, viz. one for rotation and one for operating the cable.

The aircraft is hauled inboard by means of two permanent magnets, one of which is attached to the

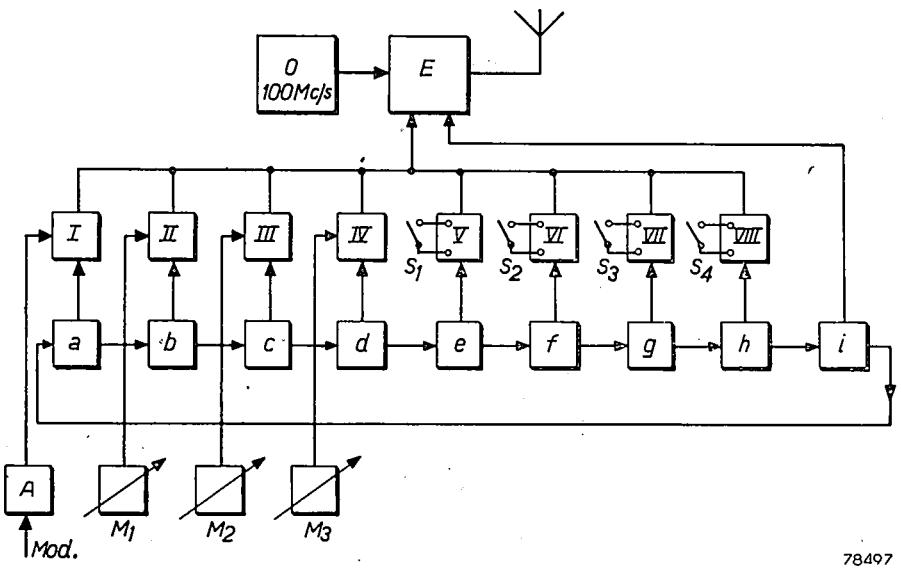


Fig. 4. Block diagram of the transmitter for the 8-channel pulse modulation system. O oscillator; E output stage; a-h channel pulse generators; i sync. pulse generator; I-$VIII$ channel gating valves. The controls are effected by means of multivibrators M_1, M_2 and M_3, and switches S_1-S_4. The first channel (sound) is controlled by a modulation voltage through amplifier A,

by these pulse generators is effected in the output stage of the transmitter through channel gating valves I — $VIII$, and by the sync. pulse generator direct.

crane cable, the other being inside the fuselage of the aircraft; the latter is then guided into its

[4]) C. J. H. A. Staal, An installation for multiplex pulse modulation, Philips tech. Rev. 11, 133-144, 1949.

[5]) Remote reversal of the motors for the lifeboats, the crane and the crane cable is unnecessary: a tumbler switch automatically reverses the direction of rotation when the end position is reached. The motor for the catapult equipment runs only in the one direction.

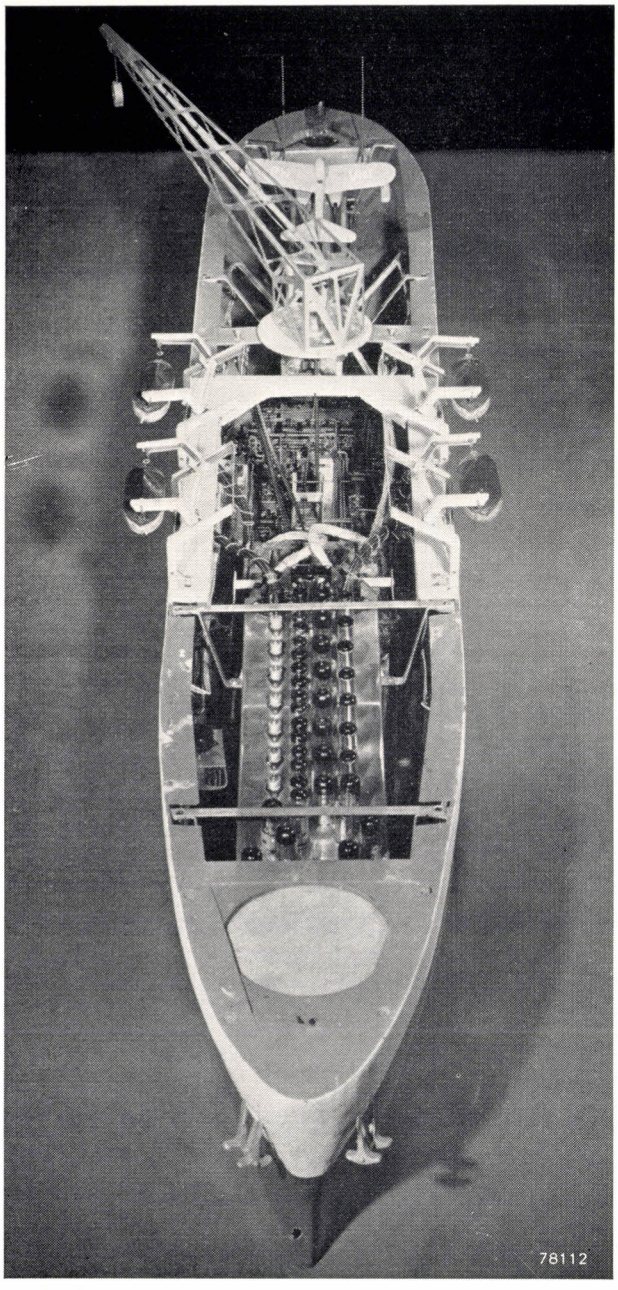


Fig. 5. View of the interior of the 8-channel radio-controlled vessel.

position with the help of two guide pins; the tail is caught by a hook forming part of the catapulting equipment. The machine is launched into the air by a pair of springs engaging with the landing gear.

This vessel is fitted with two types of motor. For the steering and crane, "Philishave" motors are used; so much power is required for the other operations that it was found necessary to employ more powerful motors. The latter employ permanent magnet stators, necessitating current only for the rotors. The "Philishave" motors are run from 8 pocket-lamp batteries, and the others from a nickel-iron battery which also feeds the receiver.

On this vessel a 40-valve receiver is used, the valves being of the types already mentioned. In spite of this large number of valves, only 2.8 W is needed for the filaments. The nickel-iron battery mentioned above delivers 24 V and has a capacity of 15 Ah; the high tension voltage, 164 V, is derived from this battery with the aid of a vibrator, and the total consumption of the receiver, which weighs 1.75 kg, is 27 W.

Fig. 5 shows the interior of the vessel for demonstrating this 8-channel system. The model is 2 m in length and weighs 60 kg. In all, the batteries (the nickel-iron and the dry batteries for the "Philishave" motors) weigh 19 kg. This boat, too, is capable of running for several hours without charging or renewal of the batteries.

Summary. Two radio-controlled model boats are described; one of these, the simpler of the two, is controlled by a two-channel amplitude-modulation system. The other is a more elaborate model, capable of executing a variety of manoeuvres, and controlled by an 8-channel, pulse-modulation system. Although the smaller vessel has 7 valves and the larger one as many as 40, the amount of power required by the receivers as well as the weight of the receivers and batteries, are relatively low, owing to the use of modern low-consumption miniature battery valves.

LATTICE IMPERFECTIONS AND PLASTIC DEFORMATION IN METALS

II. BEHAVIOUR OF LATTICE IMPERFECTIONS DURING DEFORMATION

by H. G. van BUEREN.

548.4:539.374:669

In a previous article the general properties of lattice imperfections such as vacancies, interstitial atoms and dislocations were discussed. With the aid of the theory as developed, it is now possible to obtain an insight into numerous phenomena of practical importance in the plastic deformation of metals, e.g. work hardening, the ageing of alloys, and variations of the electrical resistance.

Slip lines

The most important process involved in the deformation of metals is the slipping of two parts of the metallic crystal over each other. As mentioned in the previous article [1]) both the observed slip over a distance of many times the atomic spacing in a single slip plane, and the low value of the observed critical shear stress τ_{cr} (compared with the theoretical value of $G/2\pi$) can be explained by introducing the idea of special kinds of lattice defects known as dislocations. These are produced in large numbers during the actual process of deformation by so-called Frank-Read sources (the emission of dislocation loops, see I).

In principle, no limit has been set for the number of dislocations that can emanate from a source, although in practice slip distances greater than about 1000 times the atomic spacing on a single slip plane are seldom observed; actually, most slip lines represent a displacement of not more than some tens to some hundreds of times the atomic spacing.

According to various workers in this field [2]) [3]), this limitation in the number of dislocation loops produced should be attributed to the stress set up by the emitted dislocations at the originating source itself; this stress is opposed to the applied stress causing the deformation and, provided it is large enough, may check the activity of the source.

The stress field around a dislocation is fairly complex. Without entering into derivations, we may say that at a distance r from a dislocation axis, stresses occur of the order of magnitude:

$$\tau \approx \frac{Gb}{2\pi r}, \quad \ldots\ldots\ldots \quad (1)$$

where G is the modulus of rigidity of the material and b the Burgers vector of the dislocation. At the centre of a dislocation loop of radius R, as produced by a Frank-Read source, this stress bears the character of a shear stress, opposed in direction to the shear stress that activates the source. The magnitude of this "back stress" set up by the dislocation loop is, in accordance with (1):

$$\tau_b \approx \frac{Gb}{2\pi R}. \quad \ldots\ldots\ldots \quad (2)$$

If there is more than one dislocation loop, the total back stress at the source is practically equal to the sum of the back stresses that each of the dislocations would produce if they acted individually.

The critical shear stress necessary to activate a Frank-Read source is:

$$\tau_0 \approx Gb/l, \quad \ldots\ldots\ldots \quad (3)$$

where l is the length of the source (see I, eq. (3)). As soon as the applied shear stress exceeds this value, the source commences to produce dislocation loops.

Fisher, Hart and Pry [2]) state that once a Frank-Read source has been activated, the minimum shear stress necessary for the source to continue to emit dislocation loops is about $^1/_3$rd of the value indicated by (3). If this were not so and the full shear stress τ_0 were required to maintain the activity of the source, each Frank-Read source would (assuming the applied stress remained constant) become inactive after emitting only one

---

[1]) H. G. van Bueren, Lattice imperfections and plastic deformation in metals, I. Nature and characteristics of lattice imperfections, in particular, dislocations. Philips tech. Rev. **15**, 246-257, 1954 (No. 8/9) (hereinafter referred to as I).

[2]) J. C. Fisher, E. W. Hart and R. H. Pry. Phys. Rev. **87**, 958, 1952.

[3]) N. F. Mott, Phil. Mag. **43**, 1151, 1952 and **44**, 742, 1953.

dislocation loop, since the back stress of that loop would in itself be enough to reduce the shear stress at the source to a value below the critical value (3). In order that the source shall actually become inactive, however, the opposing stress must attain a fairly high value, viz. of the order of τ_0.

In general, a dislocation loop emitted by a Frank-Read source does not in its entirety leave the crystal, but is — at least in part — retained in the region of the source and can thus ensure a permanent back stress. This being the case, we can interpret the limited activity of the Frank-Read source as being the result of the progressive increase in the total back stress when the number of dislocation loops increases. If the number of dislocation loops becomes so great that the total permanent back stress approximates to τ_0, the source will have become inactive. Dislocations can be halted in their progress through the crystal owing to various causes: the principal cause is the presence of other dislocations in the crystal. As mentioned in I, it is probable that a network of dislocations may be present in the undeformed crystal (of which the elements may possibly function as Frank-Read sources). Elastic stress fields will occur around such dislocations, and local stresses may attain such high values that a dislocation loop is but little affected by the applied stress that tends to move it. Secondly, when a dislocation loop intersects elements of a dislocation network, it will contain jogs which, upon further movement, will usually initiate further lattice defects, in the manner described in I. The number of newly created lattice defects increases according to the distance over which the loop progresses, and the energy needed to accomplish this can ultimately no longer be supplied by the applied stress; the dislocation loop then becomes stuck. Thirdly, a moving dislocation can combine with another dislocation moving in a different slip plane, to form a special kind of dislocation which is effectively immobile (*sessile* dislocation).

When parts of the dislocation loops emanating from each Frank-Read source become stuck in the crystal, increased deformation, resulting in the activation of more and more sources, must produce a steady increase in the total dislocation density. This is a significant conclusion of which repeated use will be made below.

Let us now return to the experimental data. The slip bands which can be seen on the surface of deformed metals under the ordinary microscope, in general contain a large number of slip lines, closely packed and extending over large distances — each representing a displacement along its slip plane equal to some hundreds of times the atomic spacing. Recent electron-microscopic studies [4]), however, have revealed that these slip bands are only special instances of a much more uniform distribution of slip lines.

The whole surface of the deformed metal crystal (those examined were aluminium, copper and silver) appears to be fairly uniformly covered with short, fine slip lines (10^{-3} cm to 10^{-2} cm in length), the so-called elementary lines, each of which corresponds to a displacement equal only to some tens of times the atomic spacing (*fig. 1*). Only here and there do the above mentioned, prominent, slip bands occur (*fig. 2*).

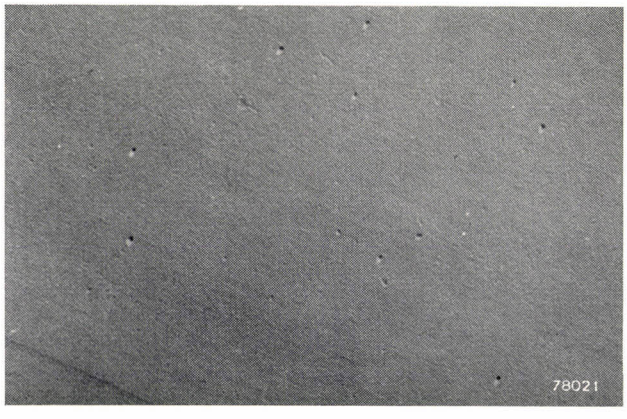


Fig. 1. Elementary lines on the surface of an aluminium crystal which has been stretched by a few per cent. Magnification 20,000x (Figs. 1 and 2 are taken from H. Wilsdorf and D. Kuhlmann-Wilsdorf, Z. angew. Physik **4**, 341, 1952.)

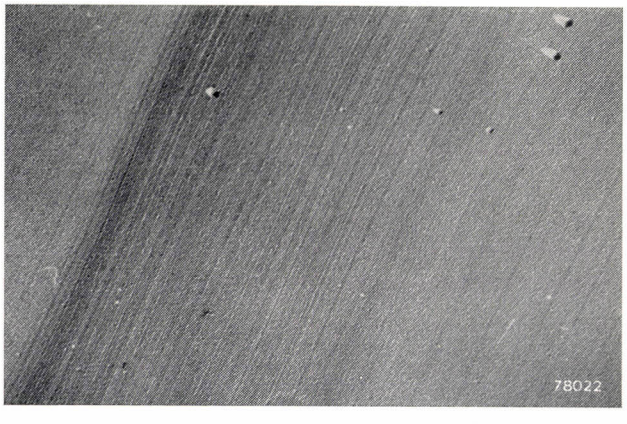


Fig. 2. Surface of aluminium crystal showing slip band gradually merging into elementary structure. Magnification 15000x.

If we regard the formation of elementary lines and slip lines as characteristics inherent in the crystal as a whole, and not merely as a surface effect, we are led to the inference that Frank-Read sources in general give rise only to some few tens

[4]) D. Kuhlmann-Wilsdorf and H. Wilsdorf, Acta Metallurgica **1**, 394, 1953.

of dislocation loops which progress only for a short distance along their slip planes. Where such a "zone" of dislocations intersects the surface of the metal, a "step", i.e. an elementary line, occurs in the surface. The height of the step, or the displacement which the elementary line represents, is then equal to some tens of times the atomic spacing; the length of the line is of the same order as the diameter of the dislocation zone (which will in part still exist within the crystal).

Sources lying some distance below the surface of the crystal will not in general produce slip lines. According to the above ideas, such sources will be surrounded by a number of dislocation loops running wholly within the crystal. Calculations on the behaviour of Frank-Read sources in a crystal support this conclusion [5]). The pronounced, extended slip lines are to be regarded as more or less accidental. Apparently, some favourably placed Frank-Read sources are capable of emitting a much greater number of dislocation loops than are the majority of sources. From the observed fact that the stronger slip lines often occur in groups (slip bands), it is probable that zones of a certain size may exist in the crystal which offer very little obstruction to the movement of dislocation loops or, alternatively, that the formation of an abnormally large number of loops in a single slip plane tends to promote the evolution of equally large numbers of loops in neighbouring slip planes. With regard to the occurrence of slip bands and to the effect of temperature on their structure (which although not pronounced is nevertheless present), no satisfactory explanation has yet been proposed.

According to the foregoing, the observed low critical shear stress of metals is entirely due to the presence of many sources of dislocations, e.g. Frank-Read sources. It is to be expected that, in crystals of such small dimensions that the probability of existence of a source is low, the critical shear stress will be much higher than in larger crystals, and may be expected to approximate to the theoretical value $G/2\pi$. Quite by chance, such small single crystals have become available. It was observed [6]) that under certain atmospheric conditions filaments or "whiskers" appear on the surface of several metals, of a thickness of about 10^{-4} cm. Such whiskers probably contain only one or two dislocations. Experiments have shown [7]) that these whiskers can be subjected to very considerable elastic deformation, without plastic deformation setting in. In the case of tin, a deformation corresponding to a shear stress of the order of several tens of kilograms per mm² has to be applied to these whiskers to produce plastic deformation, whereas larger crystals are plastically deformed at stresses of the order of only 0.1 kg/mm² (The theoretical shear stress $G/2\pi$ of tin crystals is about 200 kg/mm².)

Work hardening

A result of the above-mentioned increase in the number of dislocations in a metal, due to obstruction of the dislocation loops emanating from a Frank-Read source, is that as the deformation progresses, so the stress required to continue deformation increases; i.e. work hardening occurs. The quantitative effect of the premise which relates work hardening to interaction between dislocations, is not easily demonstrated, owing (amongst other things) to the uncertainty that exists as to the manner in which the dislocations are distributed through the crystal, i.e. uncertainty as to the structure of a deformed crystal.

Interaction between dislocations begins to be of importance when the forces between them are, on the average, of the same order of magnitude as the forces resulting from the applied stress in the metal. Calculation has shown that this happens when the average distance between one dislocation and another is about 10^{-4} cm or less; in other words, when the dislocation density, defined as the total length of the dislocation lines within 1 cm³ of the metal, is 10^8 cm^{-2}, or more.

This conclusion can be reached by the following schematic reasoning. Let us imagine two parallel dislocations of Burgers vectors b, separated by a distance d. The stresses occurring at dislocation *2* in consequence of the presence of dislocation *1* are, according to expression (1), of the order of magnitude

$$\tau \approx \frac{Gb}{2\pi d}.$$

The observed critical shear stress of a metal, that is, the theoretical shear stress required to activate a Frank-Read source, is given in (I) as:

$$\tau_{cr} \approx \frac{Gb}{l},$$

where l, the length of the source, is usually found to be of the order of several times 10^{-4} cm. If, during the deformation process, the average distance d between dislocations approaches the value $l/2\pi$, i.e. about 10^{-4} cm, the mean force acting between the dislocations becomes of the same order of magnitude as that resulting from the applied stress. The interaction thus assumes importance when the average spacing of the dislocations is roughly 10^{-4} cm.

If the deformation, and hence also the dislocation density be increased, the average distance between the dislocations is reduced roughly in proportion to the square root of the density; accordingly, the force necessary to accomplish further deformation, i.e. further movement of the dislocations along each other, increases approximately in direct proportion to the root of the dislocation density.

---

[5]) H. G. van Bueren, Acta Metallurgica **1**, 464, 1953.
[6]) K. G. Compton, A. Mendizza and S. M. Arnold. Corrosion, **7** 327, 1951.
[7]) C. Herring and J. V. Galt. Phys. Rev. **85**, 1061, 1952.

Examination of metals by X-ray diffraction has yielded some information regarding the density of dislocations in an undistorted metal crystal: this lies somewhere between 10^7 and 10^8 cm^{-2}. In severely deformed materials, dislocation densities of 10^{10} to 10^{12} cm^{-2} are often found [8]). These figures are only rough estimates, and exact values must await future determinations; however, they have received approximate support, for example, from measurements of the internal damping in metals and by examination of the crystal boundaries.

The dislocation densities mentioned agree well with those which can be derived (along the lines of the above simple calculation) from the observed strain hardening in highly deformed materials, where shear stresses of 10 to 100 times the original value are required to achieve further plastic deformation.

Regarding the relationship between the stress σ and the deformation ε, i.e. the form of the work-hardening curve, entirely satisfactory agreement between theory and observed results has not yet been found. If it is assumed that dislocations are distributed uniformly throughout the deformed crystal — which is an assumption that cannot easily be reconciled with the mechanism of formation of dislocation from Frank-Read sources — an expression of the following form is obtained:

$$\sigma \propto \varepsilon^{1/2}. \quad \ldots\ldots\ldots \quad (4)$$

Careful experiments have shown that in hexagonal metals such as magnesium, zinc and cadmium, which usually exhibit slip only along the basal planes, and also in cubic metals such as copper, gold and aluminium, when only one slip system is active, the strain hardening characteristic is practically linear instead of having the anticipated parabolic form of (4); it is also found that the observed increase in the critical shear stress during deformation is much smaller than would be compatible with the above.

By contrast, when several slip systems are involved simultaneously in the deformation — as almost invariably occurs in cubic metals — a parabolic strain-hardening curve (*fig. 3*) is obtained which agrees reasonably closely with the simple formula (4).

It is probable that the difference between these two cases is that with only one slip system there is little intersection of dislocations, whereas when multiple slip occurs, there is a great deal of intersection, as a result of which new lattice defects are formed. Qualitatively, this would explain the greater strain hardening. The theoretical considerations outlined above would then no longer be valid; it will not be possible to find a way out of this difficulty until more is known about the structure of deformed crystals.

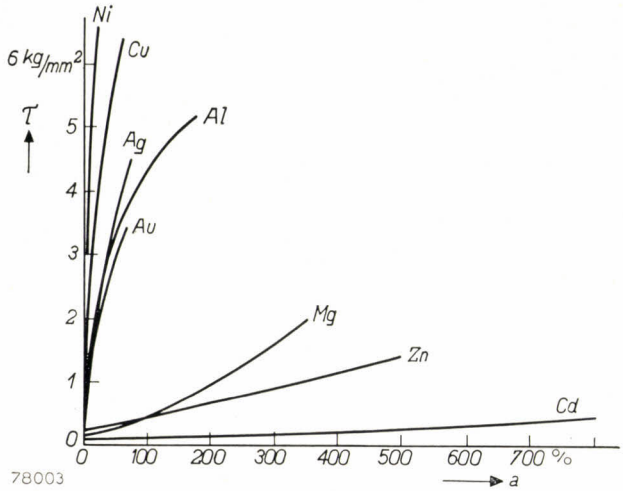


Fig. 3. Work hardening curves of hexagonal and cubic metal single crystals. The curves represent the shear stress τ as a function of the shear a. Cubic metals (Cu, Ag, Ni etc.) exhibit much more strain hardening than the hexagonal (Mg, Zn etc.); moreover, the strain hardening in cubic metals shows an almost parabolic curve, in contrast with an almost linear relationship in hexagonal metals (taken from G. Masing, Lehrbuch der allgemeinen Metallkunde, Springer, Berlin 1950, p. 362.)

Deformation bands, polygonization and recovery

Both microscopic and X-ray diffraction studies have given indications of the presence of obstacles to the movement of dislocations. Apart from slip lines, whether concentrated as slip bands or not, other formations are observed on the surface of deformed metals, which have been given the collective name of deformation bands. An example is shown in *fig. 4*. In the neighbourhood of

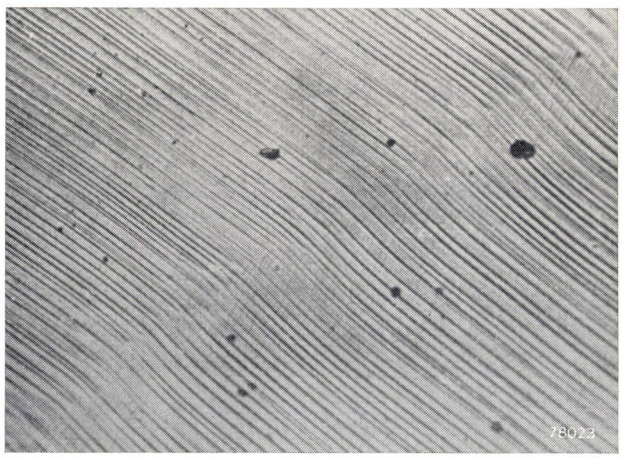


Fig. 4. Deformation band on the surface of an aluminium crystal with 20% extension. (Magnification 250x). Note the curvature of the slip lines. (Figs. 4, 6 and 8 are taken from R. W. Cahn, J. Inst. Metals **79**, 129, 1951.)

[8]) See e.g. A. H. Cottrell, Dislocations and plastic flow in crystals, Oxford University Press, 1953, pp. 100 and 153.

such bands, the pattern of the slip lines intersecting the bands is curved, thus indicating curvature of the lattice planes. Such curvature can be the outcome of the presence of an excess of dislocations of one sign. A single edge dislocation necessarily produces some curvature of the crystal lattice, since the number of lattice planes on either side of the slip plane differs, but this curvature is too slight to be perceptible. It is different, however, when dislocations of the same sign pile up somewhere within the crystal, as for example when the successive loops emanating from a Frank-Read source become stuck. The number of lattice planes on either side of the slip plane then differs considerably, and a perceptible curvature is produced. The present conception of a deformation band is illustrated diagrammatically in *fig. 5*. Activated by a shear stress, a number of Frank-Read sources form dislocation loops of which only the cross section perpendicular to the purely edge portions are shown. All dislocations on one side of a particular source have the same sign, and those originating from the sources shown on the left-hand side of the figure (and which will have moved to the right) differ in sign from the dislocations arising from the sources on the right-hand side. In the zone where the dislocations meet, their movement becomes restricted and the dislocations pile up against one another, as illustrated in fig. 5*b*. This figure also shows the curvature in the slip planes resulting from this congestion; this curvature corresponds to that of the slip planes depicted in fig. 4.

Such congestions of dislocations represent high concentrations of energy. Considerable stresses will occur around the deformation band, in consequence of which nearby dislocations in other slip systems will commence to move. If these dislocations reach the slip planes of the deformation band, they may be blocked by the dislocations already there. It has been shown [9]) that a very effective obstacle to the movement of the original dislocations can thus be formed in cubic metal crystals. According to Mott [3]) this process is responsible for the fact that the dislocations in the deformation band do not return when the applied stress is removed. The congestion can be eliminated only by movement of the dislocations out of their slip planes; such "climbing" movements are non-conservative, however (see I), and will occur with a reasonable frequency only at a sufficiently high temperature (several hundred degrees centigrade). In metals which have been deformed and subsequently annealed, the deformation bands do actually appear to have undergone a change, in the manner shown in *fig. 6*. The slip planes intersecting the band are now no longer curved, but are sharply kinked.

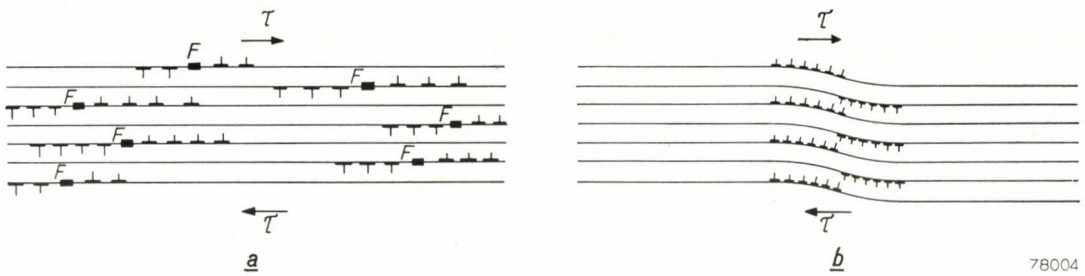



Fig. 5. *a*) When a shear stress τ is applied to a crystal, numerous Frank-Read sources F are activated, which then produce dislocation loops. The symbols $\perp$ and $\top$ in the figure represent the cross-sections of the edge portions of these dislocation loops in the plane of the drawing. The horizontal line shows the orientation of the slip plane, and the vertical line that of the extra atomic half-plane. *b*) With continued loading, dislocations of like sign pile up in a number of slip planes, resulting in curvature of these planes. The configuration depicted corresponds to a deformation band. (Figs. 5*b* and 7 are taken from the book of A.H. Cottrell; see reference [8])).

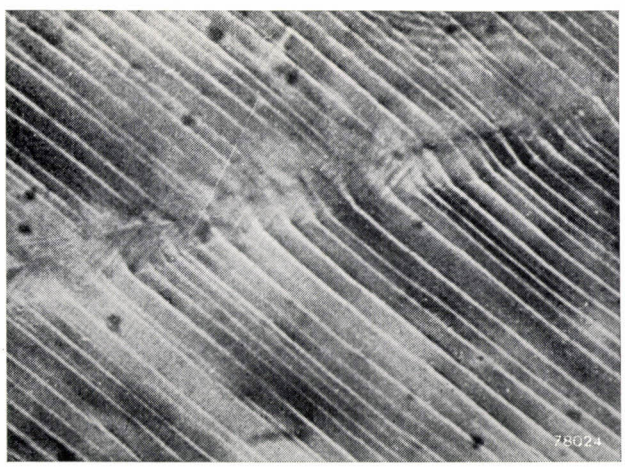


Fig. 6. Polygonized deformation band on the surface of an aluminium crystal elongated 31% and annealed for 1 hour at 450 °C after stretching. Magnification approx. 200x.

[9]) W. M. Lomer, Phil. Mag. **42**, 1327, 1951; A. H. Cottrell, Phil. Mag. **43**, 645, 1952.

This can be explained by assuming that the climbing movement of the dislocations out of the original deformation band produces a configuration as shown in *fig.* 7. This configuration represents much less energy than the original one, because the extra half-plane of each dislocation now merges into the more loosely packed half of the next dislocation.

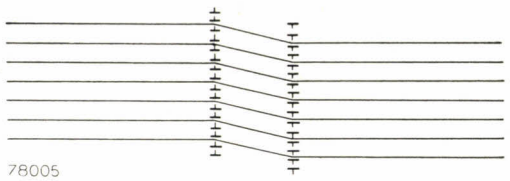


Fig. 7. At sufficiently high temperatures, the dislocations in a deformation band such as that shown in fig. 5*b* may move out of their slip planes and assume positions of lower potential energy. The result is a polygonized deformation band.

This re-grouping of the dislocations into configurations which bear the characteristics of crystal boundaries (the dislocations form divisions which separate two zones of different orientation) has been called *polygonization*. Since the stresses in the lattice are in this case considerably reduced, the mobility of the dislocations, and therefore also the susceptibility of the crystal to deformation, are greater; the work hardening is thus to a large extent eliminated. This is one of the most important causes of the well-known phenomenon of *recovery*.

In most metals recovery takes place during annealing at 400—500 °C; at higher temperatures recrystallization occurs, and here we have to assume that the lattice defects produced by plastic deformation almost entirely disappear, since the properties of a recrystallized metal are nearly enough the same as those of the original, undeformed metal.

Polygonization is also revealed by X-ray diffraction examination which was, in fact, the medium through which this phenomenon was first discovered. The diffraction pattern of an undeformed metal consists of a number of sharply defined points, which become elongated by deformation to form lines, or so-called asterisms, most of which are oriented radially (*fig. 8a*). After annealing at the recovery temperature, the nature of these asterisms is found to have undergone a change, in that the continuous diffraction streaks have broken up into discrete spots (fig. 8*b*). This modification in the diffraction pattern suggests the occurrence of a substructure in the crystal, i.e. numerous crystal elements of slightly differing orientations. This substructure can be identified with that which occurs after polygonization, that is after the formation of a number of polygonization boundaries.

We will now break off our discussion of present-day conceptions of the behaviour of dislocations, and turn our attention to certain other effects in which the lattice defects produced by plastic deformation play a part.

a *b*

Fig. 8. *a*) Transmission Laue photograph of aluminium crystal elongated by 7%. The deformation is revealed by the elongated diffraction streaks (asterisms). *b*) The same crystal after 36 hours annealing at 590 °C. The asterisms are now broken up into discrete spots.

Effects of lattice imperfections in alloys

The effect of lattice imperfections in alloys is threefold: such defects facilitate diffusion of both the foreign atoms and the atoms of the host-metal; they promote the formation of nuclei of a new phase and they can also function as "capture centres" for foreign atoms. As a result of the first two of these effects, phase transitions, recrystalli-

zation, precipitation etc. are all considerably assisted by the presence of lattice defects. Plastic deformation, therefore, which is itself a source of lattice defects, must appreciably accelerate these effects. As an example, it is sufficient to mention the experiments carried out in connection with the increase of the precipitation rate of copper in aluminium, caused by cold working [10]). The precipitation rate of copper in aluminium containing 4% of copper, previously rolled to 10% of its original cross-section, is some 1000 times higher than in the undeformed metal. Microscopic examination shows that the precipitation takes place preferentially in the slip planes, i.e. in those parts of the lattice where the deformation is concentrated.

The ageing of alloys, viz. the gradual variation in hardness that occurs for example after quenching, is also influenced by the presence of lattice defects, As previously explained in this Review [11]), the quench-ageing of carbon steel, for example, is based on the precipitation of the dissolved C-atoms in the phase that is stable at lower temperatures (in this case Fe_3C). In order to accomplish the precipitation, the C atoms must diffuse throughout the lattice. It is therefore evident that the precipitation rate, i.e. the ageing rate, increases steeply when the quenched metal is deformed, e.g. rolled [10]) [12]).

It is interesting to note that the deformation does not appear to affect the precipitation activating energy, but merely multiplies the precipitation rate by a factor that is independent of temperature. It seems that the number of sites to which a dissolved atom can jump is increased, but that the energy necessary for such a jump remains the same. The effect is probably a result of the large number of vacancies created during the deformation process (see I).

The fact that lattice defects can also function as capture centres for foreign atoms is borne out by the effect known as strain ageing. When a metal containing foreign atoms in solution is deformed, it will subsequently exhibit a gradual variation in hardness. The explanation offered by Cottrell and Bilby [13]) has already been discussed in a previous article [14]); it is founded on a trapping of the solute atoms in dislocations which are produced in large numbers during the process of deformation. The best example of strain ageing is found in α-iron containing carbon or nitrogen; the C and N atoms are located at interstitial sites, but they are slightly too large for the available space and therefore show a tendency to diffuse towards the dilated lattice zones in the region of edge dislocations (or dislocations preponderantly of the "edge" type). Here they are trapped, but the dislocation is then blocked, for, in order to move further, it must either take the atom with it or break away from it. Both processes clearly hinder the motion of the dislocation, and the metal as a whole becomes harder. Intimately connected with this phenomenon is the occurrence of the sharp yield point in iron containing C or N. According to the theory, the yield point corresponds to the stress needed to break away the dislocations from their locking points, viz. the trapped C or N atoms.

It is difficult to calculate exactly the ageing rate as a function of the concentration q of the foreign atoms and of the diffusion coefficient D of these atoms in the parent metal. The difficulties are partly due to the peculiar stress field surrounding the dislocation, and partly to the fact that the atoms in solution normally show a tendency not only to diffuse towards the dislocations, but also to form precipitates of the phase which is most stable at low temperatures. In the original theory of Cottrell and Bilby, the last-mentioned effect is disregarded. It has been shown by Fast, however, from experiments based on the internal damping in iron containing either carbon or nitrogen [14]), that the equilibrium concentration of free C or N atoms at ordinary temperatures is so small that these atoms can produce little or no strain ageing. We are thus forced to conclude that the C or N atoms responsible for blocking the dislocations are either present in much higher concentrations than correspond to equilibrium, or that they are produced by the carbide or nitride precipitates present. In the latter event, it can be demonstrated that if the precipitate occurs in the metal in a finely dispersed state, the rate of ageing is determined as a first approximation by the product Dq. This could explain why nitrogen in α-iron results in much more rapid ageing than carbon, at the same temperature.

A phenomenon for which a really satisfactory explanation has not yet been given but which is probably largely determined by the behaviour of dislocations and other lattice defects, is that of creep. This is the slowly increasing plastic deformation of a loaded material that occurs after the instantaneous value indicated by the stress-strain curve has been reached. It is possible to make a distinction between various stages of creep: the most important stage in practice is the steady-state creep which occurs only at elevated temperatures. A special article on the subject of creep is shortly to appear in this Review, however, so a discussion of the mechanism or mechanisms of this effect will be deferred for the present.

In conclusion we shall turn to the question of the variation in the electrical resistance of metals when deformed, particularly at low temperatures.

---

[10]) M. L. V. Gayler, J. Inst. Metals **72**, 243, 1943.
[11]) J. D. Fast, Philips tech. Rev. **13**, 165-171, 1951.
[12]) G. C. Smith, Progess in Metal Physics **1**, 163-234, 1949.
[13]) A. H. Cottrell, Progress in Metal Physics **1**, 77, 1949. B. A. Bilby, Proc. Phys. Soc. A 63, 191, 1950.
[14]) J. D. Fast, Strain ageing in iron and steel, Philips tech. Rev. **14**, 60-67, 1952.

Variation in resistivity due to plastic deformation

Metals possess electrical resistance because the electrons, which as carriers of the current follow a mean path in the direction of the applied electric field, are scattered by lattice defects. A metal with a perfect lattice would have no electrical resistance; hence the resistivity can be regarded as an indication of the divergence from the perfect lattice structure.

At ordinary temperatures, by far the greater part of the resistance is produced by scattering due to thermal vibrations of the lattice; in alloys, scattering is also produced by the alloying atoms. At low temperatures (in the region of the boiling point of hydrogen or lower), the relative effect of lattice vibrations is very slight, and the resistivity of pure metals accordingly drops sharply below that of alloys.

It appears that the plastic deformation of a metal can increase the resistivity to a fairly considerable extent. At ordinary temperatures, the relative increase in resistivity of a heavily worked metal usually amounts to a few per cent; the high melting-point metals such as tungsten and molybdenum are exceptions as their resistivity can rise 10-50 per cent. This increase can be attributed to the large number of lattice defects produced by plastic deformation.

To a first approximation, the various contributions towards the total resistivity arising from one cause or another are additive (Matthiesen's law), so it may be anticipated that the relative variation in resistivity due to plastic deformation at low temperatures, where the effect of lattice vibrations is only slight, will be greater than at ordinary temperatures. Investigation is therefore best carried out at low temperature.

Measurements on metal wires stretched a few percent at the temperature of liquid air [15]) show that the relative increase $\Delta\varrho/\varrho$ of the resistivity at this temperature is of the same order of magnitude as the relative variation in the length $\Delta l/l$ of the wire. The results of some of these measurements are shown graphically in *fig. 9*.

The relationship between the increase in resistivity and the extension is not linear; analysis of the curves shows that it may be approximately represented by:

$$\frac{\Delta\varrho}{\varrho} = a\left(\frac{\Delta l}{l}\right)^n, \quad \ldots\ldots \quad (5)$$

where a is a proportionality constant which is of the order of unity at the temperature of liquid air, and n is generally about 3/2.

Of the various kinds of lattice defects that can be held responsible for the increase in resistivity, the first that come to mind are, of course, dislocations. There are indications, however, that dislocations are responsible for only a small part of the total increase. It follows from various experiments, and is also acceptable from current theoretical points of view, that the dislocation density in a lightly deformed metal (with which we are here concerned) will be at most 10^{10} cm^{-2} (see above). The scattering power of a dislocation can be calculated approximately [16]), and it is then seen that, to explain the increase in resistivity observed in the above-mentioned experiments, a much greater dislocation density is required, viz. from 10^{11} to 10^{12} cm^{-2}.

Dislocations thus appear to be responsible only for a minor part of the observed variation in resistivity. On the other hand, vacancies and interstitial atoms produced by the plastic deformation of a metal can easily cause the observed effects. As mentioned before, such lattice defects occur in large numbers as a consequence of intersection among the dislocations during their progress through the crystal lattice. Jogs are thus produced which leave a trail of vacancies or interstitial atoms in their wake as they move forward. Assuming that the dislocations are almost all generated by Frank-Read sources and that the jogs are formed by the intersection of dislocation loops with the existing network of dislocations, an extension of 10% will yield a concentration of vacancies of about 10^{19} per cm^3.[17])

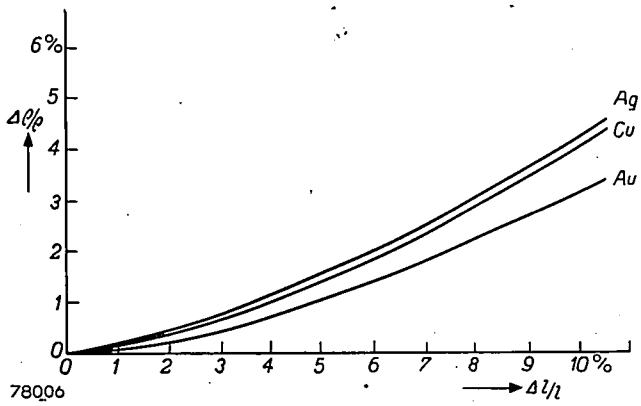


Fig. 9. Relative variation $\Delta\varrho/\varrho$ of the resistivity of silver, copper and gold, due to plastic elongation, plotted against relative variation in length $\Delta l/l$ of the wire, at the temperature of liquid air (unpublished work of J. H. Manintveld, Delft).

[15]) J. Molenaar and W. H. Aarts, Nature, **166**, 690, 1950; M. J. Druyvestein and J. A. Manintveld, Nature **168**, 868, 1951.

[16]) D. L. Dexter. Phys. Rev. **85**, 936, 1952.

[17]) H. G. van Bueren, Acta Metallurgica **1**, 607, 1953.

The scattering power of vacancies and interstitial atoms for electrons can also be computed[18]); the results obtained by various workers vary to some extent, but on average it may be taken that a concentration of such lattice faults of 10^{18} to 10^{19}/cm^3 will be sufficient to explain the observed variations in resistivity.

The deformation ε, and the number of jogs formed, are proportional to the total area covered by the loops emitted by a Frank-Read source, i.e. to the square of the mean radius of the loop. The number of lattice defects left in the trail of the jogs will be proportional to the number of jogs and to the distance progressed by a jog, that is, the mean radius of the loops or, in total, to the 3rd power of this radius. The extent of the scattering of the electrons and hence the increase in resistivity can be said to be proportional to the number of new lattice defects per cm^3; hence the relationship between the increase in resistivity and the deformation can be expressed in the following form:

$$\Delta\varrho \sim \varepsilon^{3/2},$$

in agreement with the observed relationship (5).

If the material, deformed at the temperature of liquid air, is subsequently raised to a higher temperature, some of the extra resistivity is found to disappear. (This refers to the absolute value $\Delta\varrho$ of the extra resistivity; the relative value of the extra resistivity $\Delta\varrho/\varrho$ drops, of course, since the effect of the lattice vibrations (which determine ϱ) in relation to that of the other lattice defects (which determine $\Delta\varrho$) increases with temperature.) A close study of this resistivity recovery[19]) has proved that more than one recovery process is involved, each of which depends on the temperature in accordance with an exponential law in which an activation energy appears. A distinction is made nowadays between three different recovery processes (*fig. 10*). The first of these takes place (certainly, at least, in silver, copper and gold) at temperatures in the region of —100 °C; this is not characterised by a clearly-defined activation energy, but by a range of energies. During this recovery, the extra resistivity drops by 10-20 % or more. The second process occurs (in the metals mentioned) within the range of ordinary temperatures; in this case the value of the activation energy can be much more closely determined. When this recovery has taken place, about half of the extra resistivity is still left. The third recovery process sets in at several hundred degrees centigrade and is intimately connected with the mechanical recovery mentioned in a previous section. By contrast, the two first mentioned processes in the recovery of the resistivity are accompanied — if at all — by hardly any variation in the mechanical properties. This is a further argument in support of the assertion that dislocations are only in part responsible for the extra resistivity.

The recovery processes are undoubtedly closely linked with the diffusion of the vacancies and interstitial atoms to sites where they have less influence, or none at all, on the resistivity. The first of the recovery processes mentioned is at present regarded as the result of the recombination of vacancies and interstitial atoms, and the association and dissociation of groups of these defects. Regarding the quantitative effect of these processes we are still in the dark. It is probable, in view of the relatively broad band of activation energies, that a number of processes take place simultaneously. Very probably too, the second process can be attributed to the diffusion of vacancies to the metal surface, to grain boundaries and to dislocations where they are partly eliminated and partly blocked, whereby their influence on the resistivity is reduced. The activation energy for diffusion of vacancies tallies well with that involved in the second recovery process. The progress of the dislocations is to some extent hindered by the blocking effect resulting from the trapped vacancies. This agrees with recent findings[20]), according to which the second

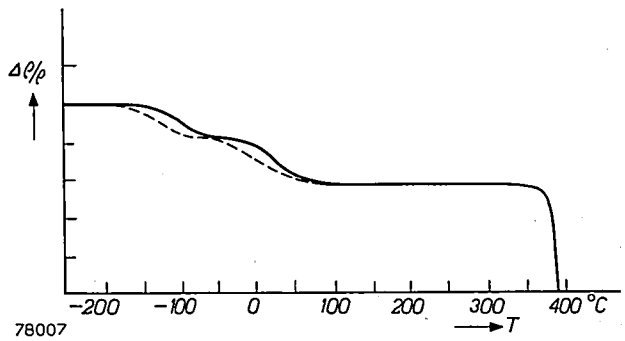


Fig. 10. Recovery of the extra resistivity of a metal produced by deformation at —180 °C after heating at higher temperature. After stretching the (Cu) wire, the temperature is increased to —150 °C for 15 minutes; the wire is then cooled again to liquid air temperature (180 °C) and the resistivity again measured; other parts of the wire are annealed at —130 °C for 15 min, the resistivity being again measured at the temperature of liquid air, and so on. The form of the curve is very dependent on the time during which the wires are maintained at the various temperatures; the dotted part of the curve shows the variation in resistivity for annealing periods of 45 min. From these curves the activation energy of the various recovery processes can be derived. (The curve is due to J. A. Manintveld [19]), except that the part above ordinary temperatures is purely diagrammatic.)

[18]) The most recent computation is that of P. Jongenburger, Phys. Rev. **90**, 710, 1953.

[19]) J. A. Manintveld, Nature, **169**, 623, 1952.

[20]) T. H. Blewitt, Phys. Rev. **91**, 1115, 1953.

recovery process is accompanied by a slight increase in the mechanical hardness. The third process is closely related to the movement of dislocations during mechanical recovery (polygonization), whereby, according to Mott [3]), a large number of vacancies or interstitial atoms disappear, (during the climbing movement of the dislocations), and also the effect of the dislocations on the resistivity is reduced, probably as a consequence of their more favourable arrangement in the lattice.

In this and the previous article, an attempt has been made to present a picture, albeit incomplete, of the formation of lattice defects and their effects in plastic deformation. As this field has been the subject of experimental and theoretical research only a short time, it is not surprising that this review contains so many references to the inadequacy or uncertainty of the theoretical interpretations. Because of this fallability, moreover, the manner in which the views of other workers are presented is unavoidably biassed by the views of the writer. However, there can be no doubt that lattice imperfections constitute a very important factor in the present field of investigation, and that a study of their effects on a wide range of phenomena will ultimately lead to a concise picture of what is now largely a detached collection of fact and hypothesis.

Summary. Following an introductory article, the plastic deformation of metals by the mechanism of slip is discussed with reference to microscopic observations. The following points are dealt with: the various kinds of slip lines visible on the surface of deformed metals, the work hardening resulting from the stress fields round the dislocations, and deformation bands which constitute experimental evidence of the blocking of dislocation movements. Polygonization and recovery are briefly discussed, as well as some phenomena related to lattice defects in alloys, viz. ageing phenomena. In conclusion, some comments are made on the variation in resistivity due to plastic deformation, and its possible theoretical explanation.

ABSTRACTS OF RECENT SCIENTIFIC PUBLICATIONS OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN

Reprints of these papers not marked with an arterisk * can be obtained free of charge upon application to the administration of the Philips Research Laboratory, Eindhoven, Netherlands.

2040: F. A. Kröger: The location of the activator in fluorescent ZnS-Cu (J. chem. Physics **20**, 345-346, 1952, No. 2).

It is argued, — against the opinion of others — that in green-fluorescent ZnS-Cu the activator ion (Cu^+) is located at a normal cation site. Charge compensation is effected either by Cl^- on a normal anion site or by replacing two Zn^{2+} ions by a monovalent Cu^+ ion and a trivalent ion. The blue fluorescence of ZnS-Cu-Cl with high copper concentration is attributed to a monovalent complex, Cu_2^+. At high temperatures these preparations show a green fluorescence. Centres containing more than two Cu atoms act as "killers", influencing the temperature at which the blue fluorescence is suppressed in favour of the green. See Nos. **1931** and **1902**.

2041: B. H. Schultz: Approximative formulae in the theory of thermal regenerators (Appl. sci. Res. **A3**, 165-173, 1952, No. 3).

The application of Hausen's theory to the temperatures and efficiency of regenerators involves very laborious computations. In this article it is shown that, for some limiting cases, more simple formulae can be found with the aid of perturbation theory. The formulae are series expansions for the fluctuating temperatures and the efficiency, in terms of the reduced surface Λ and the reduced heat capacity Γ of the regenerator, for the case of $\Gamma \gg 1$ and $\Lambda \gg 1$, under balanced load conditions.

2042: I. Pelchowitch and J. J. Zaalberg van Zelst: A wide-band electrometer amplifier (Rev. sci. Instr. **23**, 73-75, 1952, No. 2).

It is shown that by appropriate feed-back methods it is possible to construct a wide-band electrometer amplifier with a stable flat response characteristic. Theoretically, the signal-to-noise ratio has been found to be most favorable in the case where the mean frequency of the amplifier is low. Some practical values of the circuit components are given.

2043: J. D. Fast: The allotropic transformation of hafnium and a tentative equilibrium diagram of the system zirconium-hafnium (J. appl. Physics **23**, 350-351, 1952, No. 3).

The temperature of the transformation in hafnium is at least 550 °C higher than the value recently reported by Dewez and is estimated in this paper as 1950 $\pm$ 100 °C. A tentative equilibrium diagram of the system zirconium-hafnium is given. From

this it would appear that in the neighbourhood of 100% Hf the transition temperature is very sensitive to small admixtures of Zr.

2044: J. M. Stevels: Expériences et théories sur le facteur de puissance des verres en fonction de leur composition, III (Verres et Réfractaires **6**, 3-7, 1952, No. 1). (Experiments and theories on the power factor of glasses; in French).

French translation of **R 189**.

R 198: M. E. Wise: Dense random packing of unequal spheres (Philips Res. Rep. **7**, 321-343, 1952, No. 5).

A heap of rigid spheres has a given typical statistical distribution for their radii, and the spheres are densely packed like the grains in a powder. "Dense random packing" is defined in a new way, in terms of a probability distribution function W for tetrahedra: each tetrahedron has its vertices at the centres of four spheres of which every sphere has to touch the other three. Thus W is defined by four radii. Boundary effects are ignored. General equations for W are deduced and also one of the possible numerical solutions for one log-normal distribution of radii; this is used to obtain approximations for several of the most interesting geometrical properties of the heap. In this way, for example, the mean density and the mean number of interstices are found, and also a statistical distribution for the largest possible spheres that can go into the interstices, and one for the number of spheres touching one sphere of given radius. Finally, the system is discussed critically as a model for a real packed heap. The mathematical methods are also of interest: these include some little-known formulae of solid geometry and formulae for integrating numerically in two, three and four dimensions.

R 199: J. Bakker: Kinetics of the emulsion polymerization of styrene (Thesis, Utrecht, Philips Res. Rep. **7**, 344-400, 1952, No. 5).

The rate of polymerization of styrene in an emulsion has been measured by a dilatometric method, while at the same time the end of the incubation period has been determined by measuring the change in electrical conductivity of the emulsion. The curves representing the conversion as a function of time proved to be S-shaped. The measurements were accurate to within 5%, the error being due mainly to the difficulty in determining the moment at which the reaction begins. Up to 50% conversion, this conversion-time function could be represented by $x^{\frac{1}{2}}(1-x)^{-\frac{1}{2}} = t/t_{\frac{1}{2}}$, independently of the amounts of soap or catalyst used, when the emulsifier is potassium myristate (x = degree of conversion, t = time, $t_{\frac{1}{2}}$ = half-conversion time). When other soaps are used — laurate, stearate or oleate — small deviations from this formula were found. Beyond 50% conversion, the function could be represented by

$$\frac{1-\varphi}{k}\,(x-\tfrac{1}{2}) - \frac{\varphi}{k}\ln 2(1-x) + 1 = t_{\frac{1}{2}}$$

where k probably depends on the nature of the soap, but not upon the amounts of soap and catalyst; φ is characterstic of the monomer. With K-myristate as emulsifier, k was found to be 1.10 ± 0.07. This parameter is related to the number of latex particles, thus making it possible to evaluate this number. Given equal quantities of monomer and equal volumes of water, the half-time value $t_{\frac{1}{2}}$ bears a relation to the amounts of soap (S) and catalyst (R), viz $t_{\frac{1}{2}} = \text{const.}\,(SR)^{-\frac{1}{2}}$. The moment at which the incubation period ends (t^*) appeared to be practically independent of the amounts of soap, but dependent upon the amount of catalyst as $t^* = \text{const.}\,R^{-2/5}$ or, after correction of the experimental values of t^* on theoretical considerations, as $t^* = \text{const.}\,R^{-3/5}$. This applies only as long as the amounts of soap used are less than 15% by by weight of the amount of monomer. When this limit is exceeded the micelles no longer disappear entirely from the solution. From considerations of the initiation theory of Smith and Ewart, another method was found for determining the number of latex particles, which, combined with the first method, provided a means of calculating the area ω_0 occupied by one molecule of soap. For K-myristate we found $\omega_0 = 35 \pm 6$ Å^2. Investigations concerning the initial stages of the reaction revealed that $t_{\frac{1}{2}}$ should depend not only upon specific properties of the soap but also upon the initial weight of monomer m_0 and upon the volume of the aqueous phase V. Experiments showed, however, that $t_{\frac{1}{2}}$ is independent of V but directly proportional to $m_0^{\frac{1}{2}}$, within the range covered by the experiments. Further experiments are required before any conclusion can be drawn as to the dependence of $t_{\frac{1}{2}}$ outside the present range. Some measurements of the degree of polymerization of the polymers obtained led to the conclusion that this is determined entirely by chain transfer, which is reconcilable with the concept of the termination reaction in the case of emulsion polymerization. In none of the experiments was any modifier used.

VOL. 15 No. 11, pp. 297-324 MAY 1954

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
THE PHILIPS INDUSTRIES

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, NETHERLANDS

A LINE CONVERTER FOR THE INTERNATIONAL EXCHANGE OF TELEVISION PROGRAMMES

by J. HAANTJES and Th. G. SCHUT.

621.397.2

*The international relay of the London television broadcasts on the occasion of the Coronation was a spectacular proof of the possibility of converting television pictures of one scanning frequency into another. On a smaller scale, this had already been demonstrated a year previously, when broadcasts from Paris were relayed to England *). This article gives a description of the "line converter", which in June 1953 enabled the Dutch and West-German TV transmitters to relay programmes from London and Paris by effecting the necessary conversions from pictures of 405 and 819 lines respectively into 625 lines.*

At present, several different television standards are in regular use in Western Europe. They all have the same frame frequency, viz. 25 complete pictures per second (two interlaced frames forming one complete picture), but one of the points in which they differ is the number of lines per frame. An exchange of TV programmes among the West-European countries is therefore possible only if there are means available for converting pictures of one number of lines into those of another. This problem became urgent when, in France, the Netherlands and Western-Germany, the wish was expressed for relay-broadcasts of the B.B.C. transmissions scheduled for the Coronation. As is well known, such a relay network was indeed established [1]) and in general very satisfactory pictures were received. The 405-line signal was relayed, via a chain of link transmitters, to Paris, where the Radiodiffusion-Télévision Française took care of its conversion into an 819-line signal [2]). A second chain of link transmitters took the 405-line signal from Cassel (Northern France) to Breda (Holland), where it was converted into a 625-line signal. Two link transmitters set up in Breda relayed this converted signal to the Dutch TV stations at Lopik and Eindhoven. Two further link transmitters beamed the signal from Eindhoven to an existing network of similar transmitters in Germany.

In this article we shall first consider some general aspects of the problem of line conversion, and then give a more detailed account of the system used in Breda.

General aspects of the problem

Indirect method

One of the ways of showing TV pictures to a large audience, e.g. in a cinema, is the following. A film is made of the image on a picture tube; the film is developed, fixed, rinsed and dried at great speed, and immediately afterwards projected for the audience. It has been possible to reduce the time required for processing the film to 1-2 minutes.

Clearly, instead of being projected, the film can be run off through a film scanner, which is the usual procedure if films are to be broadcast by television. By virtue of the intermediate image on the film, the outgoing signal is no longer dependent on the

*) TV from Paris, Wireless World 48, 298-300, Aug. 1952.
[1]) Philips tech. Rev. 14, 358-360, 1953 (No. 12).
[2]) The link Paris-Lille in the map accompanying the article referred to in [1]) brings the 819-line signal from Paris to the transmitter at Lille. The transmitter in Berlin (not shown on the map) was also included in the West-German network.

number of lines of the original picture. This is called the *intermediate film process* — an *indirect* method, as opposed to the *direct* methods to be dealt with later.

This process has certain drawbacks. First of all, the photographing of TV pictures (50 frames per second) with a film speed of 25 frames per second is no easy matter; the obvious solution of moving the film during the frame blanking of the TV picture is not feasible with ordinary film cameras, as the time required for the film to move one frame is too long. In the second place, a highly elaborate installation is necessary for the rapid processing of the exposed film. Added to this are the costs of the film material. A further complication, also involving delay, is the fact that the sound has to be recorded as well. All this tends to make the intermediate-film method rather unattractive.

Direct methods

There are also methods which are practically delay-free and can, therefore, be considered as direct methods.

It is possible, for example, to reproduce the original TV picture in the form of a potential pattern on a flat plate (target plate), and then scan this pattern with an electron beam in accordance with the required number of lines. The formation of the potential pattern can be realized in two ways:

1) directly, by means of an electron beam modulated with the original signal, or

2) with the aid of an intermediate optical image, displayed on a picture tube, and viewed by a television camera (in this case the target of the pick-up tube of the camera is the plate on which the potential pattern is formed).

In the former case a special tube is required, containing a target plate, a "writing" and a scanning electron beam. (Alternatively, two tubes may be used — each with a target and one beam; whilst projection is taking place in one tube, the other is scanning; after one complete picture their functions are reversed, etc.) With this first method it has been found rather difficult to retain the original gradation of the picture.

It has been suggested that the second method has the drawback that, apart from the "electrical intermediate image" (the potential pattern), an optical intermediate picture is also required, and that each increase of the number of intermediate pictures will inevitably impair the image quality. This apparently reasonable argument, however, should not carry undue weight, for the following reasons. The intermediate optical picture can be directly watched and checked; also a large experience has been accumulated in the presentation of TV pictures on cathode-ray tubes and the prevention of possible faults — which cannot yet be said of the first-mentioned method.

It has further been suggested that in the method using an intermediate optical picture, it is not desirable to use a camera tube in which the secondary electrons can return to the target [3]); the resulting redistribution effect would cause the formation of spurious signals or "edge flare" to a disturbing degree. Camera tubes in which this redistribution effect occurs, are the iconoscope and the image iconoscope [4]). We know, however, that the image iconoscope — to mention only this

[3]) A. Cazalas, Problème de la transformation des standards de télévision, Onde électr. **31**, 178-183, 1951; V. K. Zworykin and E. G. Ramberg, Standards conversion of TV signals, Electronics **25**, 86-91, January 1952.

[4]) For the working principle of these tubes, see P. Schagen, H. Bruining and J. C. Francken, Philips tech. Rev. **13**, 119-133, 1951.

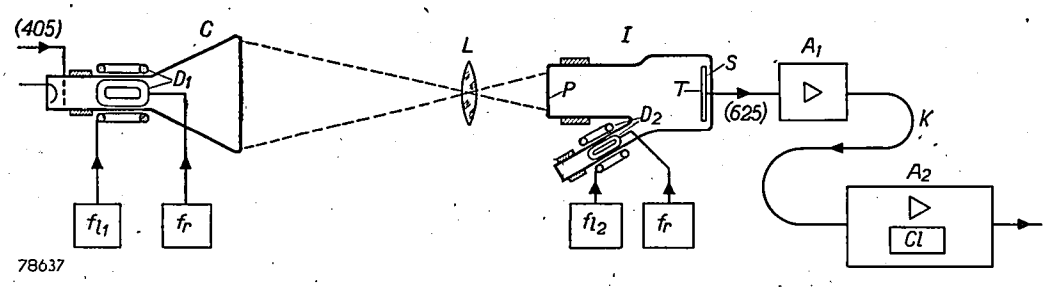


Fig. 1. Line converter with intermediate optical picture for converting a 405-line television picture into one of 625 lines. The 405-line picture is made visible in the conventional way on the screen of a (special) picture tube C (with deflection coils D_1; line frequency $f_{l_1} = 405 \times 25 = 10\,125$ c/s, frame frequency $f_r = 50$ c/s). A lens L projects the image of C on the photo-cathode P of a camera tube (image iconoscope I). Its target (T) is scanned in accordance with the 625-line system (deflection coils D_2; line frequency $f_{l_2} = 625 \times 25 = 15\,625$ c/s, frame frequency $f_r = 50$ c/s). S signal plate. A_1 pre-amplifier, K cable. A_2 main amplifier (with clamping circuit Cl, see further in this article).

tube — does not give rise to spurious signals if only the photo-cathode is sufficiently illuminated, although then the redistribution effect is rather strong. Without undue trouble, the light spot of the picture tube can be given a luminance sufficient for the image iconoscope to operate under favourable conditions.

This consideration, added to the fact that the image iconoscope is capable of producing images of excellent definition and gradation, led to the decision to use an image converter based on the system of an intermediate optical picture and an image iconoscope camera (*fig. 1*).

Before giving a further description of the equipment, we will deal with some disturbing effects which may occur, and the manner in which these have been counteracted.

Disturbing effects and their prevention

The following three disturbing effects are encountered with a line converter employing an intermediate optical image:

1) Superimposed on the photo-current of the camera tube is a signal corresponding to the original number of lines; without counter-measures, this signal would occur as a disturbing component in the output signal.

2) Interference will occur between the two numbers of lines.

3) Although the same picture-repetition rate is used over the whole of Western Europe (25 complete pictures per second), slight differences do occur, which give rise to beat effects.

These three effects will now be dealt with successively. Unless stated otherwise, we shall only consider conversion from 405 to 625 picture lines.

Presence of the original signal in the photo-current

The image on the picture tube of the line converter is not a "continuous" image, but a picture which is traced line by line by a moving light-spot. A corresponding light-spot therefore moves across the photo-cathode of the image-iconoscope. For the sake of convenience, we shall first assume that the picture to be transmitted is a surface of uniform brightness. The light-spot then has a constant brightness and thus gives rise to a constant photo-current, irrespective of position. At the end of each line, however, and also upon the completion of a frame, the light-spot disappears for a moment (due to the line blanking and frame blanking, respectively), and consequently the photo-current momentarily drops to zero.

The photo-current therefore contains Fourier components with the line frequency ($405 \times 25 = 10125$ c/s), with the frame frequency (50 c/s) and with multiples of the two. If an actual image is transmitted, then the light-spot varies in brightness and the photo-current will also contain Fourier components corresponding to the content of the image. Thus the character of the 405-line picture is still present to a considerable degree in the photo-current.

The photo-electrons impinge on the target, where they cause potential fluctuations, which are capacitively transferred to the signal plate. Hence the output signal of the image iconoscope consists of a mixture of 625-line and 405-line signals. To contend with this situation, two remedies have been resorted to, one being directed against the high frequency components, and the other against the low frequency components of the disturbing signal. Used in conjunction, these were found to deal effectively with the disturbance caused by the fluctuating photo-current.

High frequencies: Fluctuations of the photo-current can be reduced by using a long persistence phosphor for the picture tube of the line converter. The afterglow of the moving light-spot then smooths out the variations of the total photo-current, in particular those caused by the line and frame blanking signals.

The duration of the afterglow is limited by the fact that excessive afterglow would cause an objectionable blurring of movement, visible as blurs or smudges behind moving parts of the picture. It was found possible, however, to lengthen the afterglow to such an extent that at least the high-frequency components were smoothed out of the photo-current without creating objectionable blurring.

Willemite was chosen as the phosphor; it has an afterglow of approximately 13 milliseconds. After $^1/_{50}$ second, 20-25% of the original luminance remains, and after $^1/_{25}$ second still 5-6%. Willemite has a good fluorescent efficiency and the spectrum of the fluorescent light is favourable with respect to the spectral sensitivity of the photo-cathode. The colour of the light is green.

Low frequencies: The photo-current, and hence the output signal, still contain the low-frequency Fourier components of the original signal. These can be removed if the required signal contains a certain periodically recurring reference level, having a repetition frequency which is high compared to the disturbing frequencies. Due to the low-frequency disturbing components, this reference

level gradually varies, but by means of a suitable circuit, it can be restored to a fixed level. This is done in the manner illustrated in *fig. 2*.

In our particular case, the black level is used as the reference level. Since the image iconoscope does not transmit any "direct voltage", this level is not automatically present, but has to be specially introduced. An opaque strip is placed at one side of the picture on the photo-cathode. The part of the target corresponding to this strip lies at the side where the line-scanning begins and is scanned by the beam just before the beginning of each line. In this manner the black level is established, at the repetition frequency of the scanning lines.

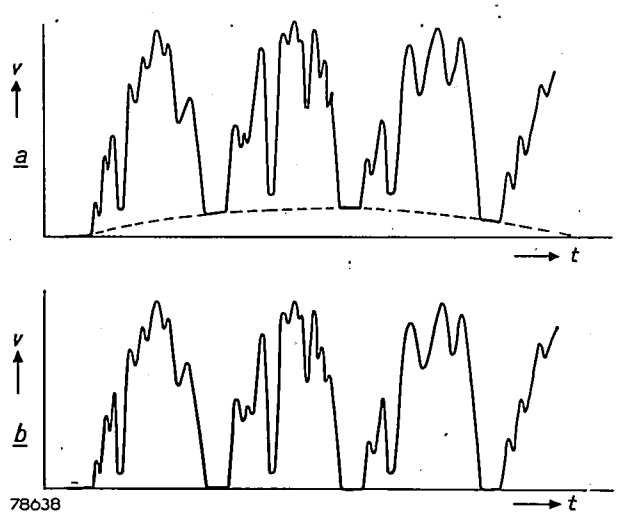


Fig. 2. Output signal v (at an arbitrary point in the amplifier A_2, fig. 1), plotted as a function of the time t.
a) Without counter-measures, the black level would fluctuate slowly as a result of the presence of low-frequency components in the photo-current of the image iconoscope.
b) By means of "clamping", the black level is fixed at the beginning of each line.

In addition, at the beginning of each line, the control grid of one of the tubes of amplifier A_2 (fig. 1) is earthed for a short moment ("clamping"), by means of an electronic switch.

The clamping circuit (*fig. 3*) is made up of four diodes, which are so arranged that they are conducting only if point A has a positive and point B has a negative potential. In that condition, the control grid g_1 is brought to earth potential. At the beginning of each line, a multivibrator (M_2) applies a positive pulse to A and a negative pulse to B, and thus each time reduces the signal level at g_1 to zero. The multivibrator M_2 is controlled by another multivibrator (M_1), which in turn is triggered by the line-synchronizing pulses of the 625-line system. With the aid of these multivibrators the most favourable position and width of the pulses applied to A and B can be adjusted.

By thus fixing the black level for each line, not only is the disturbance due to the low-frequency component in the photo-current largely eliminated, but also low-frequency disturbances of different origin, such as hum (from camera or amplifiers), and spurious signals in the vertical direction.

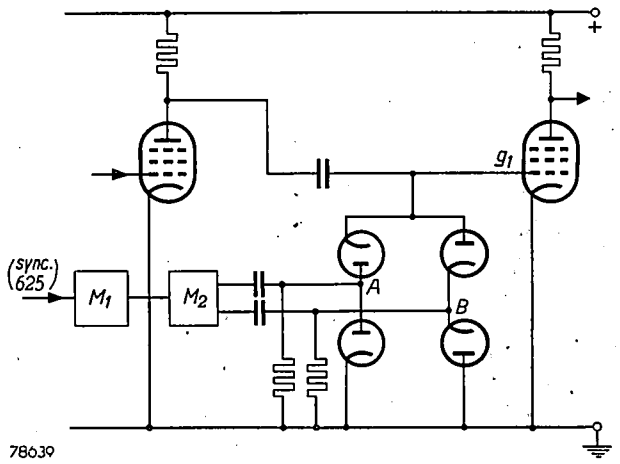


Fig. 3. The clamping circuit used, incorporated between two pentodes of the amplifier A_2 (fig. 1). M_1, M_2 multivibrators. g_1 control grid, which is periodically restored to earth potential via the four diodes.

Interference between the two numbers of lines

The second difficulty encountered is caused by the fact that the traced lines (405) and the scanned lines (625) have a tendency to form interference patterns on the target of the camera tube. This is to be attributed to the co-existence of various factors. Essentially it is caused by the changes in the mutual positions of the two line-types: at one place a scanned line will more or less coincide with a traced line, at another place it will come approximately in between two traced lines. In the latter case the signal obtained will be the weaker of the two. As a consequence of this, an interference pattern consisting of horizontal dark and light bands will be superimposed on the 625-line pattern.

At first sight one would expect the number of dark bands to equal the difference between the numbers of lines, i.e. $625 - 405 = 220$ (approx. 10% of which will be invisible because of the periodic frame blanking).

Such an interference pattern, however, is hardly discernible as far as the image iconoscope is concerned, but another pattern, containing about 100 dark bands, does occur. The cause of this lies in the fact that the "scanning sensitivity" of a target element is greatest just before and smallest immediately after scanning; we define the "scanning

sensitivity" here as the reciprocal of the ratio of the charge supplied by the photo-current in a short time interval to the resultant charge contributed to the signal current.

For the output signal we have

$$i_{sign} = [\int_{t_1 - 1/25}^{t_1} s(t) \cdot i_{ph}(t) \cdot dt]/1/25$$

in which i_{sign} is the signal current at the moment of scanning (t_1), and $s(t)$ and i_{ph} are the scanning sensitivity and the photo-current respectively, at a moment t.

The occurrence of an interference pattern with approximately 100 dark bands is illustrated by *fig. 4*, in which the measured relative sensitivity is plotted as a function of the time. A description of the method of measurement is given in the caption. This curve clearly shows that whatever is traced during the *last* $1/50$ *second before scanning*, contributes far more to the output signal than that traced $1/50$ second previously.

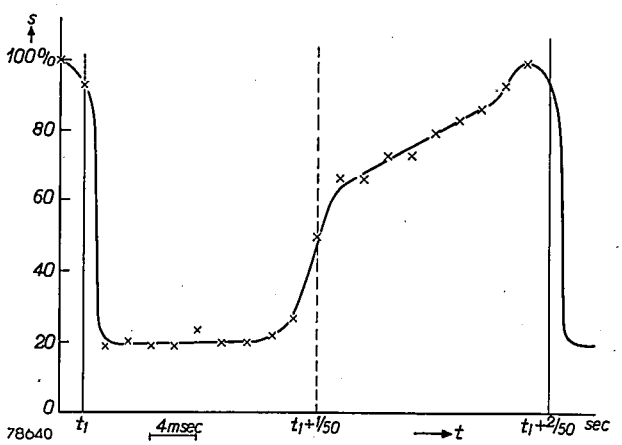


Fig. 4. Sensitivity s of a target element of the image iconoscope, plotted as a function of time t. After each scanning of an element (t_1, $t_1 + 2/50$, ...), s drops from its maximum value (represented here as 100%) to its minimum value. At $t_1 + 1/50$, neighbouring elements are scanned by the interlaced frame.

The curve was measured as follows. Light flashes of duration 4 milliseconds were projected on a small area of the photocathode with a repetition frequency of 25 c/s. The time lag between the illumination and the scanning of the corresponding area of the target was varied and the output signal of the image iconoscope measured as a function of this time lag, at constant values of photo-current and scanning current. For different values of these currents, slightly different curves were found, but all having a similar form, viz. a sudden drop of s on scanning, and an increase beginning midway between two successive scannings.

If the tube is used as a line converter, therefore, each scanning produces a signal which is mainly derived from the *immediately preceding* frame tracing: The lines of one scanned frame will thus mainly interfere with the lines of the immediately preceding traced frame. This causes an interference pattern with $\frac{1}{2}(625—405) = 110$ dark bands (about 100 of which are visible). This "secondary" pattern, as has been mentioned already, predominates over the primary interference.

A satisfactory method of removing this interference pattern was evolved, which consists in tracing the lines on the picture tube of such a thickness that they overlap. Although this involves a certain loss of definition in the vertical direction, the definition in the horizontal direction remains unimpaired if the thickening of the lines is effected by "spot wobbling".

This is a small vertical oscillation superimposed on the normal rectilinear movement of the light-spot, so that the latter traces sinusoidal lines. The amplitude is about twice the line separation, and the frequency is chosen so high that the oscillation is not visible.

In our line converter, spot wobbling is effected with the aid of a simple valve oscillator using a pentode, and an additional pair of deflection coils on the neck of the picture tube. These coils are incorporated in the oscillatory circuit of the oscillator. The frequency is approx. 10 Mc/s, so that one line includes roughly 600 sine-wave cycles. The amplitude is variable by adjusting the screen-grid voltage of the pentode.

In June 1953, after the London broadcasts, some programmes from Paris were also relayed to the Netherlands. It was then found that the above-mentioned interference occurred to a far less degree with a conversion of a larger number of lines into a smaller number (819 into 625). Even without spot wobbling, no disturbing interference was discernible.

Further examination of the origin of the "secondary" interference pattern

Fig. 5 shows four different "cross-sections" of the lines on the target plate, marked *1*, *2*, *3* and *4*. *AB*, *CD* and *EF* are cross-sections of lines of the first *traced* frame, *A'B'*, *C'D'* and *E'F'* of lines of the second (interlaced) frame; *PQ*, *RS*, etc. are cross-sections of lines of the first *scanned* frame, and *P'Q'*, *R'S'* etc. of lines of the second (interlaced) scanned frame.

For simplicity, the following assumption are made: the two traced frames are exactly complementary (thus $A'B' = BC$, etc.), and the two scanned frames are also complementary ($P'Q' = QR$, etc.); further, the upper *traced* line and the upper *scanned* line start at the same level (both A and P are at the level h_1), and the number of traced lines to the number of scanned lines is in the ratio of 3 : 4. From this last assumption it follows that after three traced lines (i.e. four scanned lines) the initial condition is restored, so that the situation between h_1 and h_2 is representative of the whole target. Furthermore, we shall, for the time being, assume that the afterglow of the picture tube is zero.

Fig. 5 shows on which scanned lines a signal of the immediately preceding traced frames is obtained, as well as the ratio of the corresponding signal strengths. This depends on the extent to which the odd-numbered scanned lines (under *2*) overlap the odd traced lines (under *1*), and the overlap of the even-numbered scanned lines (under *4*) on the even traced lines (under *3*). The extent of the overlap in each case is shown by the shaded parts in fig. 5; the size of each shaded area is a measure of the signal strength at the corresponding scanned line. We see that at the levels h_1 and h_2 a strong signal is produced and between these, at a level h_m, a weak signal.

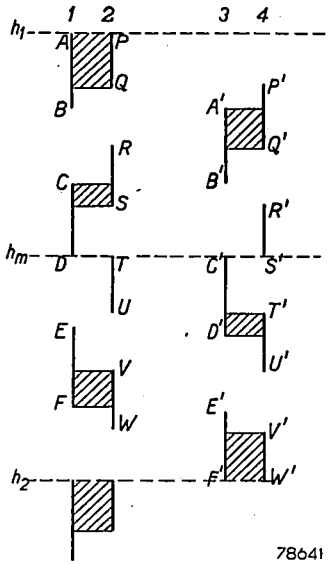


Fig. 5. An illustration of the occurrence of the "secondary" interference pattern. AB, CD and EF are "cross-sections" of lines *traced* on the target of an odd-numbered frame, and PQ, RS, TU and VW are cross-sections of odd-numbered *scanned* lines. The cross-sections $A'B'$, $C'D'$, etc. and $P'Q'$, $R'S'$, etc. similarly correspond to even-numbered frames. In this example it is assumed that P lies at the same level as A (at the level h_1), and that the ratio of the number of traced lines to the number of scanned lines is 3 : 4 (thus below h_2 the band h_1-h_2 is repeated), and the interval between the lines is equal to the line width ($AB = BC = A'B'$, etc. and $PQ = QR = P'Q'$, etc.). The size of each shaded area is approximately a measure of the signal strength per scanned line. Strong signals are thus produced at h_1 and h_2, and weak signals between them, at h_m. This explains the occurrence of light bands in the picture at h_1 and h_2 and dark bands at positions corresponding to h_m.

When an image is built up from these signals, it will exhibit light bands at heights corresponding to h_1 and h_2, with dark bands in between. This applies to both interlaced frames. The number of dark bands is equal to the difference between the number of lines per frame; this is what we have defined above as the "secondary" interference pattern. (The "primary" pattern, with double the number of dark bands, viz. the difference between the number of lines per *complete picture*, is lacking in fig. 5. This is a consequence of the assumption that the tracing lines are of the same width as the interspacing, although actually they are slightly narrower. This pattern shows light bands at h_1 and h_2 as in the secondary interference, but in addition there is also a light band at h_m. The latter is formed because the line $R'S'$ is scanning the traced line CD. This produces only a weak signal, since CD does not belong to the immediately preceding traced frame; for this reason primary interference is not very pronounced with the image iconoscope.)

The *intensity* of the secondary interference pattern depends on the time interval between tracing and scanning, i.e. on the time lag between the tracing beam in the picture tube and the scanning beam in the camera tube. It also depends on the afterglow or decay time of the phosphor. This can be clarified if we examine more closely what actually happens on the surface elements of the target.

Consider first the surface elements which receive their charges from, say, an "odd numbered" tracing frame. These surface elements can belong to either of two groups: elements that will be scanned in the next scanned frame, and elements for which this will take place in the subsequent (interlaced) scanned frame. If the time lag between tracing and scanning for the first group is Δt, then it is $\Delta t + {}^1/_{50}$ sec for the second group. The elements of the two groups are by no means uniformly distributed over the target plate; there are places where elements of the first group predominate, as well as places with a majority of elements of the second group, and furthermore there are places where both groups are represented at about equal strength. The places with mainly elements of one group correspond to the bright bands of the primary interference pattern, since there the scanned lines coincide more or less with the traced lines, so that the time lag has in the main the same value (Δt in the first, third, etc. bright band; $\Delta t + {}^1/_{50}$ sec in the second, fourth, etc. bright band).

The same applies to an "even-numbered" tracing frame. For this, too, there are certain areas of the target where elements with time-lag Δt prevail and which, therefore, correspond to the first, third, etc. bright bands in the picture; and areas where the predominant time-lag is $\Delta t + {}^1/_{50}$ sec, corresponding to the intermediate bright bands, as well as the mixed areas corresponding to the dark bands of the primary interference pattern.

Consider now the photo-current i_{ph} which impinges on an element ΔS, having a time-lag Δt. In *fig. 6*, i_{ph} is plotted as a function of the time. At an instant t_1, the tracing beam in the picture tube passes a surface element of the fluorescent screen, corresponding to ΔS. This element begins to emit light, at an intensity which decreases according to the afterglow characteristic of the particular phosphor. The photo-current decreases correspondingly (fig. 6*a*). The target element under consideration is scanned at the instant $t_2 = t_1 + \Delta t$. In fig. 6*b* the same is shown for an element having a time-lag $\Delta t + {}^1/_{50}$ sec; at $t = t_3$ the photo-electrons begin to impinge upon the target element, and at $t_4 = t_3 + \Delta t + {}^1/_{50}$ sec the scanning beam passes the element.

The thin line in fig. 6 represents the scanning sensitivity of a target element (cf. fig. 4). The signal can be found from this diagram by integration of the product of s and i_{ph} over a period of ${}^1/_{25}$ sec, but without going into this further, it is apparent that the signal produced in the case of fig. 6*a* is stronger than that of fig. 6*b*, since in fig. 6*a* the state of high sensitivity coincides with a high momentary value of the photo-current, while the contrary is the case in fig. 6*b*.

This difference in signal strength, which causes the secondary interference pattern, is obviously dependent on Δt, the value of which may vary between 0 and ${}^1/_{50}$ sec. If one reduces Δt, the signal strength in fig. 6*a* decreases and that in fig. 6*b* increases, and at a certain value of Δt both signals will be equally strong and the secondary interference pattern will disappear. Secondary interference could be avoided in this way; in practice, however, the earlier mentioned method of

spot wobbling is more satisfactory, since then Δt can be made as large as possible, so that the effect of movement blurring is reduced to a minimum.

To complete this account, we should mention still another cause of the occurrence of the secondary interference pattern, viz. the slight overlapping of the lines of the interlaced frames. As, however, this effect is of minor importance in an image iconoscope, we shall not pursue it further here.

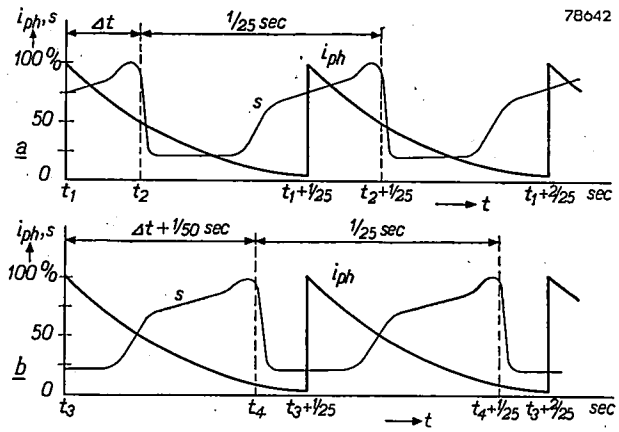


Fig. 6. Effect of the time lag Δt between tracing and scanning. The photo-current i_{ph} impinging upon a target element, and the sensitivity s of this element are plotted as functions of time (i_{ph} follows the afterglow characteristic of the phosphor, and s is the characteristic shown in fig. 4; both quantities are represented on an arbitrary scale).
In (*a*), a high sensitivity coincides with a high momentary value of i_{ph}, which is not the case in (*b*). This is the reason why a stronger signal is produced in case (*a*) than in case (*b*).

Beats caused by inequality of the frame frequencies

The third difficulty mentioned occurs if there is a slight difference between the picture frequencies of the two TV systems. As a rule, the mains frequency — nominally 50 c/s in Europe — is selected for the frame frequency (twice the picture frequency). If, however, the grid systems of the countries in question are not interconnected — and the British system is not linked up with the continental ones — small differences in the picture frequency will occur. The time lag between tracing and scanning of the target in the camera tube will then change continuously. This causes the output signal strength of the camera to fluctuate with a frequency equal to the difference between the two picture frequencies. In the final 625-line picture this is visible as a brightness fluctuation having the difference frequency.

If the picture frequencies are unequal, it is moreover possible that some of the aforementioned disturbing effects will be noticeable. If, due to inadequate clamping, the low-frequency components of the photo-current have not been completely eliminated, a dark horizontal band is seen across the screen, and also the amount of movement blurring and the intensity of the secondary interference pattern fluctuate.

By using suitable values of the mean photo-current and the beam current of the image iconoscope, these difficulties can be largely eliminated. More effective, however, is *the exact synchronizing of the two picture frequencies*, i.e. during the relay, the 625-line system is made to operate, independently of the continental electricity supply, with the British picture frequency, which is determined by the frame-synchronizing pulses in the British signal.

During the relays of June 1953, synchronized frame frequencies were used. At certain moments, however, when the chain of link transmissions between London and Breda was disturbed, the line converter temporarily received no frame-synchronizing pulses, or only badly distorted pulses, with the result that on the 625-line receivers the picture became unsteady, which is very unpleasant for the viewers. In such cases the exact synchronization of the picture frequencies was dispensed with and the difficulties reduced in the manner described above, by selecting the most suitable adjustment for the mean photo-current and the beam current.

Description of the equipment used

The line converter is shown in *fig. 7*. It was installed in duplicate in a trailer which was situated, for the relay of June 1953, at the foot of a church tower in Breda (see figs. 2 and 3 of the article referred to in [1])). On top of the tower were a beam receiver, directed to Antwerp, and two transmitters beamed at Lopik and Eindhoven.

The equipment consists of two parts: the 405 (or 819)-line section, which comprises the picture tube and its auxiliary components, and the 625-line section, viz. the camera tube and its associated equipment. Each is housed in three portable units (*1*, *2*, *5* and *3*, *4*, *6*, respectively, fig. 7).

The picture tube and its auxiliary equipment

The picture tube is of a special construction. It has a flat window of only 12 cm diameter and produces a picture of very high definition: in the centre the resolving power is more than 900 lines. The tube was operated at an anode voltage of 25 kV and a mean beam current of approx. 40 μA. The phosphor employed (willemite) then produces a luminance [5]) of approx. 5000 nit (≈ 16 000 apostilb) in the brightest parts of the image; the average luminance is then ample to

[5]) 1 nit = π apostilb = 1 candela (international candle) per m² = 1 lumen per m² per unit solid angle.

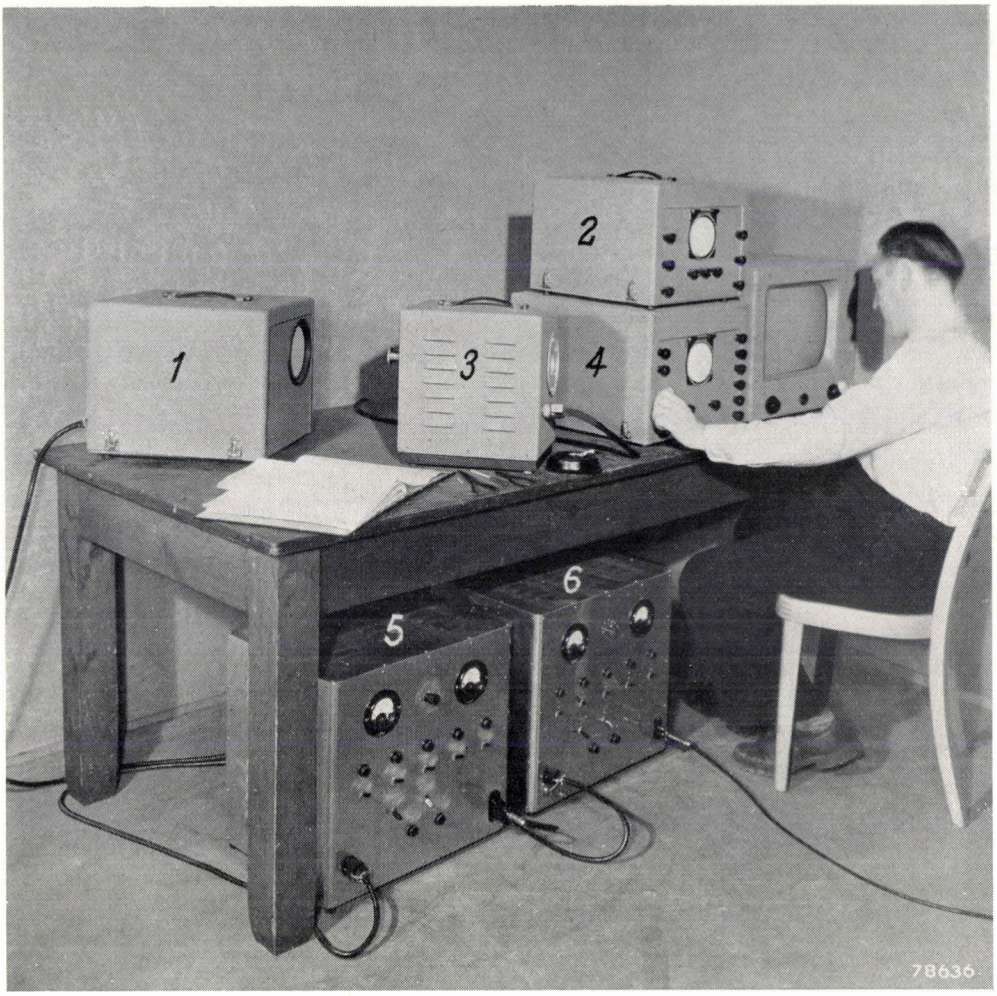


Fig. 7. The six portable units, constituting the line converter used at Breda. Unit *1* contains the picture tube, the video-output amplifier and the oscillator for spot-wobbling. Unit *2* contains the first video amplifier, the saw-tooth generators for the picture tube and the controls. Unit *3* contains the camera and its image iconoscope, type 5854, two output valves for the deflection currents and a pre-amplifier for the picture signal. Unit *4* houses the second picture-signal amplifier, the saw-tooth generators and the controls. *5* and *6* are power-supply units. To the extreme right is a monitor.

permit the image iconoscope to operate without disturbing spurious signals.

The saw-tooth generator for the horizontal deflection can be set by means of a switch to 405, 819 or 625 lines (the last line-number makes it possible to test the equipment without a British or French signal being available).

It is essential that the synchronization of both deflection systems of the picture tube is very stable. For this reason the *horizontal* deflection circuit incorporates a flywheel circuit, which can be adjusted to any of the three TV standards. This flywheel synchronization considerably reduces the influence of pulse-shaped disturbances and of temporary irregularities in the line-synchronizing pulses [6]).

This method, however, does not give satisfactory results if the incoming line-synchronizing pulses are slightly frequency-modulated (e.g. a modulation with 50 c/s from the mains supply to the synchronizing pulse generator). If this should occur, and flywheel synchronization is used, then the sides of the frame, and indeed all vertical lines of the picture, will be distorted. For this reason provision is made for cutting out the flywheel synchronization. During the transmissions of 1953 it was found that the synchronization pulses were nearly always good enough to permit flywheel synchronization.

In a method often used for synchronizing the *vertical* deflection, the synchronizing signals are applied to an integrating *RC* network. The voltage across the capacitor of this network undergoes a substantial change only during the frame-syn-

[6]) P. A. Neeteson, Philips tech. Rev. **13**, 312-322, 1952.

chronizing signal, and the passing of a certain critical value determines the moment when a frame change takes place. A simple method of this kind will not be adequate for a line converter, which has to give reliable operation at any of three TV standards each with a frame-synchronizing signal of different shape. Satisfactory results were achieved, however, by using a circuit with *double integration*.

It has been described how, by means of clamping, the black level of the 625-line signal was fixed. Something similar takes place with the incoming signal, by the aid of a clamping circuit which operates on the control grid of the output tube of the video amplifier preceding the picture tube. The back porch of the line-synchronizing signal is then used as a reference level. As, however, the incoming signal could not always be counted on having a well-defined back porch, it was necessary that this clamping too, could be cut out at will.

The camera tube and its auxiliary equipment

The camera (*3* in fig. 7) contains an image iconoscope [4]), type 5854, two output valves for the saw-tooth deflection currents and a pre-amplifier for the picture signal. The camera is provided with a Leitz "Elmar" lens, having a focal length $f =$ 58 mm and an aperture f : 3.5. With regard to the choice of the lens, it should be taken into account that the image reduction used here is only (5-6) : 1, which is, of course, far less than that for which normal camera lenses are designed. For line conversion, a lens turret and a view finder [7]) were unnecessary, so that the camera could be kept very small (fig. 7).

Owing to the high luminance of the light spot, the output signal of the image iconoscope was of such strength that no special provisions regarding the signal-to-noise ratio were required. It was thus possible to employ a pre-amplifier of a conventional type. This circuit is shown in *fig. 8*.

The first tube is a double-triode ECC 81, the two halves being connected in parallel in order to obtain a low equivalent noise resistance. (The equivalent noise resistance of a valve is the value of the resistance, which, if inserted in the grid circuit, would cause a noise output equal to that produced by the tube itself.)

In order to prevent the second tube, a pentode EF 80, from also contributing substantially to the noise, this tube has been connected up as a triode, in order to avoid the partition noise due to the random division of the cathode-current between screen-grid and anode. The control grid of this tube is connected, via a resistor R_1 of 0.68 MΩ, to the input, which causes a certain negative feedback. In itself, this does not improve the signal-to-noise ratio, but it permits the use of a higher input resistance which does produce a better signal-to-noise ratio.

The third and the fourth tubes are pentodes EF 80, connected in the usual manner, the latter tube being arranged as a cathode follower. From the cathode of this tube, the picture signal is fed to a second amplifier via a cable.

The high input resistance of the pre-amplifier causes a considerable attenuation in the higher frequencies. For this reason the second amplifier has been given a frequency-response that compensates for this attenuation. The clamping circuit incorporated in this amplifier has been dealt with earlier in this article.

[7]) Normal television cameras contain a view-finder either of the optical type (with large-aperture lens) or of the electronic type (small picture tube with associated components). Both types are fairly bulky in comparison with the other components of the camera.

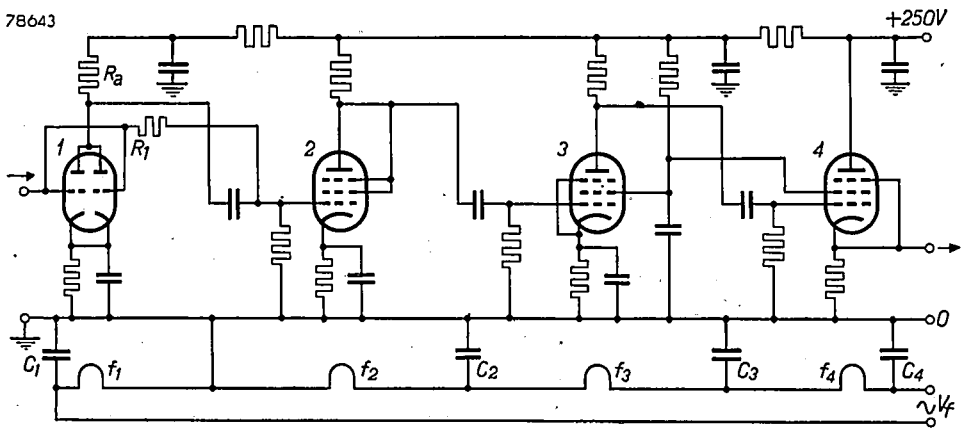


Fig. 8. Circuit diagram of the pre-amplifier in the camera. *1* double triode ECC 81 (the two sections connected in parallel) with anode resistor $R_a = 4200\ \Omega$. *2* pentode EF 80, connected as a triode. *3* and *4* normally connected pentodes EF 80, the latter operating as a cathode follower. R_1 (0.68 MΩ) negative-feedback resistor. The filaments (f_1, f_2, f_3, f_4) are connected in series (filament voltage V_f) and are earthed for high frequencies via the capacitors C_1, C_2, C_3 and C_4.

In the second amplifier, the blanking and synchronizing pulses are added to the picture signal, to form the complete video signal. These pulses are derived from a signal generator. The repetition-frequency of the frame-synchronizing pulses is governed by a sinusoidal auxiliary voltage. This voltage is derived, either from the saw-tooth current for the vertical deflection of the picture tube, or, when this is not feasible, from the mains. In the former case, the 625-line system is given the same picture frequency as the incoming signal, with the advantages outlined above.

The high speeds at which the photo-electrons hit the target in the image iconoscope render it fairly insensitive to stray electric and magnetic fields, but certain precautions are nevertheless required if maximum definition is to be attained. Especially the magnetic fields of the filament-current transformer and of the filament-current leads can be disturbing factors. For this reason the filaments of the pre-amplifier tubes are connected in series (fig. 8); so also are those of the output valves for the deflection currents. In this way the total filament current is considerably smaller, so that it was possible to locate the filament-current transformer away from the camera.

Some final observations

Very satisfactory results were obtained with the equipment described above. Some deterioration of the definition was, of course, unavoidable, but this was only slight. The gradation of the incoming signal was reproduced with great fidelity in the converted image, without gamma correction being necessary. Spurious signals from the image iconoscope were hardly discernible, thanks to the ample luminance of the intermediate optical picture (this would have been more difficult to attain with a *larger* picture tube). Another consequence of the high luminance was that the signal-to-noise ratio of the output signal was only slightly less than that of the incoming signal. Neither the grain, nor any other visible inhomogeneities of the phosphor presented any special difficulties.

Summary. A description is given of the line converter which was used to convert the British (405-line) television signal into a 625-line signal for transmission in the Netherlands and Western Germany on the occasion of the Coronation in London in June 1953. The apparatus was also capable of converting signals of the French standard (819 lines) into the 625-line system. The line converter uses a picture-tube to produce an intermediate optical image of the incoming signal. This image is viewed by an image iconoscope (type 5854) which scans it in the 625-line system. The disturbing effects which occur during this process are dealt with, as well as the means for eliminating them. The latter include: the use of a long afterglow phosphor in the picture tube, spot-wobbling in the picture tube, clamping of the black level in the output signal at the beginning of each scanning line, and the equalizing of the picture frequency of the 625-line system to that of the incoming signal (when the conditions are suitable). The picture tube has a screen of 12 cm diameter, the long afterglow phosphor being willemite. The luminance reaches a maximum value of 5000 nit. Its mean value is ample to ensure that the image iconoscope is almost free from spurious signals, and that the signal-to-noise ratio of the signal is only slightly diminished on passing through the line converter.

AN INSTRUMENT FOR MEASURING GROUP DELAY

by H. J. de BOER and A. van WEEL. 621.317.342.018.782.4:621.397.6

In radio receivers it is important that each frequency component of the audio signal is equally amplified, while phase shifts in the signal components of different frequencies have little effect on the quality. On the other hand, phase shifts in the video signal of a television receiver are very important; hence the transmission time, or delay of a signal passing through the receiver should be independent of frequency. An instrument with which differences in transmission time can be measured accurately to within 10^{-9} sec is described in this article.

Introduction

Phase delay

Distortion of several kinds may be introduced into an electrically transmitted signal by the transmitting network (cable, amplifier, transmitter, etc.). One type is associated with the phase delays at the different frequencies involved.

To define "phase delay", let us consider the case of a signal $A \cos(\omega t + \beta)$ applied to the input of a network; on reaching the output, this signal will have become

$$A' \cos(\omega t + \varphi + \beta).$$

where φ is the phase shift introduced by the network.

Writing

$$\tau_f = -\frac{\varphi}{\omega}, \quad \ldots\ldots \quad (1)$$

the output signal will be

$$A' \cos\{\omega(t - \tau_f) + \beta\}.$$

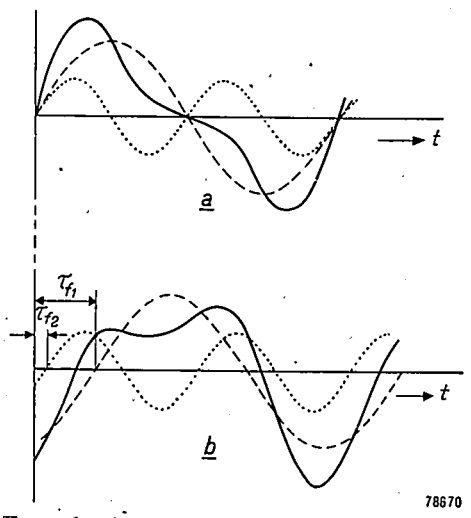


Fig. 1. Example of phase distortion. *a*) Input signal of a network. The signal contains components of angular frequency ω and 2ω as indicated by dotted lines.
b) Output signal of the same network. The phase delay τ_{f_1} of the component of angular frequency ω is assumed to exceed the phase delay τ_{f_2} of the component of angular frequency 2ω. It will be seen that phase distortion takes place.

Accordingly, the phase of the output signal at time $t_1 + \tau_f$ is the same as that of the input signal at time t_1; hence τ_f might be called the transmission time. It is generally known however as the phase delay.

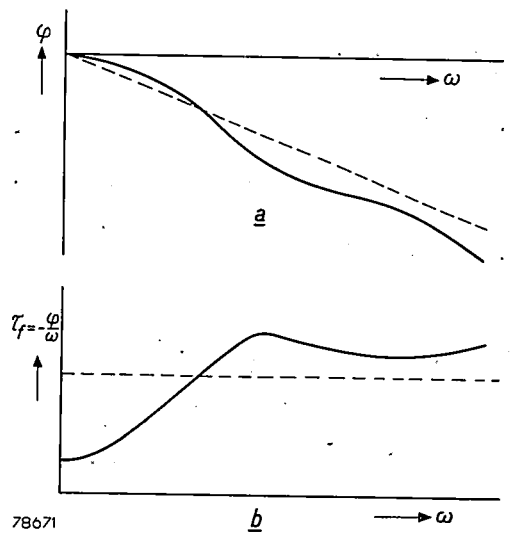


Fig. 2. *a*) Phase characteristics of a network that introduces phase distortion (full line), and that of a network which introduces no distortion (dotted line).
b) Phase delay characteristics derived from phase characteristics of (*a*).

The signal may be thought of as resolved into Fourier components which, if the phase delay depends on the frequency, will be displaced relative to one another when passing through the network. The distortion caused by these differences in delay is called phase distortion (*fig. 1*).

To avoid phase distortion, then, it is necessary to ensure that the phase delay is independent of frequency within the range defined by the Fourier components of the signal.

According to (1), this means that the phase shift φ introduced into a sinusoidal signal by the network must be directly proportional to the frequency within the range concerned.

Phase characteristics for networks with, and without phase distortion are illustrated in *fig. 2a*.

Fig. 2*b* shows that the delay is dependent on frequency in the first case, but not in the second.

Modulation phase delay

With a modulated carrier-wave signal the situation is different. Consider *figures 3a* and *b* which illustrate respectively the phase characteristic and the amplitude characteristic (amplitude gain a as a function of the angular frequency) for a typical radio-frequency network.

Let us assume that a carrier wave (angular frequency ω_0), whose amplitude is modulated sinusoidally to a depth m at a frequency $\omega_m/2\pi$, is applied to the network concerned. Such a signal (apart from a constant factor) is given by

$$S_i = (1 + m\cos\omega_m t)\cos\omega_0 t = \cos\omega_0 t + \frac{m}{2}\cos(\omega_0+\omega_m)t + \frac{m}{2}\cos(\omega_0-\omega_m)t. \quad (2)$$

According to (2) this signal can be resolved into three components that will, in passing through the network, attain values that can be deduced from diagrams 3*a* and *b*. At the output of the network a signal corresponding to

$$S_u = a_0\cos(\omega_0 t+\varphi_0) + a_1\frac{m}{2}\cos\{(\omega_0+\omega_m)t+\varphi_1\} + a_2\frac{m}{2}\cos\{(\omega_0-\omega_m)t+\varphi_2\} \quad (3)$$

is thus produced.

A trigonometrical conversion of the above (see appendix) shows that the expression for S_u represents a carrier-wave with an angular frequency ω_0, whose amplitude is also modulated periodically at a fundamental frequency $\omega_m/2\pi$. We are not at present concerned with the higher harmonics also produced during this process.

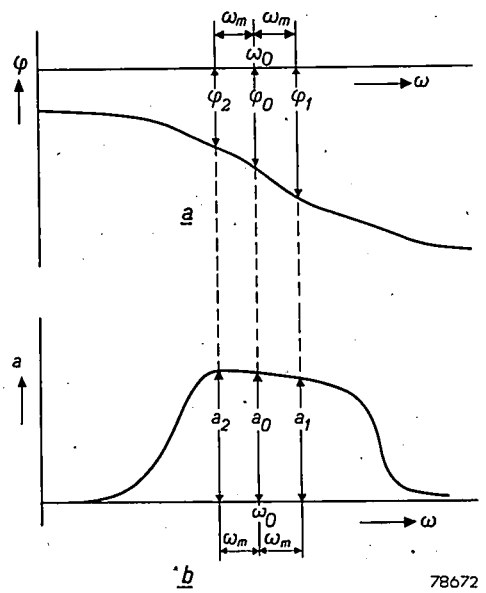


Fig. 3. Phase characteristic (*a*) and amplitude characteristic (*b*) of a typical band-pass network for high frequencies.

It can be seen from the derivation in the appendix that if the original depth of modulation is small ($m \ll 1$), the fundamental component of the modulation is displaced through an angle ψ with respect to the modulation of the input signal, ψ being given by:

$$\tan\psi = \frac{a_1\sin(\varphi_1-\varphi_0) - a_2\sin(\varphi_2-\varphi_0)}{a_1\cos(\varphi_1-\varphi_0)+a_2\cos(\varphi_2-\varphi_0)}. \quad (4)$$

The time-dependent factor $\cos\omega_m t$ in the modulation of the input signal is thus converted into $\cos\omega_m(t+\psi/\omega_m)$ in the output signal; hence the "modulation phase delay" is:

$$\tau_{fm} = -\psi/\omega_m. \quad (5)$$

Phase distortion takes place in a modulation-transmitted signal if the modulation phase delay depends on the frequency of modulation. According to (5), then, a curve representing the phase shift ψ (defined by (4)) as a function of ω_m, for the band-pass sections of a television receiver (R.F. and I.F. stages), should be a straight line passing through the origin. In these circumstances the modulation phase delay $-\psi/\omega_m$ takes the place of the "normal" phase delay $-\varphi/\omega$ described earlier (see equation 1).

To study the effect of phase distortion in television receivers, we require a sufficiently accurate method of measuring differences in normal phase delay (in the video stage) and differences in modulation phase delay (in the R.F. and I.F. stages).

In principle, this can be accomplished by plotting the phase characteristics of the networks concerned; normal phase delays can be derived direct from these characteristics (see (1)). However, for the determination of modulation phase delay from (4) and (5), that we shall now consider, the amplitude characteristics of the R.F. and I.F. stages are also required. Although the required characteristics can be measured at high frequencies (several tens of Mc/s) [1]), the results thus obtained are not accurate enough for our purpose. The values of φ_0, φ_1 and φ_2 derived from the phase characteristics are generally large in relation to the differences $\varphi_1-\varphi_0$ and $\varphi_2-\varphi_0$ which must be known in order that formula (4) can be applied.

[1]) See, e.g. G. Thirup, An instrument for measuring complex voltage ratios in the frequency range 1—100 Mc/s. Philips tech. Rev. **14**, 102-114, 1952.

An instrument (*fig. 4*) with which direct and very accurate measurements of the slope of the phase characteristic ($d\varphi/d\omega$ as a function of ω) can be made will now be described.

For a reason which will be explained later, the quantity

$$\tau_g = -d\varphi/d\omega \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

is known as "group delay"; hence the instrument is described as a "group delay meter". Differences in the phase shifts associated with different frequencies can be established by integration (e.g. graphic integration) of the group delay.

Thus:

$$\varphi_1 - \varphi_0 = -\int_{\omega_0}^{\omega_0+\omega_m} \tau_g \, d\omega, \quad . \quad . \quad . \quad . \quad . \quad (7)$$

A similar expression applies to $\varphi_2 - \varphi_0$.

The instrument is designed to measure group delay accurately to within 10^{-9} sec; according to an estimate by Ring [2]) this is the degree of accuracy which should be aimed at.

To give some idea of the basis of this estimate, it may be seen from (1) that at a frequency of 5 Mc/s (the highest frequency in a video signal for 625-line television) a phase difference of 2π radians between signals at the input and output of a network corresponds to a phase delay of 200×10^{-9} sec.

[2]) D. H. Ring, The measurement of delay distortion in microwave repeaters, Bell Syst. tech. J. **27**, 247-264, 1948.

Group delay

The quantity $\tau_g = -d\varphi/d\omega$ is described as "group delay" by reason of its significance in the simple case of a frequency characteristic that is straight within the given range of interest. In this case:

$$\varphi_1 - \varphi_0 = -(\varphi_2 - \varphi_0) = -\tau_g \omega_m,$$

where τ_g is constant.

Substitution of the above in (4) gives us:

$$\psi = -\tau_g \omega_m, \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

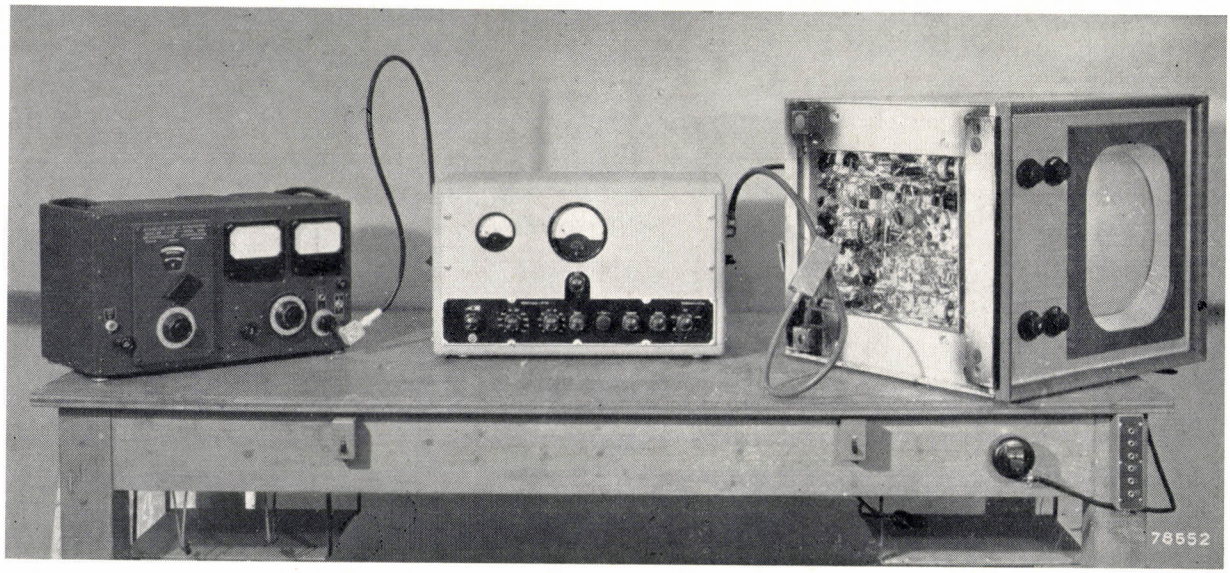


Fig. 4. The group delay meter (centre) as used for testing the R.F. and I.F. stages of a television receiver (on right of picture). At the extreme left is a signal generator producing the carrier wave required for the test. The equipment illustrated can be used for tests at carrier wave frequencies between 10 and about 100 Mc/s, the results being accurate to within 10^{-9} sec.

hence, according to (5)

$$\tau_{fm} = \tau_g. \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

so that τ_{fm} remains independent of ω_m, provided that $\omega_0 + \omega_m$ and $\omega_0 - \omega_m$ are associated with a

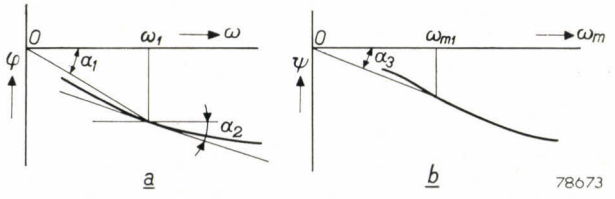


Fig. 5. *a*) Geometrical significance of phase delay and group delay. A phase characteristic is shown. The phase delay τ_f for an angular frequency ω_1 is $-\tan\alpha_1$ (see equation 1); the group delay τ_g for the same angular frequency is $-\tan\alpha_2$ (see equation 6).
b) The modulation phase delay characteristic for a certain carrier wave frequency. The angular frequency of modulation ω_m is plotted on the abscissa. The modulation phase delay at a modulation frequency ω_m corresponds to $-\tan\alpha_3$ (see equation 5) and so has in this graph the same mathematical significance as τ_f in (*a*).

portion of the phase characteristic which is adequately straight. Accordingly, a modulation signal comprising components with angular frequencies ω_m within this range will pass through the network with a delay τ_g without distortion.

A certain mathematical significance, as demonstrated in *fig. 5* (see caption), can be assigned to the different delays introduced.

Measurement of group delay

Let us assume that a signal of angular frequency ω is applied to a network whose group delay (i.e. slope of the phase characteristic) is to be determined at that angular frequency. The signal is modulated at an angular frequency, $\omega_m = p$, which is so low that the phase characteristic associated with any angular frequency band of width $2p$ may be considered straight.

The phase shift ψ associated with p and introduced into the modulation by the network will be denoted by ϑ, so that from (8),

$$\vartheta = -\tau_g p. \quad \ldots\ldots \quad (10)$$

At the fixed value p, then, ϑ is a measure of the group delay to be determined. The measurement of group delay is thus reduced to a measurement of the phase shift introduced into the modulation signal of angular frequency p by the network [3]).

The choice of a suitable modulation angular frequency p is limited by conflicting requirements. According to (10) sensitive measurement necessitates a high value of p, by virtue of which a small variation in τ_g will produce a substantial variation in ϑ. On the other hand, a low value of p is required for accuracy, since the approximation of the phase characteristic to a straight line will be better for a smaller value of p. Equation (10), too, is a better approximation in these circumstances. A frequency $p/2\pi = 25$ kc/s has been adopted as a suitable practical value for the present instrument; hence, from (10), it will be necessary to measure phase variations of 0.01° in order to determine group delay differences of 10^{-9} sec. Accordingly, the main problem is to design a very sensitive phase meter.

Principle of the phase meter

A property common to oscillator circuits is used in the measurement of small phase differences.

In principle, an oscillator is an amplifier of which all, or a portion of the output is fed back to the input to form a loop. At a given frequency, stable oscillation is possible only if the phase shift introduced into a signal at this frequency as it passes round the loop is a multiple of 2π. Every oscillator circuit contains a frequency-determining element the essential feature of which is that it imparts to the signal a phase shift, dependent (usually very sharply) on the frequency, such that the phase condition is exactly satisfied at the desired frequency. The circuit will commence to oscillate at the frequency concerned, provided that the amplification at this frequency exceeds unity. The amplitude of the oscillation at first increases, but is limited by an accompanying decrease in amplification until unity is reached.

In simple oscillator circuits the amplifier valves themselves act as amplitude-limiting elements, since their effective slope decreases as the amplitude of the grid voltage increases. However, it will be seen that in the present case, a separate, carefully designed amplitude limiter is required.

If an extra phase shift is introduced in some way into an oscillating circuit, stable oscillation is possible only at a frequency so far removed from the original oscillation frequency that the phase shift in the frequency-determining element due to the frequency variation, compensates for the introduced phase shift; in this way the phase condition (that the total phase shift in the loop is a multiple of 2π) is again satisfied.

If the condition as regards amplification is also met at the new oscillation frequency, the circuit will continue to oscillate at this frequency. It will be seen, then, that a certain relationship exists between a phase shift introduced into a signal and the change in frequency thereby produced. This relationship is the basis of the phase meter used in the present instrument. Further details are given below and in *fig. 6*.

Block diagram of the group delay meter

The diagram depicted in fig. 6 includes a selective amplifier A, the pass-band of which covers the modulation frequency employed, $p/2\pi = 25$ kc/s. The output voltage of the amplifier is applied to a modulator B, to which is connected a signal generator G producing a carrier wave whose frequency $\omega/2\pi$ is variable to cover the pass-band of the network X. The output signal of the modulator passes through network X and is then rectified by the detector C. By connecting the output of the latter to the input of the amplifier, a feedback loop is formed, so that oscillation can take place at the frequency p. The amplifier is therefore referred to

[3]) This principle was first applied to the measurement of phase distortion in long distance telephone communication. See H. Nyquist and S. Brand, Bell Syst. tech. J. 9, 522-549, 1930.

below as the p-amplifier. Special measures are taken to ensure that the phase condition can be satisfied in the frequency band about p.

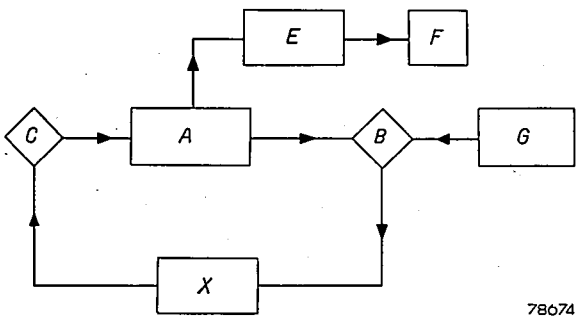


Fig. 6. Block diagram of the equipment used for measuring group delay as a function of frequency. Oscillation at a nominal frequency of 25 kc/s takes place in the loop formed by amplifier A, modulator B, network on test X and detector C. A phase shift in X produces a change in frequency which is converted by discriminator E into a voltage variation. The latter is measured by means of a valve voltmeter F. The signal generator G supplies a carrier wave of the frequency at which the behaviour of network X is to be investigated; this carrier wave is modulated with the oscillation frequency.

The exact value of the frequency of oscillation $p/2\pi$ is determined mainly by the frequency-determining element in the amplifier, but is also affected by the phase shift introduced into the modulated signal as it passes through network X. According to (10) the phase shift ϑ is related to the group delay of X at an angular frequency ω; hence a relationship exists between this group delay and the oscillation frequency. This relationship can be established in the following manner.

Relationship between group delay and frequency of oscillation

In the event of a change $\Delta\omega$ in the carrier-wave angular frequency ω, the group delay τ_g of X varies by $\Delta\tau_g$. At the same time, the angular frequency of oscillation changes from p to $p+\Delta p$, in such a way that the phase condition is still satisfied.

The phase shift of the modulation in network X thus changes (see equation 10) from $-p\tau_g$ to $-(p+\Delta p)(\tau_g+\Delta\tau_g)$, that is to say (neglecting the term $\Delta p\,\Delta\tau_g$) by an amount

$$\Delta\vartheta = -(p\Delta\tau_g + \tau_g\Delta p).$$

In the remainder of the loop, the phase shift changes by an equal amount but opposite in sign, $\Delta'\vartheta$. Let us now assume for a moment that the network X is short-circuited and the resultant loop is opened at the input terminals of the amplifier. If the phase characteristic of the network so obtained is straight in the region of 25 kc/s (it is easy to ensure that this is so), and its slope is $d\varphi/d\omega = -t_0$, we may write:

$$\Delta'\vartheta = -t_0\,\Delta p.$$

Evidently, t_0 represents the group delay of the open-circuit network (without X) for a signal modulating a 25 kc/s carrier wave. The quantity t_0 may be described as the inherent group delay of the instrument.

Equating $\Delta\vartheta$ and $-\Delta'\vartheta$, we find that:

$$\Delta\tau_g = -\frac{t_0+\tau_g}{p}\Delta p.$$

The measuring instrument is so constructed that $t_0 \gg \tau_g$; hence, ignoring the sign,

$$\Delta\tau_g = \frac{t_0}{p}\Delta p. \quad \ldots\ldots \quad (11)$$

Since t_0 and p are known, $\Delta\tau_g$ can be determined [4]) by measuring Δp. The difference of phase delay $\Delta\tau_g$ thus established differs from τ_g itself only by an (unknown) constant.

A closer investigation shows that by using $\Delta\tau_g$ instead of τ_g in (7), and applying the result in (4) and (5), the modulation phase delay τ_{fm} is found, apart from the same constant. The omission of the constant does not invalidate the calculation since we are concerned only with differences in modulation phase delay.

Particulars of the instrument

The amplifier

The oscillation frequency should depend only on the group delay of the network on test; in the absence of such a network, then, the frequency should be constant.

In normal circumstances, highly selective circuits (e.g. crystal-controlled) are used in the design of stable oscillators. The salient feature of such circuits is that their phase characteristic is steep in the region of their resonant frequency (high t_0); the frequency variations produced by slight phase shifts within the circuit are accor-

[4]) The variation in group delay thus established relates to the portion of the loop through which the signal is transmitted on the carrier wave. Besides network X, this portion includes parts of the modulator and detector circuits. Special measures are then necessary to ensure that the amount contributed to the group delay by the modulator and the detector does not depend on ω in the range of frequencies concerned, so that only network X can produce a variation in the group delay. These measures involve the use of different modulators and detectors for different frequency bands.

dingly very small. This, however, applies to any phase variation, including the phase shift which is to be measured; hence this type of oscillator is not suitable for the present purpose. The sensitivity of measurement would be small, and, of course, the ratio of the effects of desired and undesired phase variations is not improved.

Stability of frequency in the oscillator must be obtained not by reducing the *effect* of undesired phase variations, but by avoiding such variations as far as possible.

This is accomplished by including amplifier stages of two different kinds in the instrument. The first kind of stages includes an anode circuit having an LC-circuit damped by a parallel resistor R. This LC-circuit, tuned to 25 kc/s, is the frequency-determining element of the oscillator. Apart from the sign, the phase shift produced by this amplifier stage is equal to the phase angle α of the anode impedance. A simple calculation shows that in the region of resonance:

$$\mathrm{d}\alpha/\mathrm{d}\omega = 2\,RC.$$

Accordingly, this also represents the amount contributed by the amplifier stage to the inherent group delay t_0 of the measuring instrument.

Secondly, a number of resistance-coupled amplifier stages with low anode resistance are used; these also cause slight phase shifts due to the valve and wiring capacitances in parallel with the anode resistance, but it is found that the magnitude of such shifts is not much affected by variations in the resistance and capacitance values, provided that a low value is chosen for the resistance. The stage gain is then small; hence several stages must be used to obtain the required overall gain. Variations in the phase shift that may occur within this portion of the instrument are so small that they may be ignored, and also the total contribution of the resistance-coupled amplifier stages to the inherent group delay t_0 of the instrument is so insignificant, that t_0 may be taken as equal to the contribution of the tuned stage only. Hence (see above),

$$t_0 \approx 2\,RC.$$

The value of RC is so chosen that $t_0 = 60$ μsec. The requirement $t_0 \gg \tau_g$ is thus satisfied for practical values of τ_g, and at the same time adequate sensitivity is maintained.

The amplitude limiter

Since an amplitude-limiting element in an oscillator cannot be linear, it inevitably introduces harmonics of the oscillation into the circuit. If a signal component at twice the fundamental frequency, as well as the fundamental itself, passes round the loop and reaches the input of the limiter, a beat frequency will be produced at the output with the fundamental frequency $p/2\pi$. This is known as conversion.

In general, this undesired beat signal will not be in phase with the desired signal; hence the resultant signal of frequency p (the p-signal) suffers a phase shift the angle of which will depend on the amplitude and phase of the undesired signal. If this phase shift were the same for all the frequencies within the measuring range, an equal amount would be added to each of the values of $\Delta\tau_g$ read from the instrument; this would not matter, since we are concerned only with group delay differences. Unfortunately, however, the amplitude of the undesired signal, and with it the effect on the phase of the resultant signal, will depend on the self-adjustment of the amplitude limiter, which will be such that the amplification of the p-signal by the loop is exactly unity. Since the amplification of the p-signal in the X-network will not as a rule be constant over the entire measuring range, the effect of the undesired p-signal will also vary.

Two measures have been adopted to limit this interference. Firstly, an almost distortionless amplitude limiter is used, and secondly the selective circuit included in the amplifier substantially attenuates the higher harmonics.

The discriminator

Frequency variations are measured by means of a frequency discriminator, producing at its output a D.C. voltage which is proportional to the difference between the frequency of the signal applied to the input and a certain fixed frequency. At the values of t_0 and p employed, a variation in group delay of 10^{-9} seconds corresponds to a frequency variation of about 0.5 c/s (see equation (11)); hence the fixed frequency must remain constant to within considerably less than 0.5 c/s if the group delay is to be measured accurately to within 10^{-9} sec. Moreover, the discriminator must be adequately sensitive.

To satisfy the latter requirement, the output voltage of the discriminator (which is directly proportional to the group delay variations to be determined) is measured with a valve voltmeter calibrated directly in nanoseconds (10^{-9} sec). The four measuring ranges covered by the voltmeter correspond to 0—1000, 0—300, 0—100 and 0—30 nanoseconds.

"Delay switching"

If an oscillation is produced in the loop in the absence of an unknown network X, the amplifier can be so tuned that the oscillation frequency coincides with the fixed frequency at which the output voltage of the discriminator is zero; measurement on an unknown network subsequently introduced into the circuit then produces the absolute group delay of this network. But we are concerned with variations in the group delay within a certain range of frequencies, rather than with the absolute value. If the absolute delay is relatively long (e.g. 500×10^{-9} sec) and the variations amount to no more than a few times 10^{-9}, the latter would have to be read from the 1000×10^{-9} sec scale.

To overcome this difficulty, means are prodived whereby known phase shifts can be introduced into the amplifier. This is accomplished quite simply by connecting different capacitors in parallel with a known resistance. A phase shift $\Delta\vartheta$ in the measuring instrument affects the oscillation frequency in exactly the same way as a change $-\Delta\vartheta/p$ in the group delay of the unknown network (see equation (10)). It is thus possible to "switch out" a known group delay. The instrument is so designed that group delay can be switched out in steps of 100×10^{-9} and 10×10^{-9} sec.

Testing circuits containing a detector

Many of the circuits to be tested (e.g. television receivers) include a detector and, when this is so, the detector normally included in the measuring instrument can be omitted; for example, the equipment illustrated in fig. 4 does not include a detector.

A detector in the circuit on test is usually followed by another network — in television receivers, the video stage — through which the signal passes before reaching the p-amplifier. If the phase characteristic of this network is not horizontal in the region of the angular frequency p, it affects the value of t_0 (see equation 11). The video stage of a television receiver, however, is a broad-band amplifier containing no selective circuits; hence its phase characteristic is so flat that it does not seriously affect t_0. Accordingly, equation (11) continues to apply to delay variations in that portion of the receiver in which the carrier wave is modulated by the signal.

A signal which, in the set on test, passes through an uneven number of valves after leaving the detector, would be exactly π radians out of phase after travelling once round the measuring circuit, so that oscillation would not take place if special measures were not taken to correct this phase shift. Accordingly, a special switch is provided on the instrument with which the phase can be increased by π radians.

The arrangement used for testing a television receiver is illustrated in fig. 4.

Tests at video frequencies

Whereas the preceding arguments are based on the assumption that the p-signal travelling through the circuit on test is always in the form of modulation, the modulator really produces, apart from the modulated carrier wave, other signals, including one with an angular frequency of p. After passing through the circuit on test this direct p-signal will affect the phase of the desired signal in a manner that, since it cannot be predicted, may well upset the group delay measurement.

During measurements on band-pass portions of the circuit on test, the direct p-signal in this circuit will be so attenuated that it cannot reach the p-amplifier, but if the pass band of the network also includes angular frequency p, as it does in video-frequency amplifiers, this suppression will not take place. In these circumstances, special measures must be taken to ensure accuracy in the measurement. For this reason, an adaptor has been designed which, combined with the group delay meter described, enables measurements at frequencies between about 0.1 and 10 Mc/s to be carried out.

The adaptor contains a balanced modulator, which does the work of the modulator in the original equipment, and a balanced detector. In principle, these balanced circuits should produce no p-signal at the output [5]), but in fact the suppression is never complete. The phase of the final signal is most seriously affected when the phase difference between the residual signal at the output of the detector and the desired signal travelling as modulation through the circuit on test is $\pi/2$ (*fig. 7*).

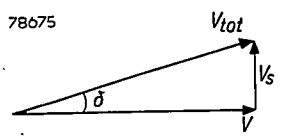


Fig. 7. Error introduced by an interfering signal V_s. The desired signal is V and the total signal V_{tot}. The phase difference δ between V and V_{tot} reaches a maximum when V_s and V are about $\pi/2$ out of phase, but is zero (and thus introduces no error) when V and V_s are either in phase, or in antiphase.

On the other hand, no error is introduced when the phase difference between the signals is either 0 or π; hence it is essential to ensure that the difference obtained corresponds as closely as possible to one of these two values. Further analysis has shown that this can be accomplished by including in the modulator and detector special features whereby the components of angular frequency p, produced by the two push-pull valves in their

---

[5]) See, e.g. F. A. de Groot and P. J. den Haan, Modulators for carrier wave telephony, Philips tech. Rev. 7, 83-91, 1942.

common signal, are kept as far as possible in anti-phase (phase balancing).

It is found by experiment that a residual signal constituting 1% of the desired signal introduces a perceptible error into the test results if produced by a disturbance of the phase balance, whereas a similar signal produced by a disturbance of the amplitude balance scarcely affects the results.

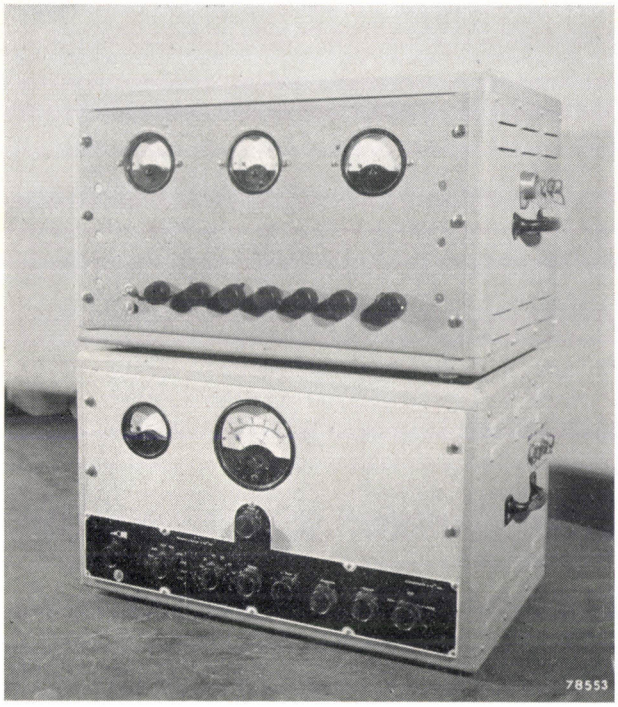


Fig. 8. Combination of group delay meter and video adaptor (top) used for group delay measurements in the frequency band between about 0.1 and 10 Mc/s.

In the circuit described here, as in the resistance-coupled stages of the amplifier, only resistors of low value are employed; hence there can be little variation in the phase balance when once adjusted.

The adaptor also contains a selective test amplifier, tuned to angular frequency p, which can be used to control and check the balance of the modulator and the detector when connected to the output of either of these units.

The combination of a group delay meter and an adaptor is illustrated in *fig. 8*. This combination is used for group delay measurements in the frequency band of 0.1 to 10 Mc/s, and the group delay meter without adaptor for similar measurements in the band of 10 to about 100 Mc/s. In principle, there is no maximum limit to the frequencies for which the above method can be employed, but in practice a limit is imposed by the fact that suitable modulators and detectors must be built for each frequency band (see note [4]). A rather complex relationship exists between the lower frequency limit and the magnitude of p. Measurements at frequencies down to about 60 kc/s have been achieved with certain precautions.

In order to calculate the phase delay it is necessary to extrapolate the measured group delay differences to $\omega = 0$. Like modulation phase delay, the phase delay thus calculated lacks an additive constant.

Some test results

As a general example of the kind of information obtainable with the equipment described here, curves plotted from some of the results are shown. The first of these relate to tests on a television receiver; the R.F. and I.F. stages of this receiver were tested together and a curve was constructed representing the difference of the group delay from that at the carrier wave frequency of the television transmitter (in this case *Lopik* at 48.25 Mc/s) to which the receiver was tuned (*fig. 9a*). This curve, together with the amplitude characteristics of the R.F. and I.F. stages (which were known), was used to compute the modulation phase delay τ_{fm} (apart from an additive constant). It is shown as passing through the origin in fig. 9*b*.

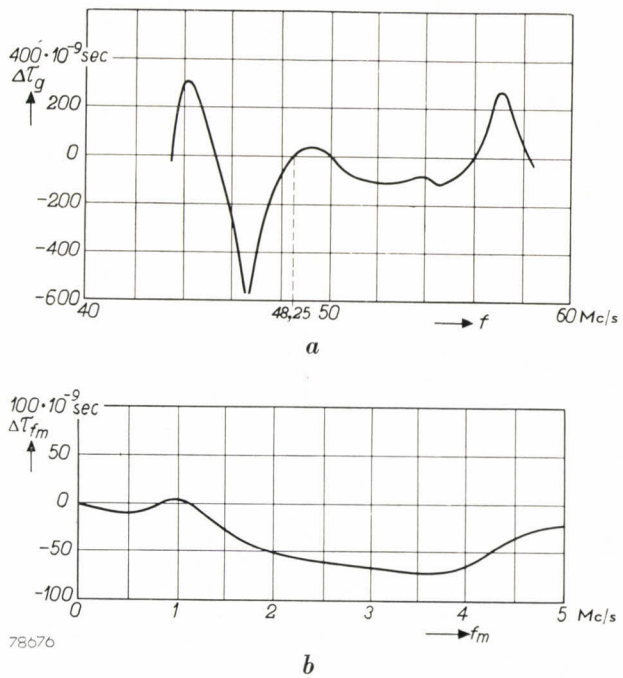


Fig. 9. Group delay characteristics of the combined R.F. and I.F. stages of a television receiver. The differences $\Delta\tau_g$ with respect to the group delay at the carrier wave frequency of the Lopik television transmitter (48.25 Mc/s) are shown. *b*) Modulation phase delay of the combined R.F. and I.F. stages of the television receiver. The quantity $\Delta\tau_{fm}$, derived from the group delay characteristic (*a*), is plotted as a function of the modulation frequency, taking into account the amplitude characteristics of the parts of the network on test.

Another curve, representing the group delay as measured in the video stage between the detector and the picture tube of the same television receiver, is illustrated in *fig. 10*. This illustration also includes the normal phase delay characteristic derived from the group delay characteristic.

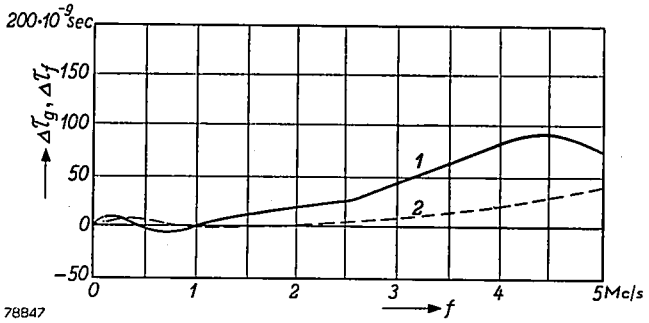


Fig. 10. Delay characteristics of the video stage of a television receiver. Curve *1* represents the differences in group delay determined by direct measurement, and curve *2* the differences in phase delay derived from *1*.

The group delay characteristic of the complete receiver, i.e. for all stages between the aerial terminals and the picture tube, will be seen in *fig. 11*. Two modulators were used to plot this characteristic, one modulating the video signal with an angular frequency p, and the other modulating a high-frequency carrier wave with the modulated video signal, which is then applied to the aerial terminals. The video signal after detection in the video detector of the receiver, passes through the video stage of the latter and is in turn detected by the detector included in the adaptor. The phase delays derived from the measured group delay characteristic are also shown in the illustration (fig. 11). Also included, as a check, are the same phase delays computed as the sum of the modulation phase delays in the R.F. and I.F. stages (fig. 9*b*), and the phase delay in the video stages (fig. 10).

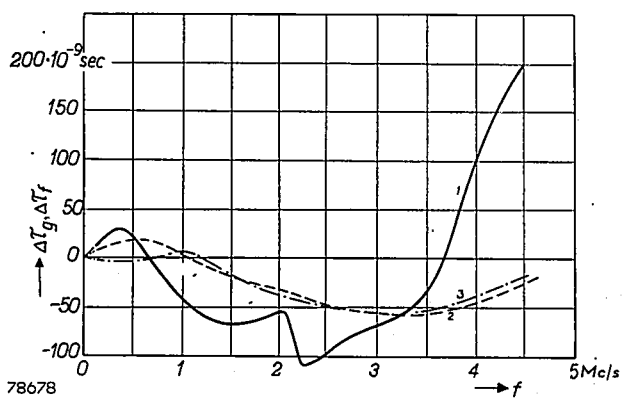


Fig. 11. Delay characteristics of a complete television receiver (from aerial terminals to picture tube). Curve *1* represents the differences in group delay $\Delta\tau_g$ determined by direct measurement, curve *2* the differences in phase delay $\Delta\tau_f$ derived from *1*; curve *3* also represents $\Delta\tau_f$, this time determined as the sum of $\Delta\tau_f$ for the video stage (fig. 10) and $\Delta\tau_{fm}$ for the R.F. and I.F. stages (fig. 9*b*).

Lastly, the delay characteristics of an amplifier specially designed for a television studio are illustrated in *fig. 12*. It will be seen that the phase delay characteristic thus obtained is very flat and hence that the amplifier concerned introduces hardly any phase distortion.

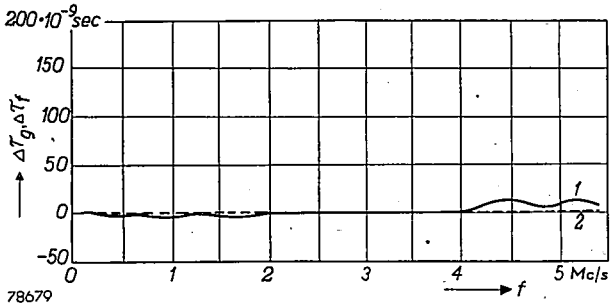


Fig. 12. Delay characteristics of a special video amplifier designed for use in television studios. Curve *1* represents the measured differences in group delay, and curve *2* the differences in phase delay.

Appendix. Derivation of formulae for modulation phase delay

When applied to the input of a high-frequency network, a modulated carrier wave as defined by formula (2) produces at the output a signal of the form (formula (3), slightly rearranged):

$$S_u = a_0 \cos(\omega_0 t+\varphi_0)+a_1\frac{m}{2}\cos\{(\omega_0 t+\varphi_0)+(\omega_m t+\varphi_1-\varphi_0)\}+ + a_2\frac{m}{2}\cos\{(\omega_0 t+\varphi_0)-(\omega_m t-\varphi_2+\varphi_0)\}.$$

This may also be written as:

$$S_u = B_1\cos(\omega_0 t+\varphi_0)+B_2\sin(\omega_0 t+\varphi_0) = = \sqrt{B_1^2+B_2^2}\cos(\omega_0 t+\varphi_0'), \quad (12)$$

where:

$$\left.\begin{aligned} B_1 &= a_0+a_1\frac{m}{2}\cos(\omega_m t+\varphi_1-\varphi_0)+a_2\frac{m}{2}\cos(\omega_m t-\varphi_2+\varphi_0)\\ B_2 &= a_1\frac{m}{2}\sin(\omega_m t+\varphi_1-\varphi_0)-a_2\frac{m}{2}\sin(\omega_m t-\varphi_2+\varphi_0)\end{aligned}\right\} \quad (13)$$

It will be seen from (12) that S_u represents a carrier wave of angular frequency ω_0 and of amplitude $\sqrt{B_1^2+B_2^2}$. Since the signal is to be detected in an A.M. detector, the amplitude is all that concerns us.

Applying (13) we find:

$$B_1^2+B_2^2 = a_0^2+\frac{a_1^2m^2}{4}+\frac{a_2^2m^2}{4}+ + a_0a_1m\cos(\omega_m t+\varphi_1-\varphi_0)+a_0a_2m\cos(\omega_m t-\varphi_2+\varphi_0)+ + \frac{a_1a_2m^2}{2}\cos(2\omega_m t+\varphi_1-\varphi_2). \quad (14)$$

Omitting, for the moment, the factor a_0m common to the fourth and fifth terms, the following expression may be used for the sum of the two:

$$a_1 \cos(\omega_m t+\varphi_1-\varphi_0) + a_2 \cos(\omega_m t-\varphi_2+\varphi_0) =$$
$$\left.\begin{matrix} a_1 \cos(\varphi_1-\varphi_0) + a_2 \cos(\varphi_2-\varphi_0) \\ a_1 \sin(\varphi_1-\varphi_0) - a_2 \sin(\varphi_2-\varphi_0) \end{matrix}\right\} \begin{matrix} \cos \omega_m t - \\ \sin \omega_m t = \end{matrix}$$
$$C \cos(\omega_m t+\psi),$$

where C is a constant, and

$$\tan \psi = \frac{a_1 \sin(\varphi_1-\varphi_0) - a_2 \sin(\varphi_2-\varphi_0)}{a_1 \cos(\varphi_1-\varphi_0) + a_2 \cos(\varphi_2-\varphi_0)}.$$

This is formula (4).

If $m \ll 1$, the last term in (14) may be omitted; the amplitude of S_u then corresponds to:

$$\sqrt{B_1^2+B_2^2} = \sqrt{a_0 + \left(\frac{a_1 m}{2}\right)^2 + \left(\frac{a_2 m}{2}\right)^2 + C \cos(\omega_m t+\psi)}.$$

This periodic function can be expanded as a Fourier series containing only cosine functions, viz.:

$$\sqrt{B_1^2+B_2^2} = b_0 + \sum_{n=1}^{\infty} b_n \cos n(\omega_m t+\psi).$$

The fundamental term is $b_1 \cos(\omega_m t+\psi)$, the phase displacement of which with respect to the modulation of the input signal S_i is ψ (see equation 2). From this the modulation phase delay equation (5) can be derived.

Summary. Phase distortion takes place in a network if the time taken by a signal to travel through the network depends on the frequency of the signal. The time, or rather the delay involved may be either "normal", or "modulation phase delay", depending on whether the signal passes through the network direct, or as modulation on a carrier wave.

The normal phase delay can be derived direct from the phase characteristic of the network, but to determine the modulation phase delay the amplitude characteristic is also required. In practice the best method of establishing the phase characteristic is to determine its *derivative* as a function of the frequency, which, ignoring the sign, is designated the "group delay".

An instrument with which group delay in networks, and more especially in television receivers, can be measured accurately to within 10^{-9} sec. is described in this article. Group delay is determined by measuring phase differences; hence the instrument is really a very sensitive phase difference meter operating on the principle that the phase shift to be measured affects the frequency of an oscillatory circuit when introduced into the latter. A discriminator is used to convert the change in oscillation frequency thus produced into a voltage variation, which is then measured with a valve voltmeter calibrated in nanoseconds (10^{-9} sec) enabling direct reading of the group delay or in most cases a difference in group delay.

Measurements on the video stage of a television receiver, involve certain complications which necessitate the use of a special adaptor. The results of some practical measurements are included in this article as an example of the kind of information obtainable by means of the instrument.

THE SPECIFIC LIGHT OUTPUT OF "PHOTOFLUX" FLASH-BULBS

by L. H. VERBEEK. 771.448.4

The very small and inexpensive flash bulb, type PF 3, which was marketed in 1952, has been enthusiastically received by photographers of all countries. This article outlines the developments which have led to the manufacture of this bulb and of the similar, larger types (PF 14 and PF 25).

The open flashlight with magnesium powder in common use for indoor snapshots a quarter of a century ago, has now been almost entirely superseded by the well-known glass-enclosed flash-bulb.

The advantages of this lamp are many — it is more reliable than the open flashlight powder, there is no risk of fire, it has a shorter flash duration, and, finally, it made it possible to synchronize the flash with the opening and closing of the shutter — an essential feature of modern flash photography.

All these advantages, however, would not have ensured the present popularity of the flash-bulb, if it had not met one essential requirement, viz. to give a satisfactory light output at a reasonable price.

This implies that the bulb should not be too large, or in other words, that the specific light output of the bulb, i.e. the light output per unit volume should be sufficiently high. (The light output is defined as the number of "visual" lumen-seconds obtained from integration of the total area of the light-time characteristic, or "flash curve", *fig. 1*).

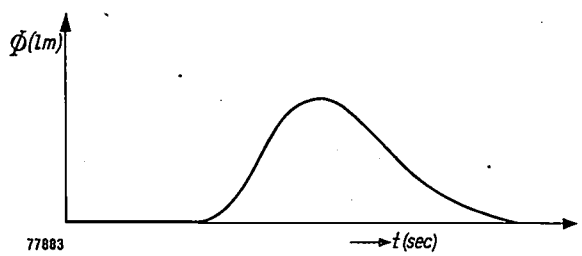


Fig. 1. Luminous flux Φ in visual lumens, emitted by a flash-bulb, plotted as a function of time [1]) (schematically). By integration over this "flash curve", the total emitted light in "visual" lumen-seconds is obtained. (For "visual" and "photographic" lumen-seconds, see the article referred to in [3]).)

Since the first introduction of "Photoflux" lamps by Philips, their development has been largely characterized by a constant striving after the greatest possible specific light output. The demands as regards total light output have not risen during the last twenty-five years; rather they have decreased owing to the use of faster lenses and more sensitive photographic emulsions. The demand for the smallest possible volume (and thus low price) for a given light-output, however, has become steadily more and more pressing. This has been partly due to competition between the various flash-bulb manufacturers, and partly also to the growing popularity of this source of lighting: the more flash-bulbs one wants to use, the more essential are both a low price and a small size. For the press photographer, the success of his work often depends upon the number of flash-bulbs he can carry in his pockets.

These demands have been answered by the impressive rise in the specific light output of the "Photoflux" bulbs since their introduction in 1932. This is shown in *fig. 2*. A certain reserve is necessary with respect to the earlier measurements (before 1945) incorporated in this graph, since at that time the measuring technique had not yet been standardized. Even if the error in the pre-1945 measurements is estimated to be as high as 20%, however, the enormous rise in the specific light output illustrated by fig. 2, is not invalidated.

The various jumps in this line of development correspond to certain alterations in the construction, or in the manufacturing methods of the "Photoflux" bulbs. The last great jump, made about a year and a half ago, corresponds to the marketing of the PF 3, which is the smallest flash-bulb ever made *). In this article a short description will be given of the modification in flash-bulb manufacture which made this last jump possible. As an introduction to this, it is desirable to consider in some detail, the steps in the development of flash-bulb design, corresponding to the previous jumps depicted in the graph.

In the very first "Photoflux" bulbs, the light was produced by the explosive reaction of two gases, approximately 30% by volume of carbon disul-

[1]) A method for recording such flash curves has been described by J. A. M. van Liempt and J. A. de Vriend, Rec. Trav. chim. Pays-Bas **52**, 160, 1933. For the measuring method standardized at present, see Am. Federal Specification WL 122.

*) Note added in proof: The development in America of a still smaller flashbulb has recently been announced.

phide vapour (CS_2) and approximately 70% of nitrogen monoxide, (NO), contained as a mixture in the bulb.

In 1934 an entirely new principle was introduced. Now the light was produced by the combustion of a pressure during combustion; consequently, a wire-filled bulb can be filled to a higher oxygen-pressure, so that a smaller bulb volume suffices for the same light output and for the same safety-margin against explosion.

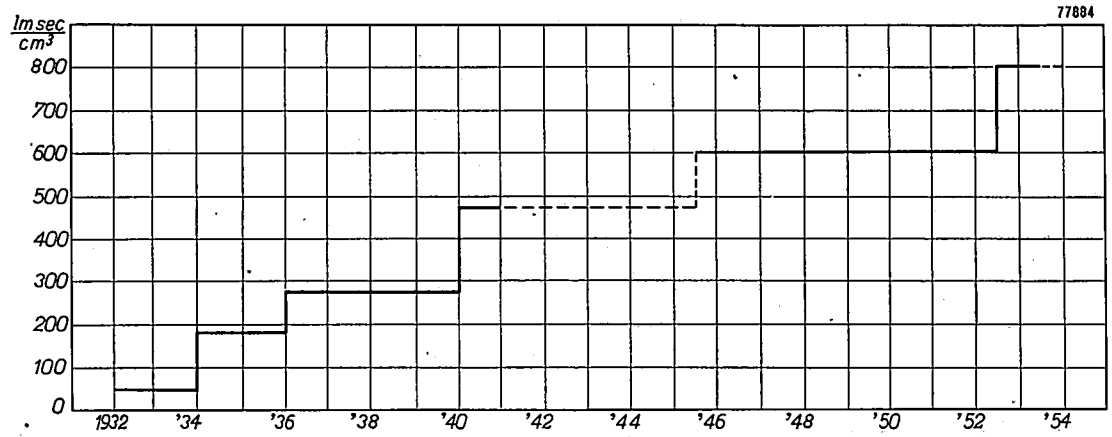


Fig. 2. The specific light output (number of visual lumen-seconds divided by the bulb volume) of "Photoflux" bulbs over the course of the years. For each year, the type with the greatest specific light output of the bulbs then available is represented.

quantity of very fine wire (only 32 microns thick) of an aluminium-magnesium alloy [2]). The bulb was filled with a loosely knit ball of this wire, together with an ample amount of oxygen. The volume was chosen so as to obtain a relatively low pressure (25 cm Hg), in order to prevent the bulb exploding upon ignition.

A different technique, developed earlier by others, and likewise producing satisfactory results, was based upon the combustion of very thin aluminium foil in oxygen. Filling the bulbs with wire instead of with foil made possible a higher degree of mechanization of the production process, so that the bulb could be made more cheaply. Also, the fact that the flash curve of a wire-filled bulb is broader than that of a foil-filled bulb (*fig. 3*), was later found to be a considerable advantage, as this permitted easier and more reliable synchronization [3]). Finally, with respect to the specific light output, the wire-filled bulb is to be preferred to the foil-filled bulb, since with a tangle of wire, the light produced in the initial stage of the combustion is absorbed to a far less degree by the yet unburnt metal than in the case of foil; thus with the same volume, more light is produced. The less steep flash curve, moreover, is associated with a lower peak-value of the

Considerations of the last-mentioned design detail are also the basis of the next step forward shown in the graph of fig. 2. It was realized that with a wire-ball, since all metal parts are easily accessible to the oxygen, it was unnecessary to have *excess* oxygen available for combustion. If *chemically equivalent* quantities of metal and oxygen are used, no superfluous gas is present to be heated to high temperatures during combustion. This means lower peak pressures, and consequently the bulb can be filled with oxygen to a higher initial pressure. Hence the required amount of oxygen can be enclosed in a

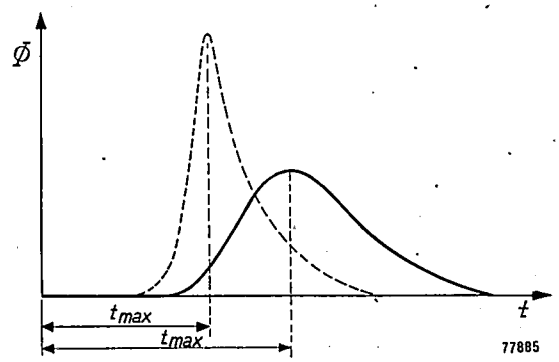


Fig. 3. Flash curves (schematic) of a bulb filled with Al-Mg wire (full curve), and of one filled with Al-foil (dotted curve). The foil-filled bulb has a higher, but narrower peak than the wire-filled bulb. Also the time to peak i.e. the period t_{max} from the moment of ignition till the reaching of the maximum luminous flux, is, as a rule, considerably shorter for the foil-filled lamp.

[2]) See J. A. M. van Liempt and J. A. de Vriend, Philips tech. Rev. 1, 289-294, 1936.

[3]) G. D. Rieck and L. H. Verbeek, Philips tech. Rev. 12, 185-192, 1951.

smaller bulb (the volume of the wire is very small in comparison to that of the oxygen and may, therefore, be neglected).

The further steps in the graph, those representing the years 1940 and 1946, are due to the introduction of a tough, protective layer of lacquer, first applied on the inside, and later also on the outside of the bulb. The reduction of the explosion risk thus obtained, permitted the use of even higher oxygen pressures. This development has already been extensively dealt with in this Review [3]).

Further research into the conditions under which explosion of a glass bulb occurs, revealed that, in order to withstand high pressures, the wall need not necessarily be very thick, provided that it is of very uniform thickness in the region of the largest diameter of the bulb, and that the latter is strictly rotationally symmetrical.

This work introduced the possibility of using bulbs made, instead of by blowing directly from the molten mass, from high-speed drawn machine-made glass tubing, which is given the required shape by locally heating and blowing in a mould. For not too large bulbs, this process can be so conducted that a very uniform wall thickness can indeed be obtained. The small bulbs for torch lamps, for example, are manufactured in this way. Although the wall thickness of these bulbs at their largest diameter amounts to barely 0.3 mm, it has been found from regular checks that on the average they can withstand static pressures of between 15 and 20 atmospheres. On the other hand, the above-mentioned investigation revealed that in the types of "Photoflux" bulbs manufactured since 1946, the peak pressures during ignition did not exceed a value of approximately 4 atmospheres. There remains, it is true (and this applies to bulbs of every shape), a certain risk that during ignition the bulb may be weakened by hot particles of ash, which may cause cracks in the glass (see the article referred to in [3])). The risk of explosion due to this cause is obviated by the tough coating of lacquer applied to the outside of the bulb. The lacquer coating itself is capable of withstanding a pressure of 3 to 5 atmospheres!

These results led to a manufacturing technique similar to that of small incandescent lamps such as torch bulbs. This method of manufacture allows the bulb volume to be used to its fullest extent as will be seen presently.

In this way, and also by using a higher filling pressure, a specific light output of about 700 lm sec/cm³ has been obtained for the PF 3 flash-bulb; for the later, slightly larger types PF 14 and PF 25, this value has been raised to 800 lm sec/cm³.

The most recent jump in the graph of fig. 2 corresponds to the introduction of this new manufacturing technique; cf. *table I*.

Table I. Data regarding the new "Photoflux" bulbs

Type	Bulb diameter (mm)	Quantity of Al-Mg wire (mg)	Light output (lm sec)	Bulb volume (cm³)	Spec. light output (lm sec/cm³)
PF 3	22	11	5500	8	700
PF 14	26.5	22	11000	14	780
PF 25	30	34	16000	20	800

The greater specific light output is not the sole reason for the lower price of the new bulbs compared with equivalent earlier types. The present method of manufacture, because it is particularly suited to large-scale mass production, was also an important factor in lowering the prices (high-speed machines, less glass scrap).

Figs 4 and *5* show the method of construction of the "Photoflux" bulbs in the old and the new manufacturing processes respectively. In the old, conventional process, construction starts from the well-known "pinch mount". This consists of a piece of glass tubing pinched at one end round the two lead-in wires, and flanged at the other end; an exhaust tube is also fused in the pinch. The ignition filament is mounted between the two lead-in wires, and is coated with a primer paste, which on heating, fires the curled wire. The bulb is filled with the

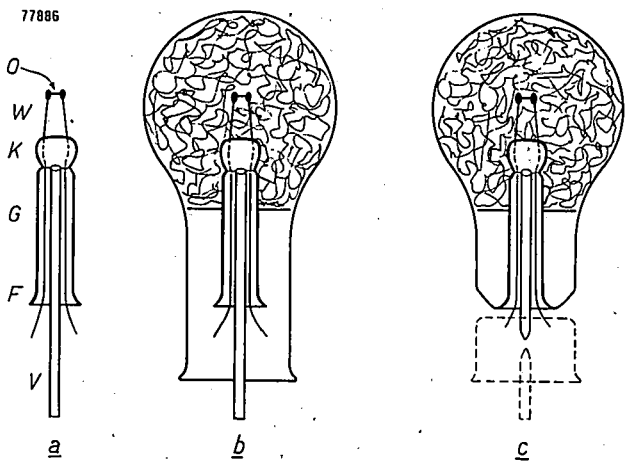


Fig. 4. Construction of a "Photoflux" bulb by the old process. The "pinch" (*a*) consists of a piece of glass tubing *G*, pinched (at *K*) around the two lead-in wires *W* carrying the igniting filament *O*, and is provided with an exhaust tube *V*. After the pinch has been placed inside the bulb (*b*), its flange *F* is fused to the neck of the bulb (*c*).

curled wire, the pinch is placed inside and the flange is then fused to the neck of the bulb. The bulb is evacuated and filled with oxygen via the exhaust tube. After sealing-in, a considerable part of the neck of the bulb and of the exhaust tube is lost as scrap glass. After pumping, a cap is cemented to the bulb.

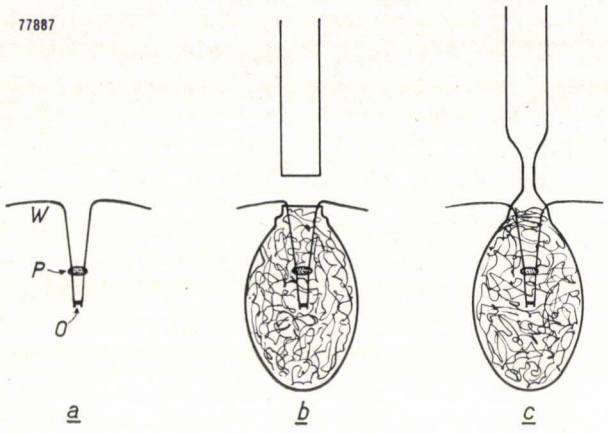


Fig. 5. Manufacture of the new "Photoflux" bulbs, with a "bead mount". The bead mount (*a*) simply consists of the two lead-in wires *W* carrying the filament *O*, fused into a glass bead *P*. This mount it suspended in the bulb (*b*), which has been blown from thin glass tubing. By fusing to this a piece of the same glass tubing (*c*), the lead-in wires are sealed-in, and the bulb is provided with an exhaust tube (which itself, after sealing off, provides the material for blowing the next bulb).

The new process (fig. 5), starts from a far simpler and smaller "mount", consisting solely of the two lead-in wires, fused into a glass bead. The ignition filament, coated with the primer paste, is mounted between the two wires as before. This "bead mount", as it is called, is suspended in the wire-filled bulb, the latter having been blown from glass tubing in the manner described above. After this, an exhaust tube of the same glass tubing is fused to the neck of the bulb. The two lead-in wires are sealed-in by the same operation. After pumping, filling with oxygen and sealing-off the tube, another bulb can be blown from the remainder of the exhaust tube, so that hardly any glass is wasted.

The PF 14 which is now being manufactured according to this technique, has a light output of approx. 11 000 lm sec, i.e. the same output as was obtained from the first "Photoflux" bulbs, marketed in 1932. The volume of the PF 14, however, is only about 1/20 of that of the old bulb. *Fig. 6* gives a graphic impression of the overall increase in specific light output since 1932 as depicted in fig. 2.

A few words should be said about the differences in the specific light output of the three bulbs of the new type (see table I). The bulb volumes are proportional to the quantities of the curled wire,

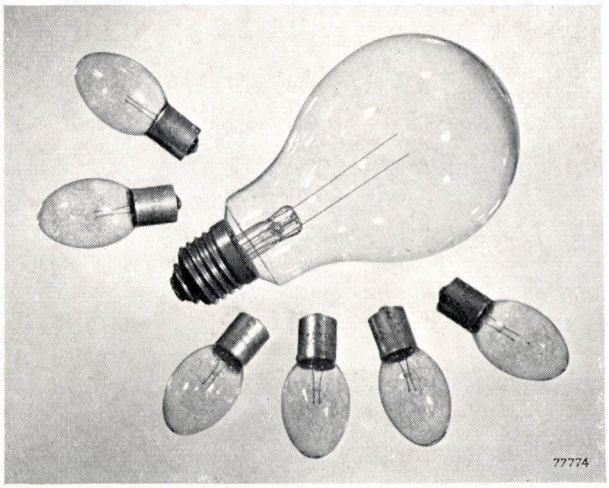


Fig. 6. A "Photoflux" bulb of 1932 with carbon-disulphide filling, compared with some flash bulbs of the modern type PF 14, which gives the same light output.

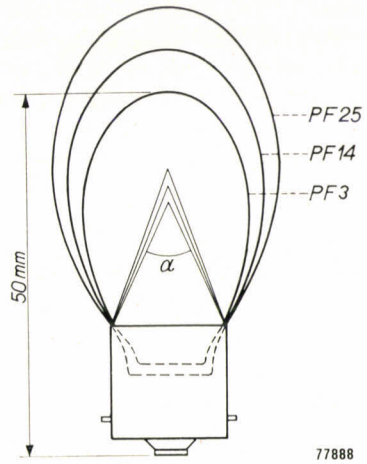


Fig. 7. The three flash bulbs PF 3, PF 14 and PF 25 have identical caps, but the bulbs are, of course, of different sizes. The angle α subtended by the cap at the centre of the bulb, is a measure of the screening effect of the cap.

since for each case equivalent quantities of oxygen are used, and the filling pressure is the same in each case. One would thus expect, provided that the light output per milligram of metal (the "efficiency") remains the same, an equal value of the specific light output for the three types. On further investigation, in fact, it has been found that the differences in the measured specific light output are *not* due to differences in efficiency, but to light losses caused by the firing paste and by the bulb cap. The three lamp types of table I have identical caps. With the smallest bulb this screens off a substantially greater portion of the light than is the case with the largest; see *fig.* 7. The firing paste, furthermore, causes a loss of light (presumably through smoke development), the percentage of which is not the same for all types; moreover, it is necessary to use firing pastes

of a different composition in order to obtain an equal ignition time (time to peak, see fig. 3) for all bulb types.

Table II shows the above-mentioned light losses for the three bulb types of table I, together with the gross specific light output obtained by adding these losses to the measured specifiic light output. From this it can be seen that the gross specific light output is practically the same for all types.

Table II. Light losses with different types of "Photoflux" bulbs

Type	Screening angle of the bulb cap	Light loss due to bulb cap %	Light loss due to ignition paste %	Gross specific light output (lm sec/cm³)
PF 3	55°	18	15	940
PF 14	45°	14	7.5	960
PF 25	40°	12.5	5	945

Summary. The light output of a flash-bulb per unit volume of the bulb is a very important factor, both with respect to the costs and to ease of handling. A graph shows how this specific light output has enormously increased since the introduction of the first "Photoflux" bulbs (1932). The article reviews the various constructional alterations in "Photoflux" bulbs which have contributed to this rise. The latest step forward has been made with the introduction of the PF 3, the smallest flash bulb in the world. By the use of a manufacturing technique similar to that used for torch bulbs, employing a bead mount, this flash-bulb has been kept very small and also low in price. The PF 14 and PF 25, at present manufactured by the same technique, have a specific light output of 800 lm sec/cm³, i.e. a value 20 times as large as that of the first "Photoflux" bulbs.

ABSTRACTS OF RECENT SCIENTIFIC PUBLICATIONS OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN

Reprints of these papers not marked with an asterisk * can be obtained free of charge upon application to the administration of the Philips Research Laboratory, Eindhoven, Netherlands.

2045: F. A. Kröger and J. A. M. Dikhoff: The function of oxygen in zinc sulphide phosphors (J. Electrochem. Soc. **99**, 144-154, 1952, No. 4).

At 1200 °C ZnS and ZnO form solid solutions: ZnS dissolves 1 mole per cent ZnO, while ZnO forms a solution with 0.3 mole per cent ZnS. The incorporation of ZnO in ZnS causes the appearance of a new absorption band and, in the presence of activators, of new fluorescence bands displaced about 150 Å toward longer wavelengths with respect to the original bands. These effects are explained by the assumption of separate oxygen levels 0.1 eV above the filled band and of new centres consisting of associated pairs of oxygen and activator ions. The formation of such pairs is explained on the basis of the principle of compensation of volume. This principle also explains the favourable effect of oxygen on the intensity of fluorescence. It is found to be the logical complement of the principle of compensation of charge, introduced earlier in connection with the incorporation of ions of a certain valency in a base lattice consisting of ions of a different valency. It is shown that oxygen gives rise to characteristic traps and has further a marked effect in combination with fluxes. The optical properties of ZnO-x(ZnS) are briefly described and explained along similar lines.

2046: J. I. de Jong and J. de Jonge: The reaction of urea with formaldehyde (Rec. Trav. chim. Pays-Bas **71**, 643-660, 1952, No. 6).

The reaction of urea and formaldehyde, giving monomethylolurea, is reversible in a neutral, acid or alkaline aqueous solution, the forward reaction being bimolecular and the reverse reaction monomolecular. The forward and reverse reactions are catalyzed to the same extent by hydrogen ions as well as by hydroxyl ions. The specific rates are directly proporional to the hydrogen ion and hydroxyl ion concentrations. There appears to be an influence of the buffer concentration on the reaction rates, probably due to a general acid and general base catalysis. The equilibrium is almost independent of the *pH* of the solution and of the buffer concentrations. The activation energies of the forward and the reverse reaction have been determined, giving the heat of reaction. Within experimental error, these activation energies may be the same for the non-catalyzed and for the catalyzed system. The reaction mechanism is discussed.

2047: J. I. de Jong and J. de Jonge: The formation and decomposition of dimethylol urea (Rec. Trav. chim. Pays-Bas **71**, 661-667, 1952).

The formation and decomposition of dimethylol urea shows a close resemblance to that of mono-

methylolurea previously studied (see Abstract No. 2046). The reaction between formaldehyde and monomethylolurea giving dimethylolurea leads to an equilibrium which can be reached from both sides. The equilibrium is independent of the pH of the solution. The formation of dimethylolurea from formaldehyde and monomethylolurea appears to be a bimolecular reaction. The decomposition of dimethylolurea is monomolecular. Both reactions are catalyzed to the same extent by hydrogen ions and by hydroxyl ions and there is an influence of the buffer concentration on the velocity. The reaction rates are found to be directly proportional to the hydrogen ion and hydroxyl-ion concentrations. The activation energies have been determined. Within experimental error, these energies are found to be about the same as the activation energies of the monomethylolurea formation and decomposition, viz. 14 and 19 kcal/mole. The reaction mechanism is briefly discussed.

2048: K. F. Niessen: Spontaneous magnetization of nickel zinc ferrite as a function of the nickel zinc ratio (Physica **18**, 449-468, 1952, No. 6-7).

The reduced spontaneous magnetization I_f^* of nickel zinc ferrite (magnetization at temperature T, divided by that at absolute zero) is determined as a function of the reduced temperature T^* (T, divided by the Curie temperature T_c) and of the nickel-zinc ratio f. The latter is defined by the following expression for the ratio between the numbers of nickel and zinc ions: $\mathrm{Ni/Zn} = f/(1-f)$. In the case $T^* \ll 1$ and in the case $T^* = 1 - \Delta$ ($\Delta \ll 1$) formulae for the reduced magnetization are given, holding for high values of f, e.g. for $^1/_2 < f < 1$. For intermediate values of T^*, a graphical method must be used. The results are compared with the experimental curves of Guillaud and Roux, which give the reduced magnetization as a function of T^* of several ferrites with various nickel-zinc ratios, especially at low and intermediate values of T^*. The curves diverge in a typical way. This can be explained qualitatively.

2049: H. Bremmer: On the asymptotic evaluation of diffraction integrals with a special view to the theory of defocusing and optical contrast (Physica **18**, 469-485, 1952, No. 6-7).

An expansion is given of the function $u(x, y, z)$ that satisfies the scalar wave equation in the half-space $z<0$ and is equal to a given distribution u_0 (x, y) in the plane $z = 0$. If u_0 is zero beyond a closed contour L in $z = 0$, $u(x, y, z)$ splits into a so-called geometrical-optical part (which vanishes outside the cylinder passing through L and having generating lines parallel to the z-axis) and a diffraction part (determined by the values of u_0 near L). Each part can be expanded into terms depending on $u_0(x,y)$ itself and on its iterative two-dimensional Laplace operators $\Delta^n u_0 = (\partial^2/\partial x^2 + \partial^2/\partial y^2)^n \; u_0$ (x, y). The expansions are in general asymptotic for small wavelengths; however, their exact validity can be proved for functions u_0 that are polynomials inside the contour L. The terms of the diffraction part consist of contour integrals along L, the corresponding terms for the geometrical-optical part do not depend on any integration. The first few terms of the latter part are essential for defocusing effects and can be connected with the brightness contrast existing in the plane $z = 0$.

2050: G. Diemer, Z. van Gelder and J. J. P. Valeton: Interference in television pictures (Wireless Engineer **29**, 164-168, 1952, No. 345).

Interference in the form of vertical lines on the left-hand side of a television picture has been studied. It can be attributed to intrinsic properties of the I_a-V_a characteristics of the power valve used for generating the line-deflection current. These properties cause irregularities in the anode current and give rise to signals of very high frequency, which may penetrate into the r.f. or i.f. amplifier of the receiver.

2051*: J. J. Went, G. W. Rathenau, E. W. Gorter and G. W. van Oosterhout: Hexagonal iron-oxide compounds as permanent-magnet materials (Phys. Rev. **86**, 424-425, 1952, No. 3).

A short survey of the comprehensive article in Philips tech. Rev. **13**, 194-208, 1952, No. 7.

2052*: J. S. C. Wessels and E. Havinga: The redox potential as a critical factor in the Hill reaction (Rec. Trav. chim. Pays-Bas **71**, 809-812; 1952, No. 7).

The photochemical reduction of various quinones and dyes by water, in the presence of chloroplasts, was studied by measurement of the redox potential during illumination. Whether this so-called Hill reaction will proceed to a measurable degree, appears to be determined by the standard potential of the redox substance. With systems having a standard redox potential of about 40 millivolts or lower, no reduction was observed.

2053: J. de Jonge, R. Dijkstra and G. L. Wiggerink: The quantum yield of the photo-decomposition of some aromatic diazonium salts (Rec. Trav. chim. Pays-Bas **71**, 846-852, 1952, No. 8).

The quantum yield of the photo-decomposition of some aromatic diazonium salts in solution is estimated, using light with a wavelength of 3560 Å. The photo-decomposition of phenyl-amino benzene diazonium sulphate is proposed as a suitable actinometer. The quantum yield of the photochemical isomerisation of the stable form of p-methoxybenzene diazo-cyanide was found to be about 0.35.

2054: J. I. de Jong and J. de Jonge: The reaction between urea and formaldehyde in concentrated solutions (Rec. Trav. chim. Pays-Bas **71**, 890-898, 1952, No. 8).

The nature and the velocities of the reactions occurring in concentrated solutions of urea and formaldehyde are very similar to those found in dilute systems. A rapid initial reaction between urea and formaldehyde, as reported in literature, was not observed, and is shown to be due to the analytical methods used.

2055: K. H. Klaassens and C. J. Schoot: The preparation of 1-methyl-2-hydroxy-3-diazobenzene-5-sulphonic acid (Rec. Trav. chim. Pays-Bas **71**, 920-924, 1952, No. 8).

The presentation of 1-methyl-2-hydroxy-3-diazobenzene-5-sulphonic acid, starting from o-cresol, is described. With a small quantity of a strong acid a suspension of 1-methyl-2-hydroxy-3-aminobenzene-5-sulphonic acid in a solution of sodium nitrite gives the sodium salt of the diazonium compound.

2056: J. G. Bos, R. J. H. Alink and C. J. Dippel: Photodecomposition of aqueous solutions of diazonium salts in the presence of mercurous ions. (Rec. Trav. chim. Pays-Bas, **71**, 945-953, 1952, No. 8).

Photodecomposition of aqueous solutions of o-hydroxydiazonium salts leads to the formation of mercury with a disproportion of mercurous ions, in contrast to p-aminobenzene diazonium salts, where mercury only results from reduction of mercurous ions by the photodecomposition products.

2057: H. Bremmer: The derivation of paraxial constants of electron lenses from an integral equation (Appl. sci. Res. **18**, 416-428, 1952, No. 6).

The paraxial trajectories in electron lenses are derived from an integral equation. The Liouville-Neumann expansions of the solutions of this equation lead to expressions for the magnification, the focal distances and the positions of the focal points and cardinal points. The number of integrations to be performed in the individual terms of the expansions is halved, as compared with the normal treatment. The focal and cardinal points are defined as osculating elements similar to those introduced by Glaser.

2058: A. H. Boerdijk: A new aspect in the calculation of toothed gearing (Ingenieur **64**, W63-W65, 1952, No. 37).

This paper deals with the calculation of the numbers of teeth of the wheels in a gearing with a prescribed gear ratio, equal to the quotient of two integers k/l. The known methods of decomposing k and l into prime factors and of developing k/l in continued fractions often fail to give a practical and exact solution. A new method is described, based on the introduction of planet gearings (or simple differentials) in the conventional gearing. Practical and exact solutions can be obtained for arbitrary values of k and l, even for very large primes.

2059: G. W. Rathenau, J. Smit and A. L. Stuyts: Ferromagnetic properties of hexagonal iron-oxide compounds with and without a preferred orientation (Z. Phys. **133**, 250-260, 1952, No. 1-2).

The article deals with the magnetisation of hexagonal iron-oxide compounds as a function of the magnetic field strength. It is shown that crystal imperfections must play a role in the formation of Bloch-Walls. In the case of random orientation of the crystallites, and with crystallites of dimensions greater than a certain critical value, demagnetization by displacement of Bloch-Walls already occurs at positive field-strengths $H \approx 4\pi I_s$ ($\mu_0 H \approx J_s$). By preparing specimens with oriented crystals, BH-values as large as 3×10^6 gauss. oersted ($BH \approx 2.5\ 10^4$ J/m^3) are obtained. Crystal growth causes an increase in the preferred orientation. An explanation of this effect is given. (See also Philips tech. Rev. **14**, 194-208, 1952, No. 7).

2060: J. Smit: The influence of elastic shear strains on the conductivity and thermo-electric force of cubic metals (Physica **18**, 587-596, 1952, No. 8/9).

Elastic shear strains cause a change in the shape of the Fermi surface of metals. The influence that such a change has upon the electrical conductivity

and the thermo e.m.f. has been calculated for monovalent face-centered cubic metals, and the results have been compared with the experimental values for Cu, Ag and Au. The theory also accounts qualitatively for the experimental values of trivalent Al.

2061: P. M. van Alphen, C. G. E. Burger, W. J. Oosterkamp, M. C. Teves and T. Tol: Detailerkennbarkeit bei Durchleuchtung und Photographie mit der Bildverstärkerröhre (Fortschr. Röntgenstrahlen-Röntgenpraxis **77**, 469-470, 1952, No. 4). (Performance of an X-ray image intensifier with visual and photographic observation; in German).

A "Bakelite" thorax phantom was used to test the performance of an X-ray image intensifier. The visual acuity during fluoroscopy in a room with fair illumination after 1 minute of adaptation and with only 1/5 of the dose rate is almost twice as good as during ordinary fluoroscopy. Photographs of the intensified image can be made with 1/10 of the energy used for ordinary contact photographs and 1/50 of the energy used for ordinary camera photographs. These photographs show a resolution of detail which lies midway between contact photography and ordinary fluoroscopy. The results are in good agreement with calculated data based upon the photon-fluctuation theory of Morgan and Sturm.

2062*: G. W. Rathenau: Grain growth observed by electron-optical means (L'état solide, Rapp. 9 Cons. Phys., Inst. int. Phys. Solvay, 25-29 Sept. 1951, Stoops, Brussels 1952, p. 55-72).

Observation of grain growth in cold rolled homogeneous Ni-Fe alloys on annealing (see abstracts **1986** and **2026**). The grain growth which accompanies the α-γ phase transformation of a Si-Fe alloy has also been studied by electron-optical means. Since, in equilibrium, the Si-content of the α phase surpasses that of the γ phase, the velocity of phase boundary movement depends essentially on the concentration gradient which determines the diffusion at the moving boundary. In agreement with expectation, the advancing boundary of a γ grain curve is often heavily convex towards the disappearing α grains. γ grains repel each other because of the high Si content of the α phase in between. The grains of the growing phase are thus enveloped by material having the structure and orientation of the disappearing phase. On lowering the temperature these α inclusions lead to roughly the initial α structure, while on quickly heating they give rise to nuclei of δ phase.

2063: K. H. Klaassens and C. J. Schoot: Preparation of 2-ethoxy-4-diethylaminobenzene-1-diazonium borofluoride (Rec. Trav. chim. Pays-Bas **71**, 1086-1088, 1952, No. 9/10).

Description of the preparation of the compound named in the title. A stable diacetyl derivative is formed by treating 2-ethoxy-4-diethylamino-1-aminobenzene with acetic anhydride.

2064: H. B. G. Casimir: Algemene beschouwingen over ruis (T. Ned. Radiogenootschap **17**, 199-206, 1952, No. 5/6). (General considerations on noise; in Dutch).

General considerations on noise (thermal and corpuscular) as an introduction to a symposium on this subject.

2065: F. L. Stumpers: De signaal-ruisverhouding bij verschillende modulatiesystemen (T. Ned. Radiogenootschap **17**, 249-260, 1952, No. 5-6). (The signal-to-noise ratio for different modulation systems; in Dutch).

A survey is given of the signal-to-noise ratio in the output of different modulation systems as a function of the same ratio in the transmission channel and of the necessary bandwidth. Pulse and carrier wave systems are considered. Attention is drawn to the noise-improvement threshold. It is shown that coded systems, like PCM and delta-modulation, are nearly immune to the noise in the channel, but that in their case the approximative nature of the pulse waveforms themselves introduces quantization noise.

2066: J. de Jonge: Evaporation of mercury from a solution of a mercurous salt at room temperature (Rec. Trav. chim. Pays-Bas **71**, 1209-1212, 1952, No. 11).

It is shown that a solution of mercurous nitrate gives off mercury vapour at room temperature, due to the occurrence of the equilibrium

$$Hg^{2+} \rightleftarrows Hg^{2+} + Hg.$$

VOL. 15 No. 12, pp. 325 364 JUNE 1954

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
THE PHILIPS INDUSTRIES

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, NETHERLANDS

A TIME-OF-FLIGHT NEUTRON SPECTROMETER

by P. E. TRIER *), J. C. HAMMERTON **) and E. WOLFENDALE *). 539.185.7

The study of neutron absorption by different elements is of great importance in nuclear physics as it yields valuable evidence on the composition and binding forces of atomic nuclei, and on the suitability of materials for atomic piles. A widely used method of measuring neutron absorption as a function of neutron energy is the time-of-flight spectrometer. An instrument of this type constructed for A.E.R.E., Harwell and used in conjunction with the 15 MeV linear accelerator described earlier in this Review, is the subject of this article.

Introduction

Neutron reactions

Neutrons are among the most penetrating particles known to modern nuclear physics. A description of the effect of a sample of material upon a beam of neutrons has been given in a paper by Taylor and Havens [1]) but a brief summary will be useful here.

A beam of neutrons passing through a sample of material suffers an attenuation due to interaction between some of the neutrons and the nuclei of the material. This interaction may take several forms, e.g. absorption and elastic or inelastic scattering of neutrons. Many elements exhibit regions of "resonance absorption" as shown by a marked increase of the attenuation of the beam at definite values of neutron energy.

For a thin sample the decrease in intensity of the beam is given by the formula:

$$I = I_0 \, e^{-\sigma_t n l},$$

where I_0 is the initial intensity of the neutron beam, I is the intensity after passing through the sample, n is the number of nuclei per cm^3 of sample, l is the thickness of the sample, σ_t is the total cross-section in cm^2, which (by this formula) is defined as the effective area presented by the nuclei to the neutron beam.

The total cross-section σ_t depends upon the velocity of the neutrons, and "resonance absorption" is mathematically expressed as a maximum occurring in σ_t at definite values of neutron energy. A careful study of these energy levels for different elements will yield valuable experimental evidence about the composition and binding forces of atomic nuclei.

Some information has been obtained about neutron absorption by studying the radioactivity caused by direct activation of samples of elements by neutrons. These experiments are neither simple to perform nor capable of any great accuracy. Quantitative evaluation of resonance cross-section necessitates the use of some form of velocity spectrometer.

Time-of-flight spectrometers

The most flexible and most widely used method of measuring neutron reaction cross sections of elements in the energy region 1 eV to 20 keV makes use of a time-of-flight spectrometer. A neutron source is periodically made to emit neutrons for a short time. The neutrons travel over a distance L and the neutron pulse, which has a continuous energy spectrum, becomes spread out along its direction of travel as the slower neutrons lag behind the faster ones. After a time T from the release of

*) Mullard Research Laboratories Salfords, England.
**) Atomic Energy of Canada Ltd, formerly at the Mullard Research Laboratories.

1) T. J. Taylor and W. W. Havens Jr., Nucleonics 5, 4. 1949.

the pulse, a particle detector in the path of the beam is rendered sensitive for a short interval ΔT and the number of neutrons arriving in this interval is recorded by electronic scaling circuits and mechanical counters. The recorded count gives a proportional indication of the number of neutrons with velocity L/T. By continuing the counting over a large number of neutron pulses and by varying T over a range, a graph of the energy distribution can be obtained.

Most of the early adaptations of this method were limited to inspection of an interval ΔT at one particular delay T per experiment, but modern time-of-flight spectrometers obtain more than one reading at a time by arranging to switch the output of the detector over a range of values for T by using a multiple bank of electronic "gating" units.

The actual measurement of total cross section is effected in the following way. The energy distribution of the neutrons in the incident beam is ascertained; a sample of the element under investigation is then interposed between source and detector. The total cross section of the element in any given energy interval is then determined from the transmission factor I/I_0 obtained from the modification of the energy spectrum of the detected neutrons.

The energy E_x of the neutrons is determined with an accuracy ΔE. The ratio $\Delta E/E$ is called the resolution of the spectrometer. This resolution is dependent upon the accuracy of determination of both L and T. The length L of the flight path is subject to uncertainty because of the finite depth of both source and detector. The time of flight T is subject to some indeterminacy since the neutron source is pulsed for a time which may not be small compared with T when dealing with high energy neutrons; furthermore the inspection period ΔT is also finite.

The resolution can be improved by reducing the depth of the detection chamber, reducing the duration of the neutron pulse or the inspection period ΔT, or by increasing the path length L. All these improvements result in a reduction of the counting rate, and since a reasonable counting rate is required, the principal factor limiting the resolution of a spectrometer is the available neutron flux at the source.

The statistical accuracy of the spectrometer is not affected by the average detector counting rate, provided that this is large compared with any random background counting rate, but it depends on the total number of counts and varies as the reciprocal of the square root of this number. Hence the need for a reasonable counting rate if experiments are not to be unduly protracted.

Two examples of early time-of-flight spectrometers, using a single inspection period per experiment are those of Luis Alvarez [2]) and Baker and Bacher [3]). In both cases the resolution was limited by the duration of the neutron pulse, that of Baker and Bacher being the best with a neutron pulse of from 50 to 500 μsec long.

The first time-of-flight spectrometer employing more than one inspection period per experiment was that of Haworth, Manley and Luebke [4]). This spectrometer aimed at a higher order of precision but the strength of the source limited the study to neutrons with energies between 0.004 and 1 eV.

This was followed by a spectrometer built at Columbia University by Rainwater and Havens [5]), using an 8 MeV cyclotron, and by a spectrometer built by the Mullard Research Laboratories and the Atomic Energy Establishment, Harwell [6])[7])[8]). It is interesting to compare the salient points of these spectrometers.

The Columbia University spectrometer had a minimum neutron pulse length of 5μsec, and a minimum inspection period of 5 μsec. The number of intervals capable of being investigated simultaneously was 16 and the neutron path length was 6.2 metres. A moderating paraffin block had to be used to slow down the higher energy neutrons to useful levels; this introduced a further uncertainty in the time at which a neutron left the source. The resolution of the Columbia University Spectrometer has since been improved by reducing the neutron pulse length to 4 μsec and the inspection period to 2 μsec. This improvement and some typical results are described in a paper by Havens and Rainwater [9]).

The Harwell time of flight spectrometer was capable of giving improved resolution at low and high energies owing to the use of a travelling wave linear accelerator with an electron pulse of 120 mA at an energy of 3.2 MeV. This ensures a sufficiently high

---

[2]) Luis W. Alvarez, Phys. Rev. **54**, 609, 1938.
[3]) C. P. Baker and R. F. Bacher, Phys. Rev. **59**, 332, 1941.
[4]) J. H. Manley, L. J. Haworth and E. A. Luebke, Rev. Sci. Instr. **12**, 587, 1941; Phys. Rev. **69**, 405, 1946.
[5]) J. Rainwater and W. W. Havens Jr., Phys. Rev. **70**, 136, 1946; J. Rainwater, W. W. Havens, C. S. Wu and J. R. Dunning, Phys. Rev. **71**, 65, 1947.
[6]) A. W. Merrison and E. R. Wiblin, Nature, **167**, 346, 1951.
[7]) A. W. Merrison and E. R. Wiblin. Proc. Roy. Soc. A **215**, 278, 1952.
[8]) F. S. Goulding, J. C. Hammerton, M. G. Kelliher, A. W. Merrison and E. R. Wiblin, Proc. Inst. El. Engrs. 1954 II, Paper M 1528 (appearing shortly).
[9]) W. W. Havens Jr. and L. J. Rainwater, Phys. Rev. **83**, 1123, 1951.

flux of low energy neutrons (due to the moderating action of the source itself) to make the use of a paraffin moderator unnecessary. The increased flux enables the path length to be increased to 10 metres; and this, combined with an overall indeterminacy in the time of flight of only 6 μsec gives the improved resolution. The minimum neutron pulse and the minimum inspection period were both 2 μsec. The number of intervals capable of being investigated simultaneously was 32.

A further improvement in resolution is now being obtained at Harwell by the use of the 15 MeV linear accelerator [10]) together with a new timing system. The 15 MeV accelerator provides a further increase of flux, allowing the path length to be increased to 20 metres. The timing system, which was designed and built simultaneously with the accelerator by the Mullard Research Laboratories in conjunction with the Atomic Energy Research Establisement, is similar to that of the previous spectrometer except that the number of intervals simultaneously available for examination is increased to 100. This Harwell spectrometer will now be described in detail.

Fig. 1 gives a general view of the installation during construction. At the left the concrete shelter, containing the linear accelerator, is visible. In the centre one of the flight-paths can be seen with the detector at the extreme right. In the background is the electronic timing system which is the principal subject of this article.

Design of the Harwell Spectrometer

General description

A block diagram of the principal units of the spectrometer is shown in *fig. 2*. The heart of the installation is the 15 MeV linear accelerator previously described in these pages [10]). The 2 μsec electron pulse is used to irradiate a γ-ray and neutron source.

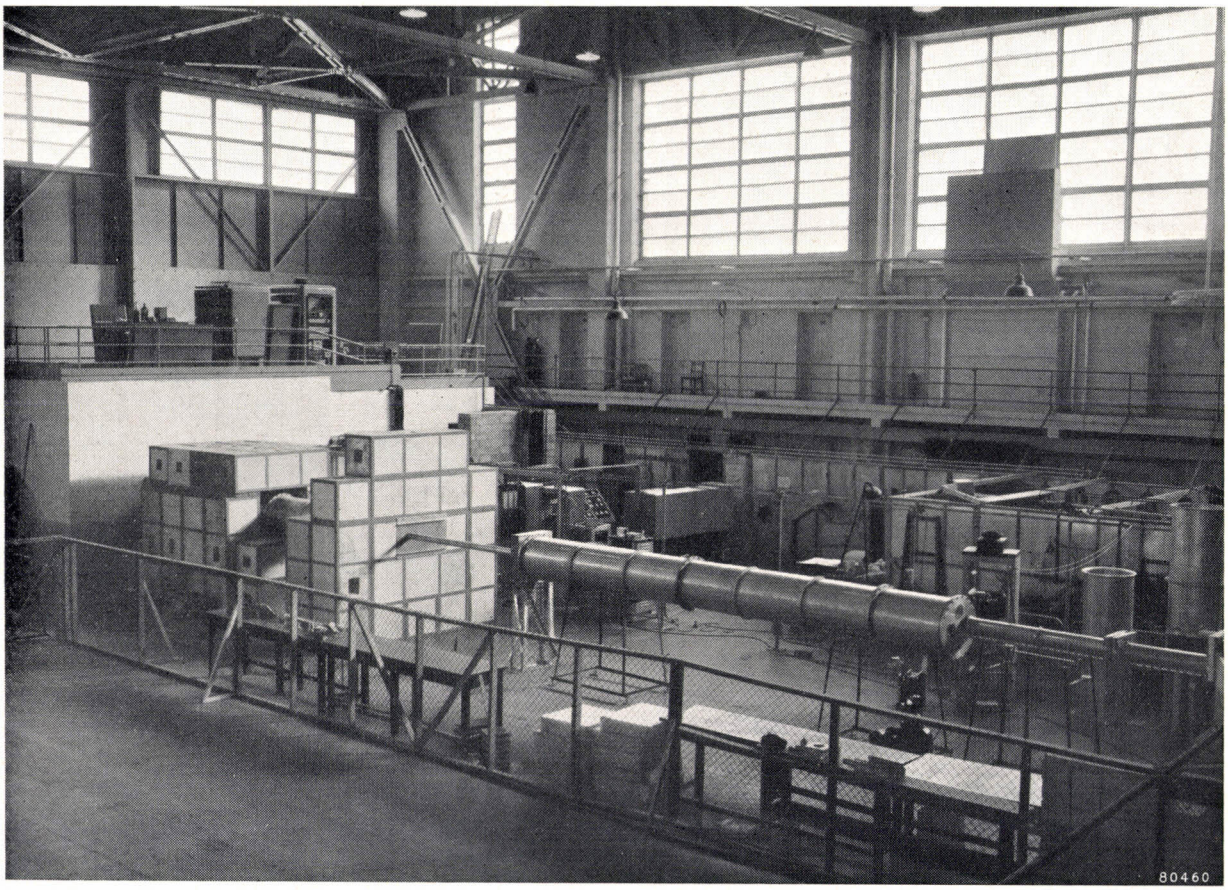


Fig. 1. General view of the neutron spectrometer at Harwell during construction.

The accelerator is housed in a heavy concrete shelter. Windows in the shelter near the neutron source provide exits for the neutrons into two roughly evacuated flight paths. Each flight path can be used separately for independent experiments.

[10]) C. F. Bareford and M. G. Kelliher, Philips Techn. Rev. **15**, 1-26, 1953 (No. 1).

A timing pulse, designated the modulator pulse, is generated in a *Master Timing Unit* and is used to initiate the high voltage pulse which is applied to the electron gun of the linear accelerator neutron detector. By now, the neutrons have become spaced out in time according to their energies. The electrical signal pulses from the detector therefore constitute a series whose distribution in time

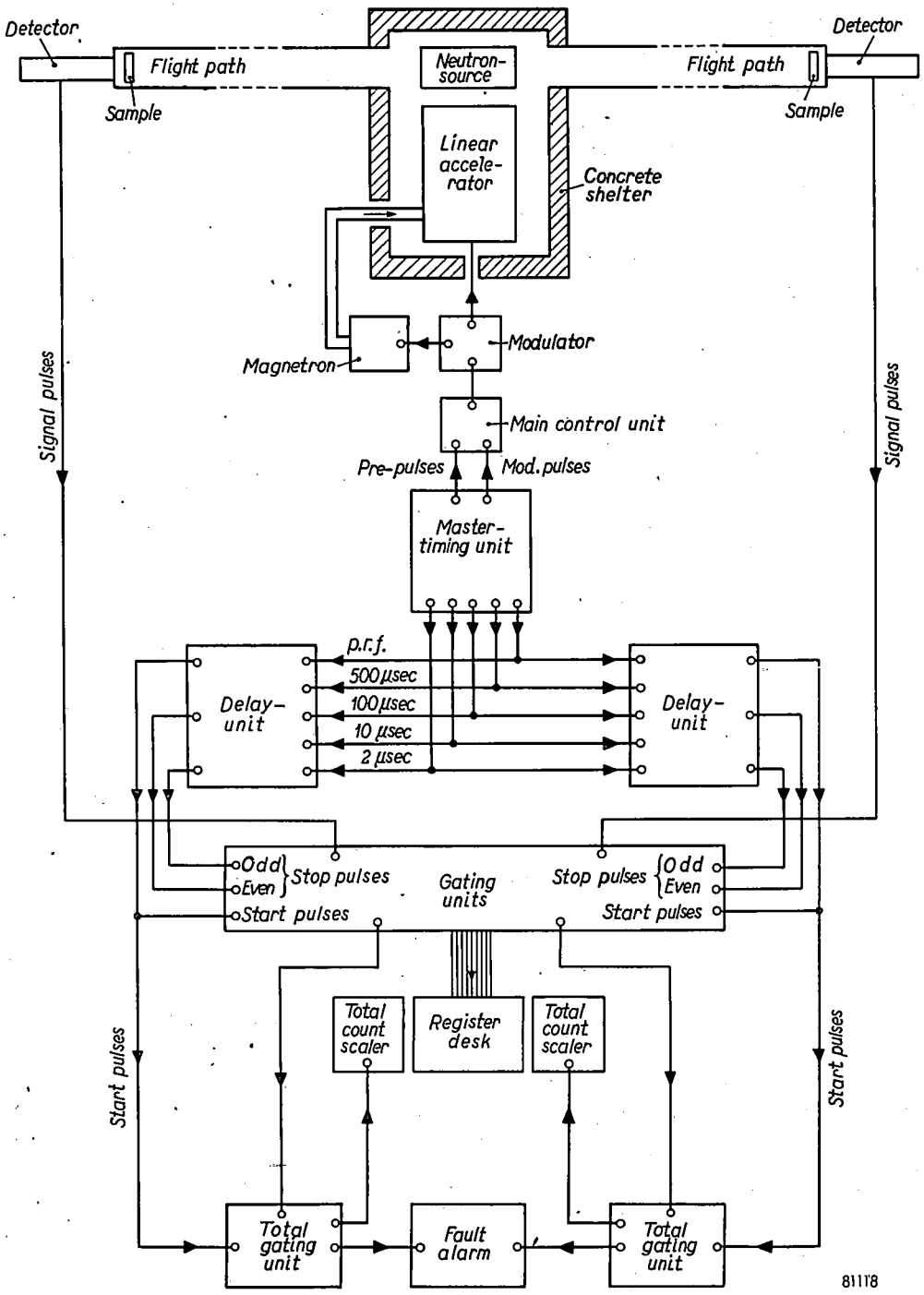


Fig. 2. Block diagram showing the principal units of the Harwell time-of-flight neutron spectrometer.

and to its magnetron. The neutron pulse of 2 μsec duration is emitted from the neutron source 1-3 μsec after the modulator pulse because of circuit delays.

The neutrons travel down each 20 metre flight path and are detected at the end of a path in a corresponds to the energy spectrum. This pulse series is fed via amplifiers to the common input rail of a bank of 100 gating units. The gating units are opened successively for equal short intervals ΔT, starting at a pre-set time T after the initiation of the

neutron pulse. Each gating unit output is fed separately via a pulse lengthener to a mechanical counter acting as a register. These counters are conveniently grouped together in a register desk. The gating units thus perform a sorting action, in such a way that the nth register will count only those pulses arriving at a time t from the initiation of the neutron pulse, such that

$$T + (n - 1)\,\Delta T < t < T + n\,\Delta T.$$

It should be pointed out here that for the release of one 2 μsec neutron pulse, each gating unit can register only one neutron arrival. The parameters of the system are however such that the probability of arrival of one neutron (let alone two or more neutrons) at any given gating unit interval as a result of a single neutron pulse release is exceedingly small. The statistical distribution is evaluated at the registers by a repetitive cycle at 200 or 400 neutron pulses per second. Even so, a run of some hours is usually required to build up an adequate count distribution.

The practical features of the Harwell instrument allow considerable flexibility of operation. The 100 gating units can be used either in a single sequence, or else in two independent banks of 50 units each for experiments on separate flight paths. A subdivision into 16 + 84 units is also possible.

The initial delay T before the initiation of the gating unit sequence is provided by a *Delay Unit* actuated by the master timing unit. The delay unit is capable of giving a delay from 0 to 2990 μsecs, adjustable in steps of 10 μsecs. The delay unit also incorporates circuits which determine the gating interval ΔT, which can be set at 2 μsecs or 10 μsecs per gate.

In order to allow the use of two independent gating unit chains, two identical delay units are provided, each having the features described above. If the whole chain of 100 gates is used as a single sequence, only one delay unit operates.

To provide a check on the correct operation of the gating units, a *Total Gating Unit* [11]) is provided for each chain, which gates the complete pulse series applied to the chain. These pulses are then counted by a *Total Count Scaler* [11]), an electronic high speed counter, which displays the count with neon indicators for units and tens and a mechanical counter for hundreds etc. The count on the total count scaler should evidently correspond to the sum of the individual gating unit counts.

Associated with the total gating units is a *Fault Alarm System*. This stops the counting and sounds an alarm whenever a fault occurs in the gating unit chain. This is an essential feature, since a normal experiment takes several hours of usually unattended operation, and an audible fault alarm therefore avoids waste of experimental time.

To avoid frequent valve failures, the entire valve complement is normally replaced at periodic intervals. All circuits are designed to be independent of individual valve characteristics within the widest possible limits, so that readjustments are not necessary after valve replacements. These features are important since the equipment contains upwards of 750 valves.

The master timing unit

The master timing unit provides the series of timing pulses on which the operation of the spectrometer is based. A block diagram of the circuit is shown in *fig. 3*. For ease of reference, the different pulse series are detailed below.

1) A pulse series with a period of 2 μsec.
2) A pulse series with a period of 10 μsec.
3) A pulse series with a period of 100 μsec.
4) A pulse series with a period of 500 μsec.
5) The *principal recurrence frequency* (p.r.f.) pulse series, which governs the neutron pulse frequency, i.e. the fundamental recurrence period of the whole system. This period can be adjusted [12]) to 5000, 2500 or 1000 μsec.
6) A pulse series, having the same frequency as the p.r.f. pulse series, which is used to initiate the time base of the monitor on the linear accelerator and is called the *pre-pulse series*.
7) The *modulator pulse series*, already mentioned above. This series has also the same frequency as the p.r.f. series but has a variable delay of from 2 to 5 μsec on the pre-pulse series.

The minimum timing interval required in the system is 2 μsec and hence the timing system is based on a primary pulse series of 2 μsec period, synchronized by a 500 kc/s crystal controlled oscillator. The other pulse series are derived from the 2 μsec series by four frequency dividers. The frequency division ratios of these dividers are 5 : 1 (for the 10 μsec series), 10 : 1 (100 μsec), 10 : 1 (1000 μsec) and 5 : 1 (5000 μsec).

The 500 μsec and 2500 μsec series are derived from the combined pulse edges of both anodes of the last two dividers, as will be described presently.

The dividers are conventional two valve multi-

[11]) Of AERE design.

[12]) The 1000 μsec period is provided for future developments.

vibrators, synchronized by downgoing edges applied to both cathodes. For reasons which will be explained below (see *Delay units*, resetting of delay *III*) divider *I* divides by 5 unsymmetrically, one valve conducting for 4 μsec and the other for 6 μsec. This ensures that the 10 μsec pulse series is delayed by 6 μsec on the 100 μs pulse series. Dividers *II* and *III* divide by 10 symmetrically. Divider *IV* divides by 5 symmetrically. The latter is achieved by applying two sets of synchronizing edges, derived from the two anodes of divider *III*, separately to the two cathodes of divider *IV*, so that each valve conducts for two and one half periods, i.e. 2500 μsec.)

The dividers deliver square wave voltages with periods of 10, 100. 1000 and 5000 μsec. The 10 μsec and 100 μsec pulse series drivers are synchronized by the downgoing edges of the first two of these voltages. The 500 μsec pulse series driver is synchronized with the combined edges from the two anodes of divider *III*. This combination takes place in the *1000 μsec mixer*.

As already mentioned above, the period of the p.r.f. pulse series can be selected at 5000, 2500 or 1000 μsec. This pulse series is derived from the *p.r.f. selector*, which selects the edges from the appropriate dividers.

The pre-pulse series is derived directly from the primary pulse series by the p.r.f. selector which gates the next primary pulse occurring after the p.r.f. pulse.

The modulator pulse series is obtained from the pre-pulse series via the modulator pulse delay circuit, which provides a continuously variable delay of from 2 to 5 μsec.

All pulses are generated by output circuits which consist of blocking oscillators. The duration of the pulses is about 0.2 μsec, and their rise time is less than 0.05 μsec when the blocking oscillators work into a load consisting of a cable with a characteristic impedance of 100 ohms. As is shown in fig. 2, the first five of the pulse series mentioned are applied to the delay units, the latter two being applied to the main control unit for the linear accelerator.

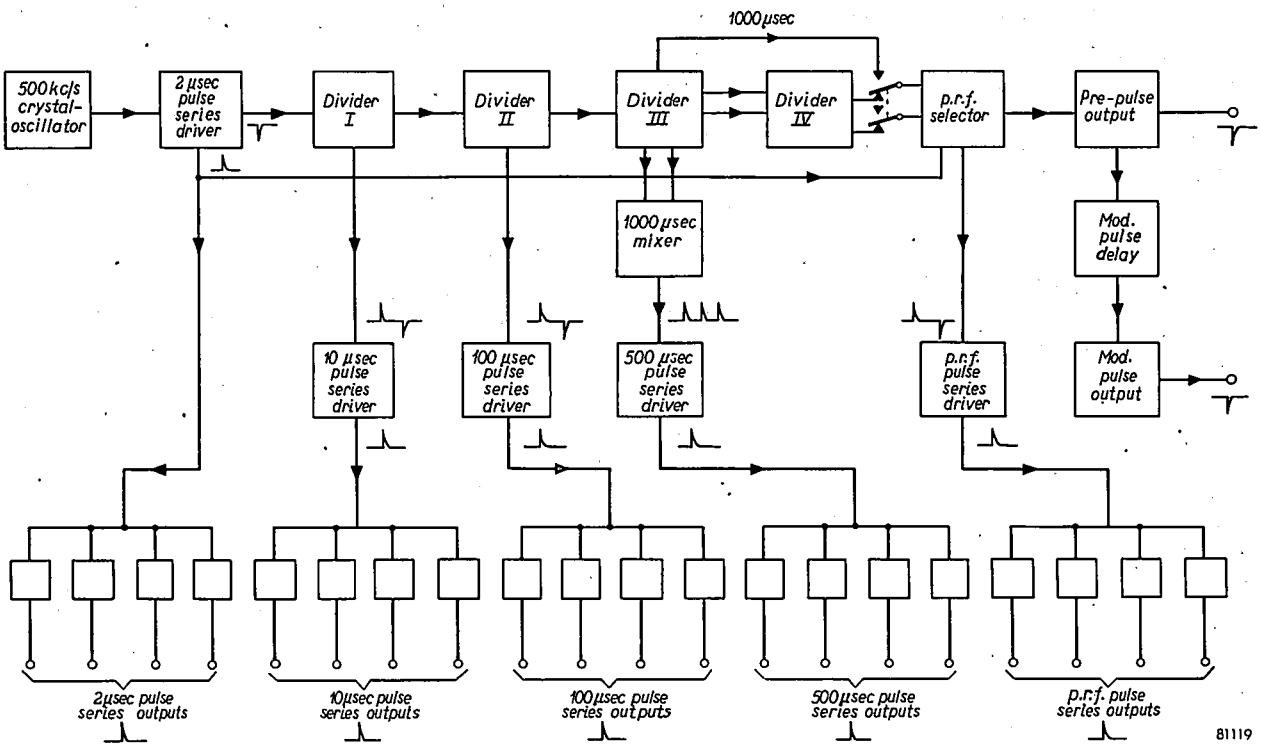


Fig. 3. Block diagram of the Master Timing Unit.

To allow for future development of the spectrometer by the addition of further flight paths, delay units and gating units, those of the pulse outputs which are connected to the delay units, are quadruplicated. Each set of four output stages is driven from a common driver.

The delay units

The delay units provide the time delay T between the initiation of the neutron pulse and the opening of the first gate in the appropriate chain of gates.

This delay is variable in 10 μsec steps from 0 to 2990 μsec.

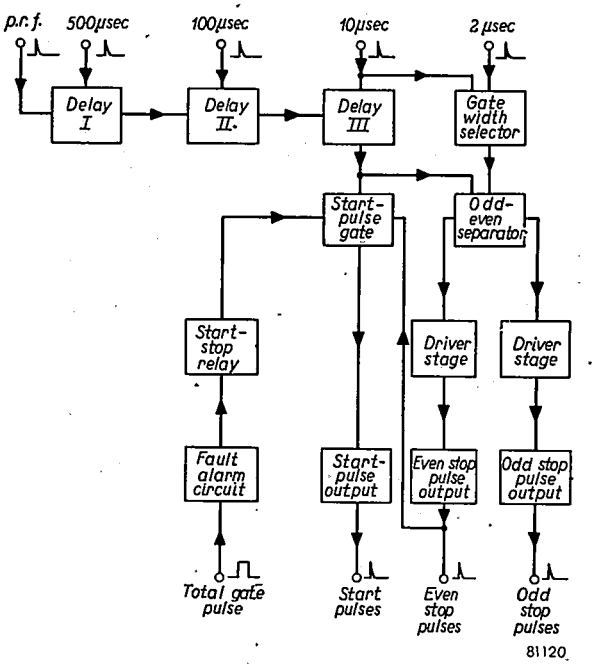


Fig. 4. Block diagram of a delay unit.

The delay units also apply three series of pulses to the gating unit chain. The first of these series is called the *start pulse series,* which initiates the gating unit chain; the others are the *odd* and *even stop pulse series* which establish the gating interval ΔT, and ensure that this interval is independent of the timing circuits of the individual gating units.

A block diagram of a delay unit is shown in *fig. 4.* The time delay T is provided by three delay circuits in cascade. Each of these circuits consist of a multivibrator which has a stable and a semi-stable state. The application of an initiating pulse triggers the multivibrator into the semi-stable state, and the duration of this state is controlled by a series of resetting pulses, which is continously fed to each multivibrator.

One example of a delay circuit is shown in *fig. 5* (delay *II*). When the circuit is in its stable state, B_1 is conducting and B_2 is non-conducting. The initiating pulses, which are provided by delay *I*, are applied to the cathode of B_2 (via terminal p). By these pulses the circuit is triggered into its semi-stable state with B_2 conducting and B_1 non-conducting. The resetting pulses are continuously applied to the cathode of B_1 (via terminal r). For this delay circuit the resetting pulses consist of the 100 μsec pulse series and hence the duration of the quasi-stable state is an exact multiple of 100 μsec. This multiple is determined by the value of resistance associated with the switch S.

Delay *II* applies downgoing initiating edges to delay *III* (via terminal q). As shown in fig. 5, at the "0" position of switch S, these edges are taken from the anode of B_2 and therefore in this position they coincide with the initiating pulses from delay *I*. In this position the circuit resets itself into the stable state prior to the first resetting pulse after the initiating pulse. At the other positions of S the initiating edges for delay *III* are taken from the screen grid of B_1 and so they are coincident with the moments of resetting. The delay time of delay II is therefore 0-900 μsec variable in 100 μsec steps.

All three delays operate in a similar manner. The initiating pulses and the resetting pulses for delay *I* consist respectively of the p.r.f. pulses and the 500 μsec pulse series. This circuit provides a delay of

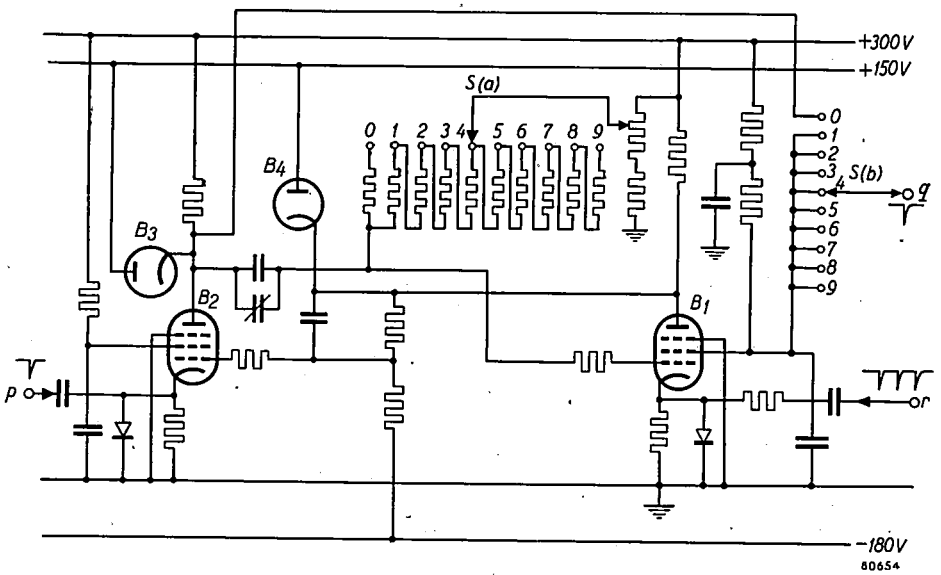


Fig. 5. Diagram of one of the delay circuits (delay *II*). It consists of a multivibrator with a stable and a semistable state.

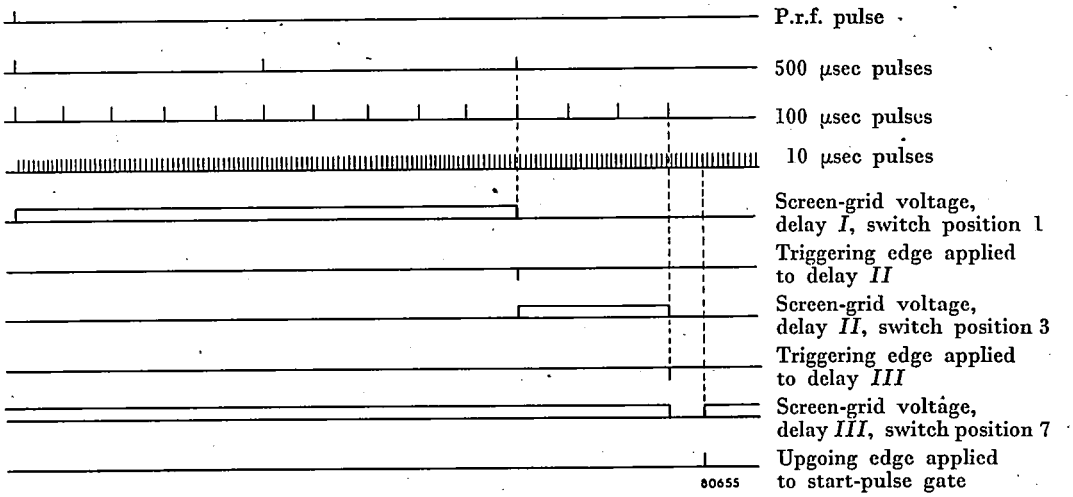


Fig. 6. Illustration of the effect of all three delay circuits in series. Delay I is set in switch position 1, delay II in position 3 and delay III in position 7. The total delay is therefore 1370 + 6 μsec minus the delay in the divider chain.

0-2000 μsec variable in two steps of 1000 μsec. The resetting pulses of delay *III* consist of the 10 μsec pulse series. As already mentioned this pulse series is delayed on the 100 μsec pulse series by 6 μsec, and as delay *III* is reset by resetting pulses in all switch positions, including the "0" position, the delay time of delay III is 6-96 μsec in 10 μsec steps.

Fig. 6 shows the effect of all the delay circuits in series. As the extra delay of 6 μsec mentioned above is not indicated on the switch positions of delay *III*, the edge which is available from delay III is delayed on the p.r.f. pulse by the sum of the three individual delay settings plus 6 μsec and minus the overall delay in the divider chain. This latter delay is approximately 1·5 μsec. It is useful as it prevents coincident triggering of the delay circuits, as all the initiating pulses occur a fraction of a microsecond after the resetting pulses. Coincident triggering can be a serious fault as it tends to render the circuit inoperative.

The way in which the start pulses and the stop pulse series are produced is illustrated in the *figures 7a and 7b* which relate respectively to gating intervals of 2 and 10 μsec. To prevent coincident triggering of the gating units it is necessary to produce two series of stop pulses, designated the odd and even stop pulse series, feeding the odd and even gating units respectively. This is accomplished by the odd-even separator consisting of a direct-coupled multivibrator, which in the case of 2 μsec

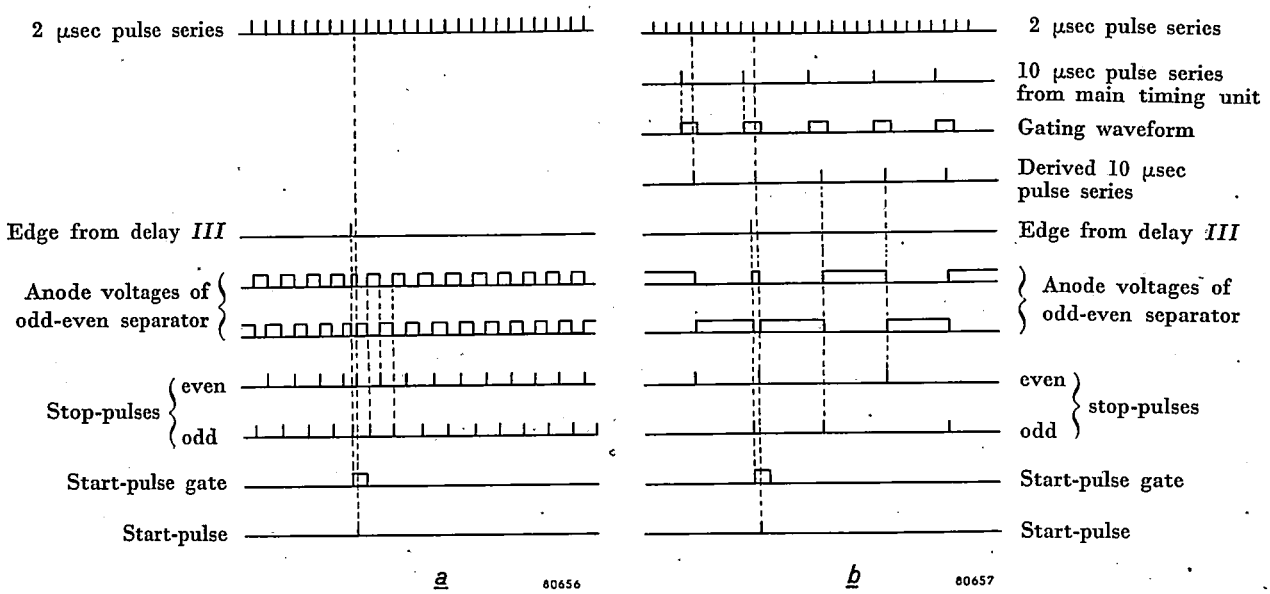


Fig. 7. Illustration of the way in which the start and stop pulses are produced *a*) in the case of a gate width of 2 μ sec, *b*) in the case of a gate width of 10 μ sec.

gate width is driven by the 2 μsec pulse series from the main timing unit. The edges in the respective anodes of the odd-even separator are used to trigger the odd and even stop pulse driver blocking oscillators.

phasing of the 10 μsec pulse series with respect to the pre-pulses, and 2 μsec by the gating of the even stop pulses. Hence, with all delays set at zero the start pulses and the neutron pulses can be made coincident in time by aligning the leading edges of

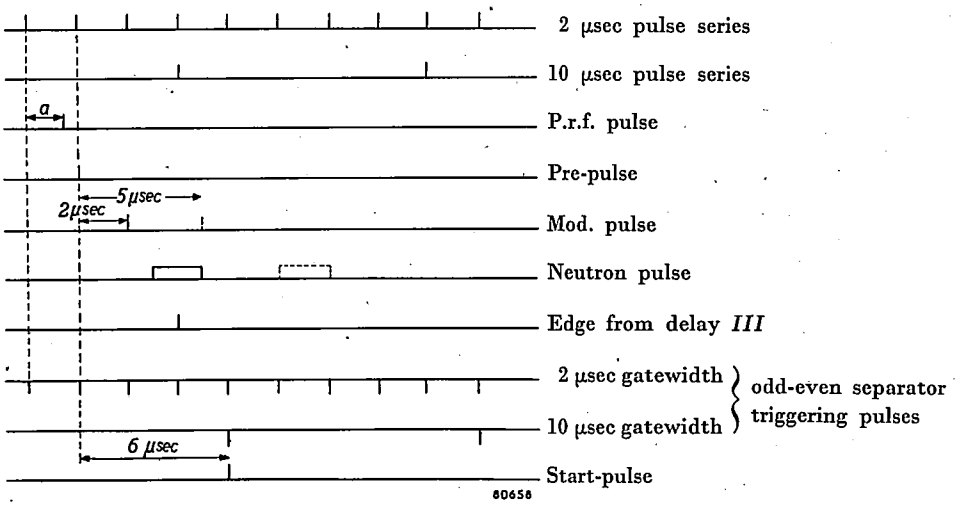


Fig. 8. Illustration of the phasing of the start pulses and the neutron pulses when all delay circuits are in the zero position.

The start pulse, which initiates the gating unit chain is derived from the even stop pulse series. The edge from delay *III* is used to produce a gating waveform, called the start pulse gate, which gates the next even stop pulse. This triggers the start pulse output blocking oscillator. It is necessary to place the stop pulses correctly with respect to the edges from delay III. This is accomplished by setting the odd-even separator in the right condition with the edges from delay *III*.

When a gate width of 10 μsec is desired (fig. 7*b*) the odd-even separator is not driven by the 10 μsec pulse series from the main timing unit, but by a 10 μsec series which is derived from the 2 μsec pulse series by gating the appropriate pulses. The waveform which accomplishes this gating is controlled by the 10 μsec pulse series from the main timing unit.

The phasing of the start pulses and the neutron pulses is illustrated in *fig. 8*. The modulator pulses are delayed on the pre-pulses by a delay which is continuously variable from 2 to 5 μsec. The modulator itself introduces a delay of from 1 to 3 μsec. The neutron pulses can therefore be delayed on the pre-pulse from 3 to 6 μsec with a modulator delay of 1 μsec, or from 5 to 8 μsec with a modulator delay of 3 μsec. With all delays set at zero the start pulses are delayed with respect to the pre pulses by 6 μsec. Of this, 4 μsec is contributed by the

the start pulse and the linear accelerator electron pulse. Any delay T now introduced by the delay circuits will be accurately defined by the setting figure of these circuits.

The circuit used for establishing the variable time delay of the modulator pulses with respect to the pre-pulses is illustrated in *fig. 9*. The two triodes of B_1 are strapped in parallel and are normally conducting. At a time coincident with the pre-pulse, both valves are rendered non-conducting

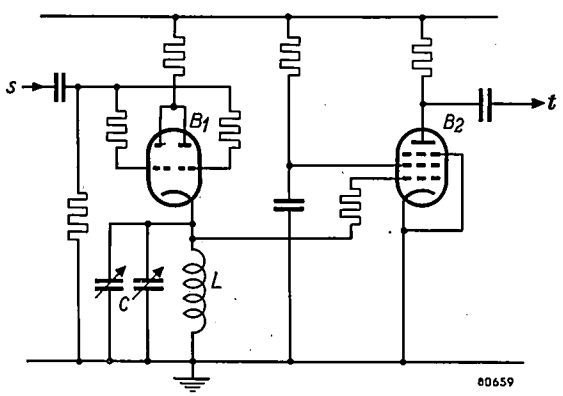


Fig. 9. Circuit diagram of the modulator pulse delay cricuit. The delay time is controlled by the LC circuit in the cathode lead of B_1.

by a downgoing edge applied to terminal *s*. A sinusoidal voltage is developed across the tuned circuit which forms the cathode load. Since in the first half cycle the cathode of B_1 is negative with

respect to earth, B_2 which is normally conducting is rendered non-conducting. At the conclusion of the first half-cycle B_2 is again driven into conduction and the downgoing edge in the anode is used to trigger the blocking oscillator which generates the modulator pulses. The waveform then tries to drive the control grid of B_2 positive and the resultant grid current damps out the sinusoidal voltage at the cathode of B_1. The delay between the pre-pulses and the modulator pulses is controlled by the tuning frequency of the cathode circuit of B_1, which can be varied by adjusting the tuning capacity.

The gating unit chain

The requirement of the gating unit chain is that the individual units shall be rendered operative for successive intervals ΔT of 2 or 10 μsec. To achieve this, every unit is provided with a series of stop pulses, odd or even depending upon its position in the chain. After the delay T, the start pulse opens gating unit number one which is closed by the next odd stop pulse a time ΔT after the start pulse. The closing of gating unit number one is accompanied by the generation of a pulse, called a "start-out" pulse, which opens gating unit number two. Number two is closed by the next even stop pulse after the interval ΔT and opens number three. Each unit therefore provides a pulse to open the next unit and the units are rendered operative for successive intervals ΔT.

Each individual gating unit incorporates a gating waveform generator, a gating valve, a start-out pulse generator and a register pulse generator (*fig. 10*).

The gating waveform generator consists of a direct coupled multivibrator with two stables states. It is normally held in one stable state by the application of the stop pulse series. The start pulse triggers it to the other stable state and the next stop pulse triggers it back to its normal stable state. It thus provides a square wave of duration ΔT equal to the interval between the start and the stop pulse.

This square wave or gating waveform is applied to the control grid of the gating valve. All the signal pulses derived from the detection chamber are applied to the suppressor grid. Hence each gating valve passes to its anode any signal pulses arriving during its gating interval ΔT.

The gated signal pulses are of too short a duration to operate a mechanical counter, and they are therefore used to trigger the register pulse generator, which is a cathode-coupled multivibrator with a stable and a semi-stable state. The duration of the semi-stable state is arranged to be 70 μsec and is used to operate a high speed relay which in turn operates the appropriate mechanical counter on the register desk. No scaling units are required as the probability of two signal pulses being gated by a particular gate during 70 μsec is extremely remote.

The screen grid of the gating valve is connected to a pulse generating circuit consisting of an 85 μH choke in parallel with a crystal diode. This circuit generates a positive pulse when the current in the gating valve is cut off at the termination of the gating interval. This pulse is fed to the next gating unit via a cathode follower.

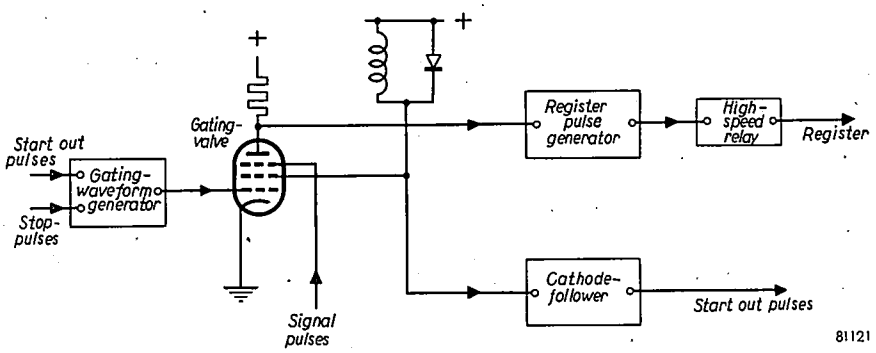


Fig. 10. Schematic diagram of a gating unit. (The first gating unit is opened by the start pulses instead of the start-out pulses.)

The total gating unit, provided at the end of each chain of gating units, is opened by the initial start pulse and closed by the start-out pulse from the last unit in the chain. It therefore gates all the signal pulses, gated by the individual units in the chain. The gated pulses are counted by the total count scaler and provide a check on the experimental results.

The duration of the total gating waveform is the sum of the gate width of the chain of gates associated with it and is used to operate the fault alarm system. Any failure in the start pulses, stop pulses or an individual gate will render the total gate inoperative. The resultant absence of the total gating waveform causes the fault alarm circuit to stop the counting and sound an audible alarm.

Accuracy of the measurements

Assuming correct operation of the frequency dividers and delay circuits, the accuracy of the electronic system is dependent upon the principal timing pulses of which all except the modulator pulses are derived directly from the 2 μsec pulse series. As this primary pulse series is derived from a crystal controlled 500 kc/s oscillator, the timing errors contributed by the principal timing pulses will be very small (less than 0.1 μsec).

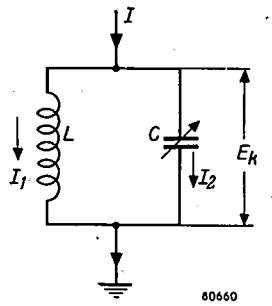


Fig. 11. Cathode circuit of valve B_1 of the modulator pulse delay circuit (fig. 9) showing the notation used in the calculation of the delay time.

Any variation in the relative phasing of the pre pulses and the modulator pulses will contribute to the inaccuracies of the electronic system, since the operation of the gating units is referred to the primary pulse series in the same way as the pre-pulses (see fig. 8) and the accelerator is actuated by the modulator pulses. The accuracy of the time delay which is provided by the modulator pulse delay circuit (fig. 9) will now be calculated.

An equivalent circuit of the cathode circuit of the valve B_1 is shown in *fig. 11*. The valve provides a current I, where $I = I_k$ when $t < 0$, and $I = 0$ when $t > 0$.

The circuit equations are:

$$-L\frac{dI_1}{dt} = \frac{q}{C}, \quad I = I_1 + I_2 \text{ and } \frac{dq}{dt} = -I_2,$$

where q is the charge of the capacitor C. Eliminating I_1 and I_2:

$$\frac{q}{LC} + \frac{d^2q}{dt^2} = 0,$$

so that for $t > 0$ we get:

$$E_k = \frac{q}{C} = -I_k\sqrt{\frac{L}{C}}\sin\frac{t}{\sqrt{LC}}.$$

As the modulator pulse is delayed with respect to the pre-pulse by half a cycle of E_k, this delay time amounts to $T_d = \pi\sqrt{LC}$ and the amplitude of E_k is $I_k\sqrt{L/C}$.

The accuracy of the time delay is dependent upon the grid voltage at which B_2 starts to conduct and upon the stability of the inductance and capacity of the tuned circuit. Measurements on samples of the pentode used (type EF 91) indicate that the spread in control-grid cut off potential is dependent on the screen grid potential, becoming smaller as the screen grid potential is decreased. With a screen grid potential of 150 V the spread in cut off potential is from 3 to 4·5 V.

The slope of the waveform at any time t is given by

$$\frac{dE_k}{dt} = -\frac{I_k}{C}\cos\frac{t}{\sqrt{LC}}.$$

Hence, at the instant when B_2 becomes conducting ($t = \pi\sqrt{LC}$) the slope is equal to I_k/C. The capacity C is variable between 88 and 345 pF. As $I_k = 19$ mA we obtain for the slope of the waveform a value of 217 V/μsec for the minimum value of C, and a value of 55 V/μsec for the maximum value of C. The possible variation in the control-grid cut off poten-

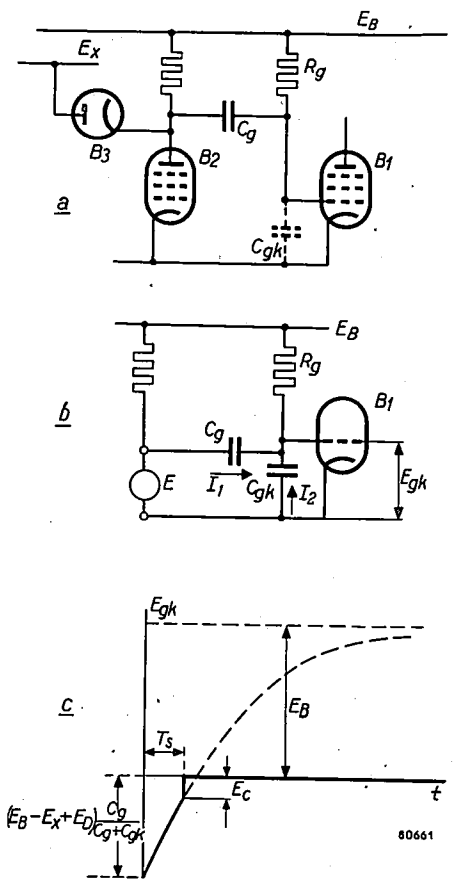


Fig. 12. *a*) Simplified diagram of a part of one of the delay circuits (fig. 5).
b) Equivalent diagram of fig. 12*a* in which the valve B_2 has been replaced by a voltage source E generating a step function voltage.
c) The voltage E_{gk} between grid and cathode of B_1 as a function of time.

tial of B_2 will thus cause at most a variation in the minimum delay time of 0·007 μsec or 0·35% and a variation in the maximum delay time of 0·027 μsec or 0·54%.

The stability of the inductance is almost solely dependent on the temperature coefficient of the permeability of its Ferroxcube core. This is quoted as 120 parts in a million per degree centigrade. The corresponding variation in delay time is therefore 0.12% for a 20°C variation in temperature.

The capacity is made up of the following: a two gang variable tuning condenser, the input capacitance of B_2, the heater to cathode capacitance of B_1, the coil capacitance, the wiring capacitance and the two valve-base capacitances. The coil capacity, the wiring capacities and the valve-base capacities are substantially constant for a particular assembly. The stability of the variable capacity is better than 0.1%. The input capacity of an EF91 and the heater to cathode capacity of both parts of an ECC81 are subject to a variation of ± 1pF. The minimum capacity is therefore subject to a maximum variation of 3 pF in 88 pF giving a variation in delay time of 1.71%. The maximum capacity is subject to the same variation with a corresponding delay time variation of approximately 0.44%.

The overall accuracy of the time delay is therefore 1.20% or 0.06 μsec at maximum delay time and 2.28% or 0.046 μsec at minimum delay time. Since the overall uncertainty of the electronic system is determined by the 2 μsec gate width, a maximum variation of 0.06 μsec contributed by the modulator pulse will have little effect.

Reliability of the installation

The reliability of the spectrometer depends upon the correct operation of the multivibrator dividing and delay circuits. The timing function of the multivibrator is governed by its semi-stable state and an examination of the factors governing this state is therefore required.

In *fig. 12a* a simplified diagram of a part of fig. 4 is shown. During the stable state, when B_1 is conducting, the flow of grid current in this valve keeps its grid-cathode potential approximately at zero. At the moment where the initiating pulse is applied to the cathode of B_2 the anode potential of this valve drops from E_B to $E_X - E_D$, where E_D is the anode-cathode potential of the clamping diode B_3 and the control grid potential of B_1 (V_{gk}) falls to a negative value

$$(E_B - E_X + E_D)\frac{C_g}{C_g + C_{gk}} .$$

E_{gk} then increases as is shown in figure 12*c* till in the absence of resetting pulses it reaches a certain critical value E_c when the circuit will reset itself into the stable state.

Assuming that the time for the anode potential of B_2 to decrease from E_B to $(E_X - E_D)$ is small compared with the duration of the semi-stable state, B_1 can be replaced by a voltage source generating a step function E as shown in the equivalent circuit in figure 12*b*.

The equations governing the circuit are:

$$E_B - E = \frac{q_1}{C_g} + (I_1 + I_2) R_g,$$

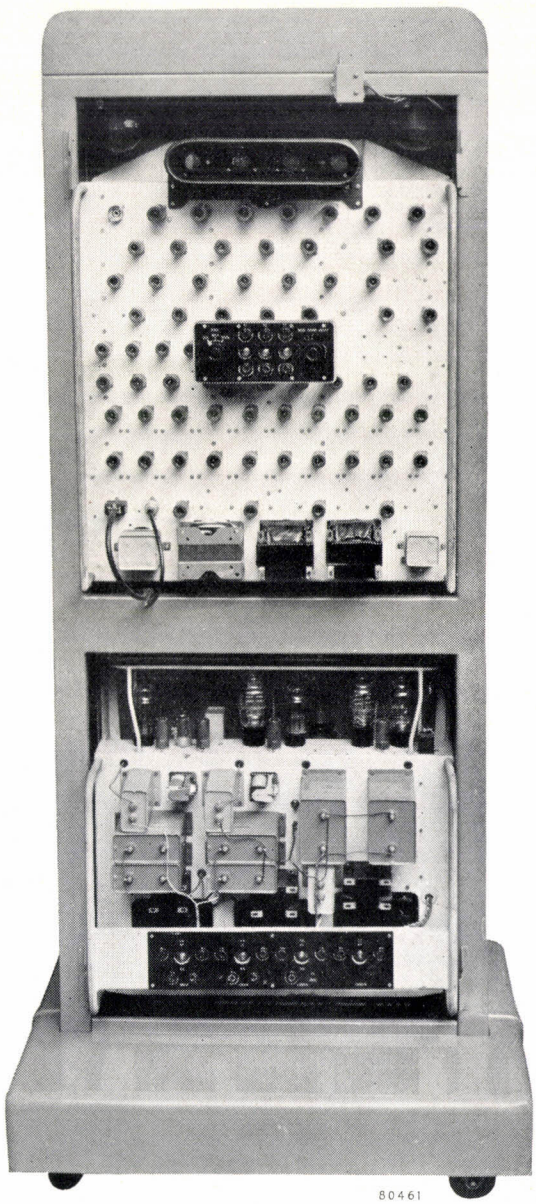


Fig. 13. The master timing unit. Front view with covers removed, showing the valve side of the main chassis (top) and the power unit (bottom).

$$E_B = (I_1 + I_2)\, R_g + \frac{q_2}{C_{gk}},$$

$$I_1 = \frac{dq_1}{dt},$$

$$I_2 = \frac{dq_2}{dt},$$

where q_1 and q_2 are the charges of C_g and C_{gk} respectively.

From these equations an expression is obtained for the voltage E_{gk}:

$$E_{gk} = \frac{q_2}{C_{gk}} = \left\{\frac{(E_B - E)C_g}{C_g + C_{gk}} + E_B\right\} \times \left\{1 - \exp\left(\frac{t}{(C_g + C_{gk})R_g}\right)\right\} - \frac{(E_B - E)\,C_g}{C_g + C_{gk}},$$

or, as $C_g \gg C_{gk}$,

$$E_{gk} = (2E_B - E)\left\{1 - \exp\left(\frac{-t}{C_g R_g}\right)\right\} - (E_B - E).$$

If we denote the time for E_{gk} to reach the critical value E_c by T_s,

$$T_s = -C_g R_g \log_e \left\{\frac{E_B - V_c}{2E_B - E}\right\},$$

where $E = E_X - E_D$.

The way in which T_s depends on the stability of the factors determining T_s can now be deduced. The actual circuit conditions are $E_B = 300$ V, $E_X = 150$ V, E_D varying from 1.0 to 3.6 V, and E_c varying from 3.4 to 4.6 V. E_B and E_X are provided by stabilized supplies of expected maximum variations of 1%. With the values mentioned we calculate that a variation of 1% in E_B or in E_B gives a variation of 0.8% in T_s.

A variation of E_D from 1.0 to 3.6 V gives a variation of 2.4% in T_s and a variation of E_c from 3.4 to 4.6 V gives a variation in T_s of 1.1%.

A variation of $C_g R_g$ is in general caused by the resistor. For high stability resistors up to 1 megohm a stability of 2.5% can be realized.

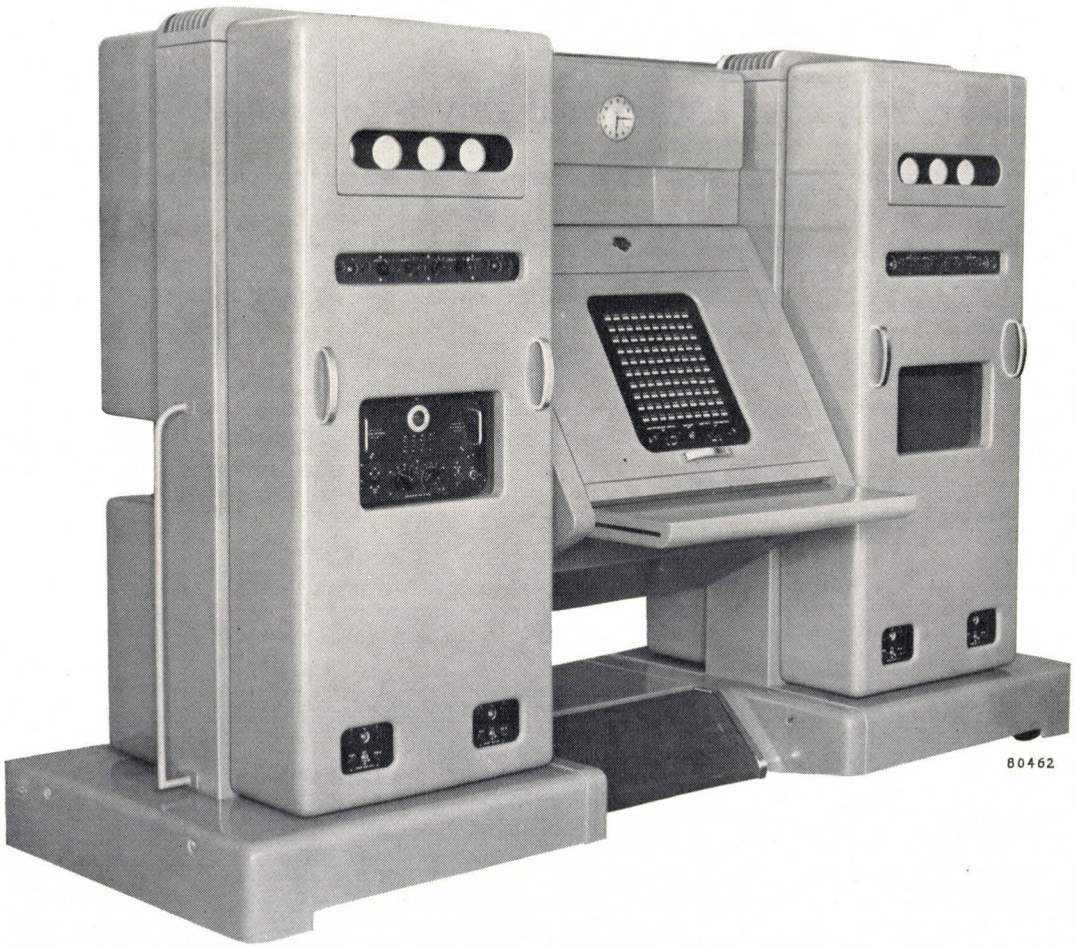


Fig. 14. The main assembly. Front view showing in the centre the register and control desk. At the top of the left and right sections are the delay unit monitor tubes; below them are the delay unit control panels. Half way down on the left section is a total count scaler (on the right there is a position for another scaler). At the bottom. left and right, are the power supply control panels.

It follows from these calculations that the duration of the reset time of each multivibrator circuit is subject to a variation of maximum value 7.6 %. Should this variation exceed the period between the resetting pulses, the danger arises that the circuit might be reset into the stable state by a wrong pulse. When the division ratio is 10, this being the maximum value occurring in the installation, the permissible variation in T_s in this respect is 10%, which is not exceeded by the figure calculated above.

mounted in the top of the rack which draw air in through a grid at the bottom. Vertical mounting of the chassis is employed so that they can be maintained without removal from the rack. *Fig. 13* shows a front view of the master timing unit with the covers removed.

The main assembly consists of two racks mounted on either side of the register desk. Each rack houses a delay unit, a total count scaler, and four power supplies, each rack again being cooled by two fans mounted in the top of the rack and drawing air

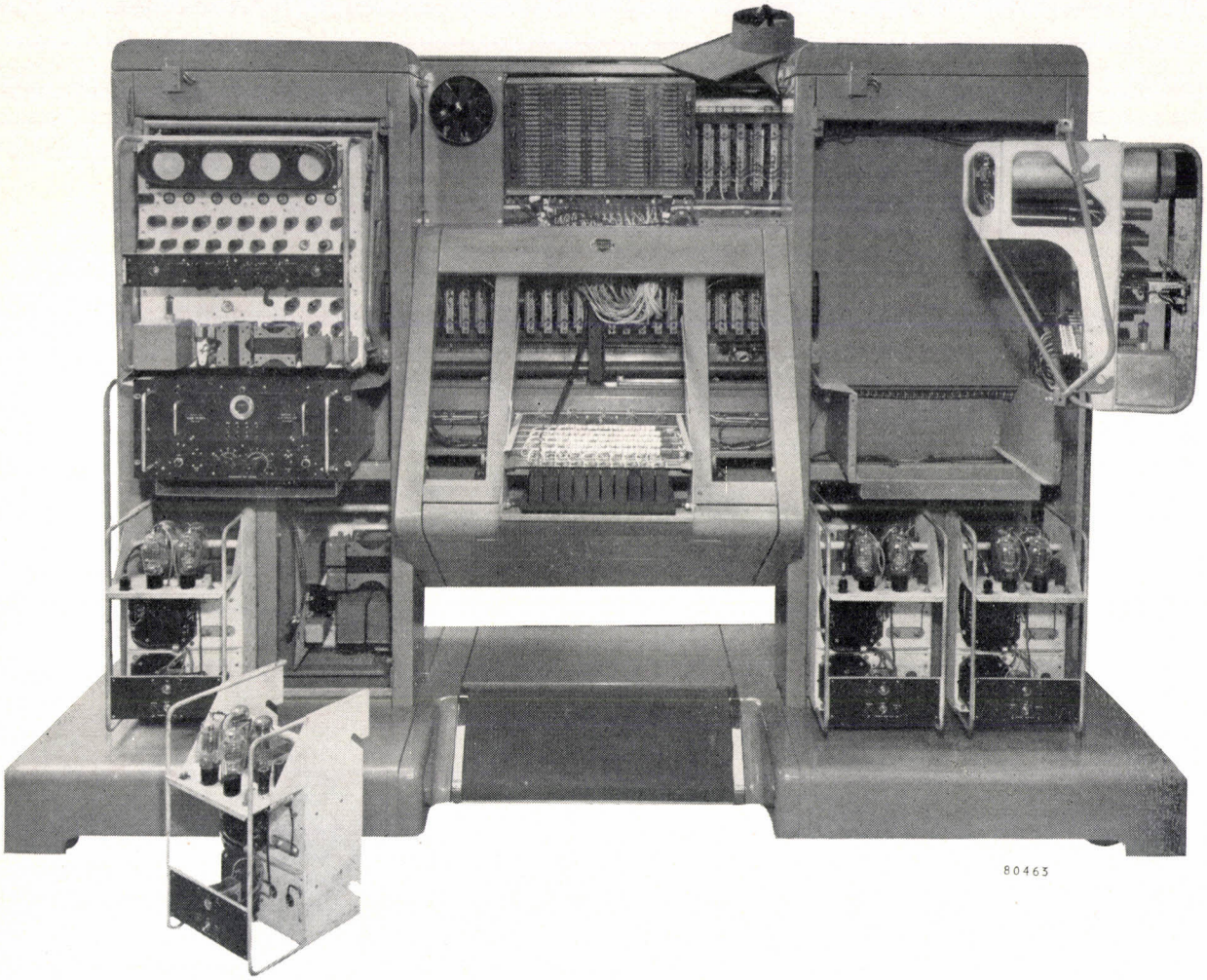


Fig. 15. The main assembly. Front view with covers removed. The right hand delay unit shows the hinged mounting framework and the input connections. One of the power supplies has been removed to show the method of mounting. The register desk has been hinged down and one of the fans has been removed, so that the rear side of the gating unit rack is visible.

Construction of the equipment

As the master timing unit is designed to feed timing pulses to four sets of delay units, it is housed in a separate rack together with its associated power supplies. The timing pulses are brought out to a series of coaxial sockets mounted on a channel at the foot of the equipment. Cooling is effected by two fans from the bottom. The delay units are mounted vertically in a hinged frame so that they can be maintained without removing them from the rack. A.C. supplies to, and D.C. supplies from each unit are brought to four tag boards mounted on either side of each rack. *Fig. 14* and *15* show the main assembly with and without covers.

The gating unit rack, which is shown in *fig. 16*, is wired as a separate unit and is supported on the back of the main framework by four brackets. In consists of 103 frames designed to take 100 gating units and 3 total gating units. Each frame is fitted with 11 spring loaded contacts to take the input and outputs to and from each unit. This provides for immediate rectification of faults in the gating unit chain as all the units are interchangeable and any faulty unit is easily replaced by one of the spare units.

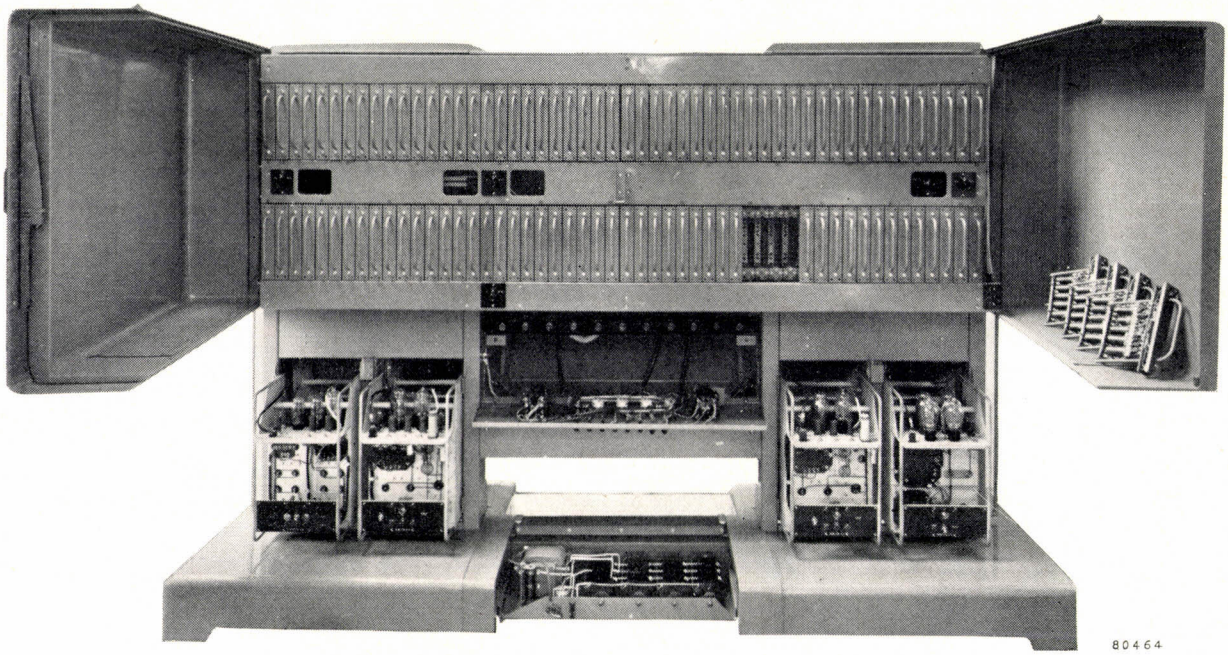


Fig. 16. The main assembly. Rear view with covers removed, showing the grating unit rack which houses 100 gating units and three total gating units. Four of the units have been removed and are lying in the right hand cover. The main input panel in the centre has been hinged down to show the main contactors. Power packs are also visible (bottom left and right). In the centre of the base is the 50 V power supply for the register desk.

The five switches dividing the rack into two separate chains of gating units are housed at the appropriate points along the bus bars which feed the start and stop pulses to the units. Cooling is effected by fans, two housed above the ends of the rack and the other two mounted horizontally drawing air from the central portion of the rack and blowing it out from the front of the unit. In order to ensure that the air circulates through the gating units, the whole rack is enclosed by two covers hinged at the end and having a wire mesh running along the bottom flange. The maximum temperature rise using this arrangement is 35 °C above ambient.

The register desk with its 100 mechanical counters is mounted at an angle to facilitate reading of the numbers. Connections from the registers to the gating units are made by a cable form wired onto a tag board on the back of the gating unit rack. Also incorporated in the register desk are the count keys, the fault alarm and the A.C. input switches.

The timing pulses are fed from the master timing unit by 100 Ω coaxial cables to coaxial sockets mounted at the back. Each cable is terminated by a 100 Ω resistor. The A.C. inputs are taken in via a hinged input panel mounted below the pulse inputs. A.C. switching is accomplished by two gravity return contactors. In order to ensure that the gating units are not operated without their bias supply, the contactor switching the A.C. to the power supplies is operated via a relay which in turn is operated by the bias supply.

The covers, with the exception of the gating unit rack covers, are fitted with interlock switches wired in series with the operating coil of the H.T. main contactor.

The electronic assembly can be seen in the background on the photograph of fig. 1, which gives a general view of the whole installation.

Acknowledgements

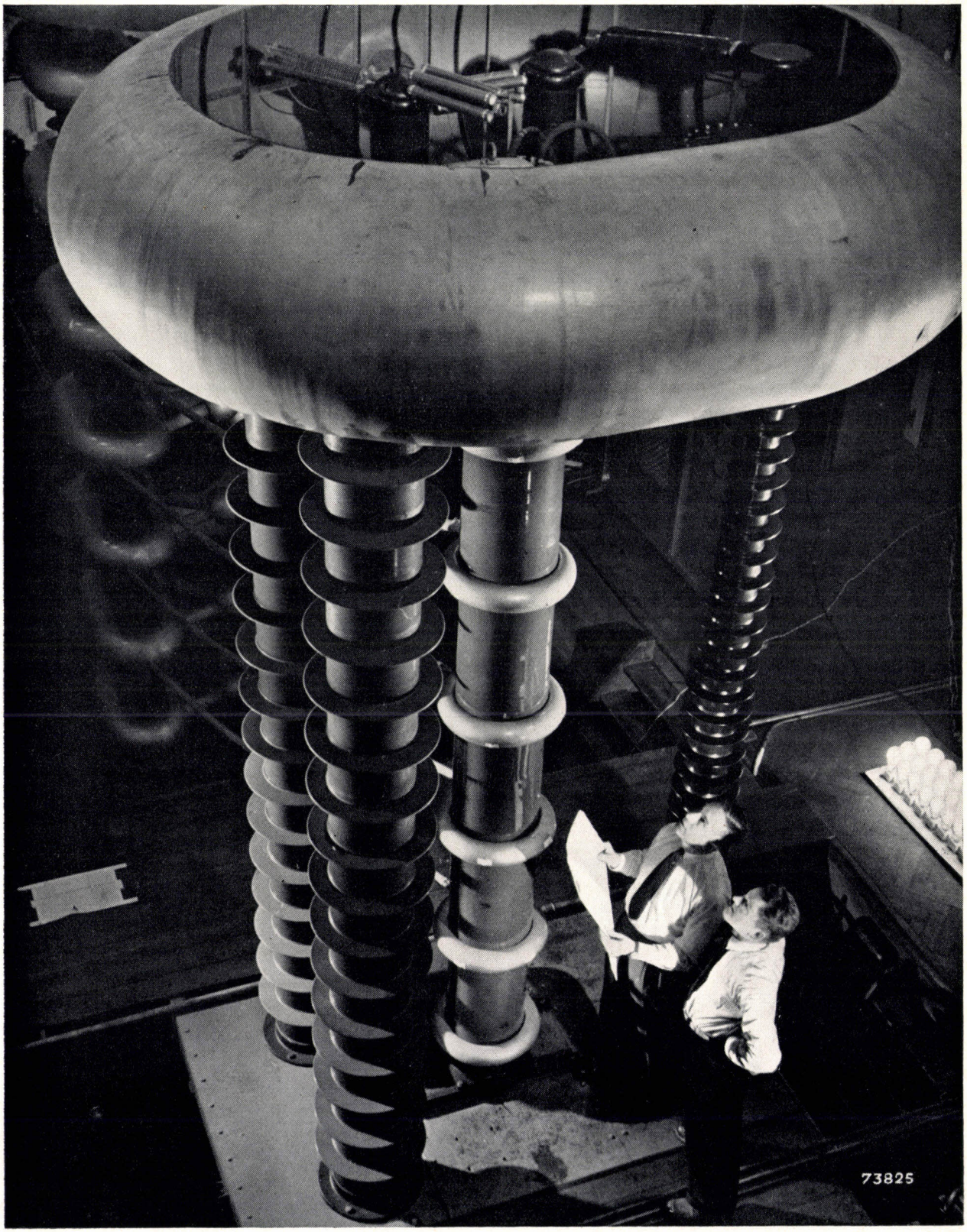
The authors wish to acknowledge the help of their colleagues at the Mullard Research Laboratories and of members of the Electronic and Nuclear Physics Divisions at A.E.R.E. during the course of the

development reviewed and in the preparation of this article. Thanks are also due to the Director of A.E.R.E. and to the Directors of the Mullard Valve Company for permission to publish this article.

Summary. In this article a description is given of a time-of-flight neutron spectrometer installed in the Atomic Energy Research Establishment at Harwell and used for measurements of the energy levels of neutron scattering and resonance absorption. The heart of the installation is the 15 MeV linear accelerator previously described in these pages. A 2 μsec electron pulse after travelling down the accelerator is used to irradiate a γ-ray and neutron source. Some of the liberated neutrons travel down a 20m flight path, after which they have become spaced out in time according to their energies. The installation contains two of these flight paths, which can be used separately for independent experiments. At the end of each flight path the neutrons are detected by a neutron detector. The electrical signal pulses from this chamber constitute a series whose distribution in time corresponds to the energy spectrum. This pulse series is fed to the input rail of a bank of 100 gating units, which are opened successively for equal short intervals ΔT starting with a time delay T from the initiation of the neutron pulse. Each gating unit output is fed to a mechanical counter. The intervals ΔT can be set at 2 or 10 μ sec and the delay time T is adjustable between 0 and 2990 μ sec. The principal recurrence frequency can at present be set at 200 or 400 p.p.s.

The functioning of the installation is based on a number of timing pulse series which are generated in the master timing unit. They are all controlled by a primary pulse series of 2 μ sec period derived from a 500 kc/s crystal controlled oscillator. The other pulse series are derived from this primary series by four frequency dividers (multivibrators). The time delay T is provided by a delay unit which contains three adjustable delay circuits in cascade. Each of these circuits consists of a multivibrator with a stable and a semi-stable state. Each gating unit contains a multivibrator with two stable states, which is triggered by two series of pulses, the start out pulse series and the stop pulse series. To prevent coincident triggering, two series of stop pulses are used, feeding the odd and even gating units respectively. The 100 gating units can be used either in a single sequence or in two independent banks for experiments on both flight paths. To provide a check on the correct operation of the installation a total gating unit is provided for each sequence of gating units. Associated with these total gating units is a fault alarm system which sounds an alarm whenever a fault occurs in a gating unit sequence. Since a normal experiment takes several hours of (usually) unattended operation, this alarm avoids waste of experimental time. At the end of the article some calculations are made concerning the accuracy and the reliability of the installation and a short description is given of some constructional details.

EXPERIMENTAL SET-UP OF A 1 MILLION VOLT ION-ACCELERATING TUBE



A source of hydrogen or deuterium ions is mounted in the top protecting cover, supported by three columns. The ions are accelerated in the 4 metre-long acceleration tube (centre). The cascade generator, supplying the acceleration voltage of 10^0 V, is visible in the background.

PHOTOGRAPHY OF THE EYE WITH THE AID OF ELECTRONIC FLASH-TUBES

by J. E. WINKELMAN *) and N. WARMOLTZ.

771.447.4:617.7

The electronic flash-tube, apart from its extremely short flash duration, good light output and high efficiency, possesses other important characteristics. One of these is the great variety of shapes in which these tubes may be constructed; another is their flexibility in operation. These characteristics are illustrated by the special flash-tube for ophthalmic puposes described in this article, which makes it possible, in a simple set-up, to obtain excellent well-defined colour photographs of the human eye.

In ophthalmology, photographic records of patients nowadays occupy an important place. Well-defined photographs, particularly those in colour, provide a far more complete record than the sketches on which the ophthalmologist formerly had to rely.

In this article we shall deal with an apparatus for eye photography that has been developed in the Wilhelmina Hospital ophthalmic clinic in Amsterdam. This apparatus serves for photographing the externally visible parts of the eyeball (the anterior segment) which includes (*fig. 1*) the cornea, the sclerotic coat or white of the eye, the conjunctiva, the anterior chamber, the iris, the pupil, the anterior surface of the lens, the eyelids and the adjoining skin, and parts of the lachrymal apparatus. *Fig. 2* shows four colour photographs made with the present equipment of the eyes of different patients. For photographing the posterior segment of the eye (the fundus oculi, including the retina), different set-ups are used, which will not be dealt with in this article.

As a photographic object the eye, in particular the diseased eye, presents considerable difficulties. Abnormalities of the eye are as a rule of very small dimensions, so that at least full-scale photographs are required. For this the camera has to be brought up very near to the object, which in turn requires the use of a very small aperture in order to produce the necessary depth of focus. If we consider, moreover, the relatively low sensitivity of colour film, then it will be clear that photographs of this kind require a large quantity of light. Even for the most powerful sources of contiuous illumination, this involves fairly long exposure times, e.g. some seconds. Difficult as it may be to keep a normal, healthy eye from moving during such a period, for a diseased eye this is well-nigh impossible.

An almost ideal solution of this problem was found in the electronic flash-tube. This light source, originally developed for highly specialized purposes [1]), but now generally applied in photography [2]), is commercially available in a number of forms. It consists basically of a glass tube, filled with rare gas at not too low a pressure, through which a

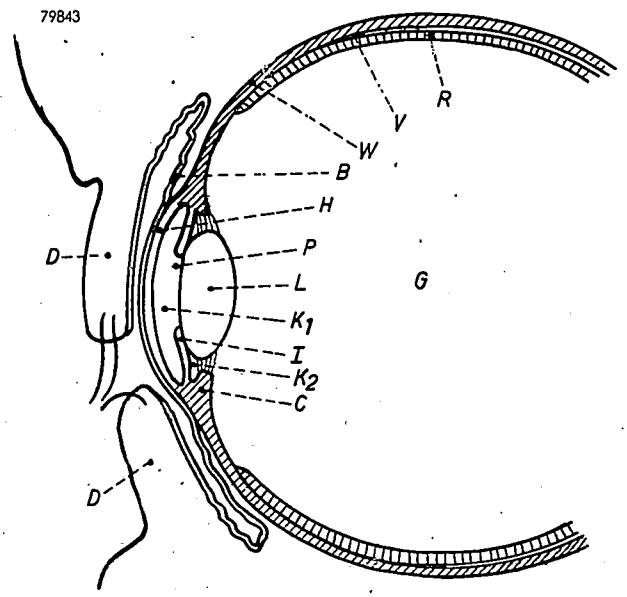


Fig. 1. Cross-section of the anterior segment of the eye. *D* eyelids, *H* cornea, *W* sclerotic coat or white of the eye, *B* conjunctiva, *L* crystalline lens, *I* iris, *P* pupil, K_1 anterior chamber. Also shown are, K_2 posterior chamber, *C* ciliary muscle, *V* choroid, *R* retina, *G* vitreous humour.

*) Ophthalmic Clinic, Wilhelmina Hospital, Amsterdam.

[1]) See e.g., S. L. de Bruin, An apparatus for stroboscopic observation, Philips tech. Rev. **8**, 25-32, 1946: N. Warmoltz and A. M. C. Helmer, A flash lamp for illuminating vapour tracks in the Wilson cloud chamber, Philips tech. Rev. **10**, 178-187, 1948.

[2]) Cf. e.g. Photogr. J. **89B**, May-June 1949 (symposium on electronic flash lamps), or: M. Laporte, Les lampes à éclaire lumière blanche et leurs applications, Gauthier-Villars, Paris 1949. (A comprehensive article on electronic flash-tubes will appear in the next issue of this Review — Ed.)

capacitor is made to discharge via two electrodes with the aid of a triggering electrode. A flash is thus produced, having a duration varying between a few microseconds and a few milliseconds, dependent on the construction of the lamp, and of such a high intensity that the total light output emitted amounts to some thousands or some tens of thousands of lumen seconds. During such extremely short periods the eye shows no discernible reaction to the illumination: this occurs later, when the photograph has already been made. The luminous efficiency of electronic flash-tubes is considerable (of the order of 40 lumen/watt), so that no excessive demands are made on the electrical supply unit. When compared with ordinary flash-bulbs, for this particular application, the electronic flash-tube has the advantage that it develops hardly any heat (to which some eye patients may be extremely sensitive). Furthermore, a series of exposures need not be interrupted for the exchanging of the bulbs. Finally, the flash-tube allows a simple solution to a very important aspect of this particular application — the setting up and focussing of the camera (see below).

In principle, various commercially available flash-tubes are suitable for the purpose described here and there have already appeared several publications on set-ups of this kind [3]). The apparatus

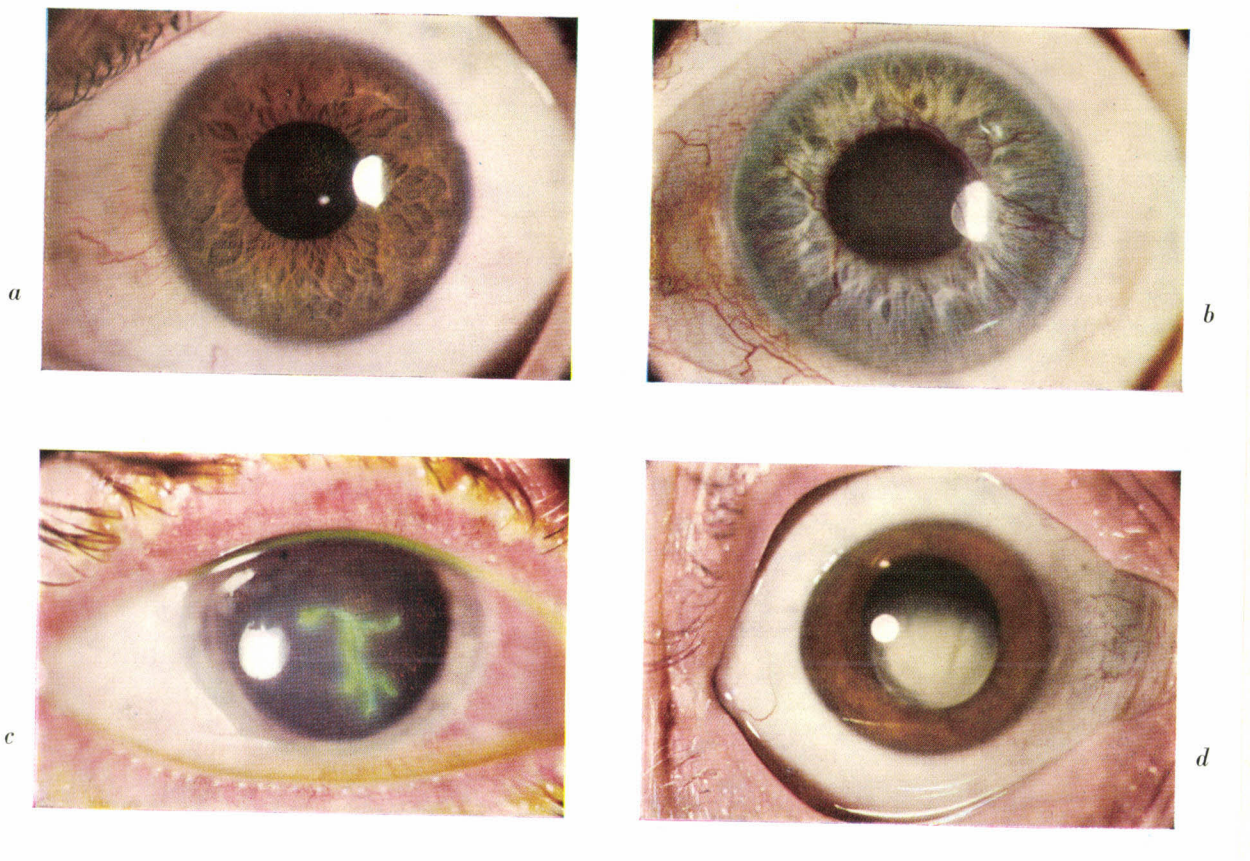


Fig. 2. Colour photographs of the eyes of different patients, made with the apparatus described, on daylight colour film (for use with a source of colour temperature 6000-6500 °K).
a) Normal eye. Note the excellent reproduction of the texture of the iris.
b) Eye in diabetes. Fine blood vessels can be seen running across the iris.
c) Eye with dendritic keratitis in the cornea. The dendritic ulcus of the cornea is coloured green with fluorescein.
d) Eye with loosened and displaced lens. On top the dark pupil is visible and below it the sagging, turbid lens.

[3]) See e.g., R.R. Trotter and W. M. Grant, Electronic flash (gas-dicharge) tube in photography of the anterior segment of the eye, Arch. Ophthalm. **40**, 493-496, 1948; N. Jeffreys, Problems of ophthalmological photography, J. Phot. Sci. **1**, 184-192, 1953 (Nov.-Dec.); G. Meyer-Schwickerath, Flash photography of the diseased eye, Photographie u. Forschung **5**, 170-173, 1953 (June).

developed in Amsterdam, however, was based on the assumption that a simple but effective set-up could best be realized by adapting the light-source to the specific requirements of this particular application. These requirements may be outlined as follows. The object to be photographed is very small — its diameter will not exceed 4 or 5 cm. In order to concentrate a substantial portion of the light of the flash tube on to such a small area by means of a lensor mirror of reasonable proportions, the flash tube should preferably be very small. Furthermore, in order to obtain a uniform illumination of the curved surface of the eye and to direct a sufficient amount of light on the iris, the light source must be placed as near as possible to the line of sight of the camera. This also requires that the source should be of small dimensions. Finally the possibility of undesirable reflections must be considered both from the cornea, which acts as a convex mirror, and from the conjunctiva. This can best be avoided by a set-up permitting a prior examination of the location of the reflections. Such an examination is possible with the aid of an auxiliary light-source, so arranged to illuminate the eye from exactly the same angle as the subsequent flash. For this purpose a filament lamp may be used, of low intensity to avoid discomfort to the patient.

All these considerations have led to the construction of the special flash-tube shown in *fig. 3*. It consists of three almost co-planar turns of rather narrow glass tube. In the centre of the spiral is an aperture approximately 1 cm in diameter. For the auxiliary lamp an experimental type of reflector lamp has been chosen which has an internal ellipsoidal mirror. This produces an image of the filament about 2 cm outside the bulb. The reflector lamp is so positioned behind the flash-tube that the image of the filament is produced within the aperture of the spiral. This image and the flash-tube are projected by means of a "Perspex" condenser lens on the eye to be photographed (*fig. 4*). The outside of the condenser is provided with a sheet of ground glass, to provide some dispersion of the light, as otherwise the structure of the source would show up on the object.

The whole lighting system is contained in a casing, mounted close to the camera on a single stand (*fig. 5*). The camera is of the 35 mm-type, provided with a mirror-reflex attachment (lens f 3.5, with a focal length of 5 cm). The distance from the lens to the patient's eye is normally 6 to 7 cm, producing a photograph about 1.5 times full size. The flash-tube is supplied from a 4 μF capacitor, charged up to about 5 kV, so that the flash energy amounts to roughly 50 joules. Nothwithstanding this relatively small energy, the illumination of the object is sufficient to allow the use of apertures of f 16 and even

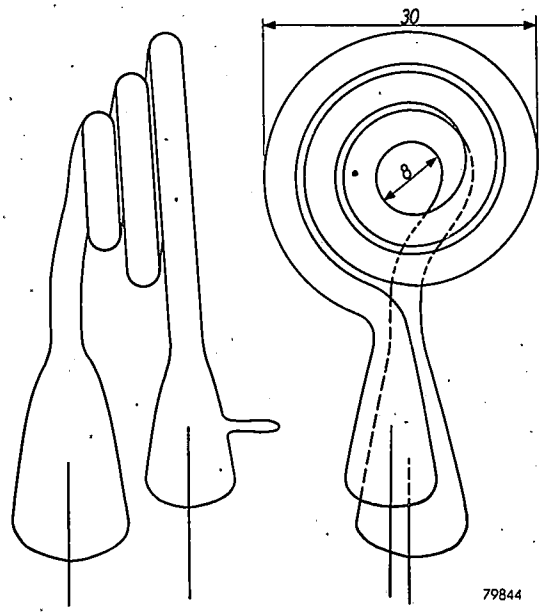


Fig. 3. Sketch of the electronic flash-tube developed for eye photography. The glass spiral has a central aperture of about 8 mm diameter.

smaller. The colour temperature of the flash-tube is approximately 6500 °K, corresponding to normal daylight, so that daylight colour film can be used.

The flash is normally triggered by contacts in the camera synchronized with the shutter, but the apparatus can, if necessary, also be used with a camera without electronic-flash synchronizing contacts. For this purpose a pedal switch is provided which successively switches off the auxiliary lamp, opens the camera shutter and triggers the flash.

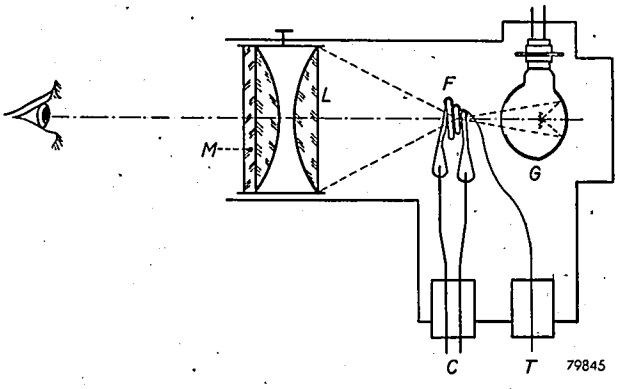


Fig. 4. The illumination system for eye photography. F electronic flash-tube, L lens, M sheet of ground glass. G is a filament lamp for focussing and sighting; its silvered bulb produces an image of the filament outside the envelope and the lamp is so positioned that the image falls within the central aperture of the flash-tube. C connections to capacitor, T lead for triggering pulse.

Switching off the auxiliary lamp is desirable in order to avoid the risk of even the slightest movement blurring and of incorrect colour reproduction due to the low colour temperature of the incandescent lamp.

Some of the photographs made with the apparatus discussed here, have been reproduced in fig. 2. In spite of the limitations imposed by the printing process, these colour reproductions give an impression of the excellent quality of the original photographs, due to the complete absence of movement during exposure and to the large depth of focus. Apart from providing a reliable record of the patient's condition, these high definition photographs have also proved of value in ophthalmic research. Great interest has lately arisen with regard to the flow of blood through the finest blood vessels of the conjunctiva, especially since Ascher's discovery [4]) of

[4]) K. W. Ascher, Amer. J. Ophthalm. **25**, 31-38, and 1174-1209, 1942; **27**, 1074-89, 1944; **29**, 1373-87, 1946.

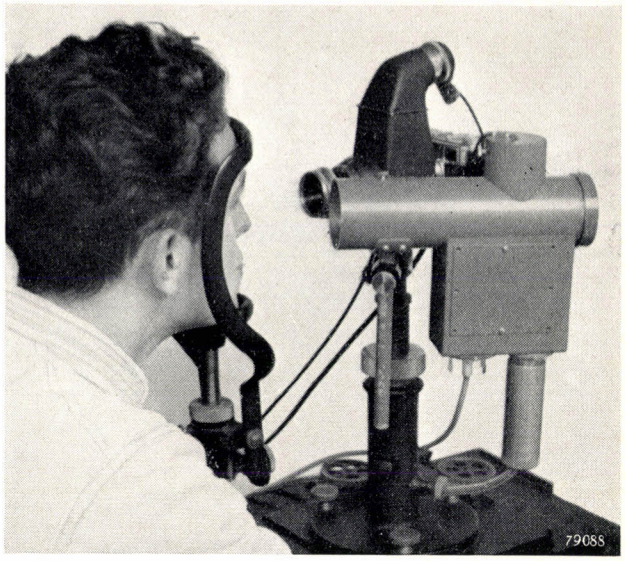


Fig. 5. Arrangement of camera and illuminating unit on a common stand of the type normally used for examinations of this kind. The patient's head is held in a fixed position by the usual ophthalmic chin-and-forehead rest.

Fig. 6. Part of the sclerotic coat and the conjunctiva of a normal human eye, enlarged about 20×. This photograph was made with the apparatus described in this article. The definition is sufficient to render visible even the fine ramifications of the blood vessels of the conjunctiva. At the right part of the cornea can be seen, and below left, a few eyelashes.

what are termed the "aqueous veins". As a rule these finest capilliaries can be observed only through a corneal microscope. Photography of this delicate network of blood vessels is a most difficult task. *Fig. 6* demonstrates that the definition of the negatives made with the set-up described (i.e. without a microscope) is sufficient to permit enlargements in which these fine blood vessels of the conjunctiva are easily visible.

The flash-tube and the further equipment were developed in cooperation with Drs H. J. J. van Boort of the Laboratory of the Light Group at Philips. Mr G. Lammens of the Amsterdam Ophthalmic Clinic who made the photographs reproduced here, played an important part in the development of the instrument.

Summary. Colour photographs of the externally visible parts of of the eye provide the ophthalmologist with an ideal means of recording his findings. Photographs of this kind require a very large quantity of light, but the hyper-sensitivity of abnormal eyes is a serious obstacle to the use of a continuous illumination of sufficient intensity. An electronic flash-tube, specially designed for use in conjunction with a filament lamp for focussing and sighting (which lights the eye from exactly the same direction) eliminates these difficulties. The lamp operates on a flash energy of only 50 joules. By means of a simple optical system for concentrating the light on the eye under examination, the illumination permits colour photographs of 1.5 × full size to be made with an aperture of *f* 16. Some colour photographs made with the apparatus are reproduced in the article, as well as a 20 × enlargement of a normal eye, in which the extremely fine blood vessels of the conjunctiva are visible, demonstrating the high definition attainable.

ABSTRACTS OF RECENT SCIENTIFIC PUBLICATIONS OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN

Reprints of these papers not marked with an asterisk * can be obtained free of charge upon application to the administration of the Philips Research Laboratory, Eindhoven, Netherlands.

2067: H. O. Huisman, A. Smit, S. Vromen and L. G. M. Fischer: Investigations in the Vitamin A-series, II. Allylic rearrangements in the Vitamin A-Series (Rec. Trav. chim. Pays-Bas **71**, 899-919, 1952, No. 8).

By splitting off water from the intermediate hydroxycompounds of the vitamin A-series, allylic rearrangement takes place in the cyclohexene ring. This rearrangement occurs throughout the whole vitamin A-series with those hydroxy-intermediates wherein the hydroxy-group is in an allyl position with regard to the double bond in the cyclohexene ring. The reason why so many attempts to prepare synthetic vitamin A in reasonable yields, starting with "key" intermediates other than the C_{14} aldehyde lead to failure, is outlined.

2068: J. A. Keverling Buisman and P. Westerhoff: Investigations on sterols, V. Thio-derivatives of provitamine D (Rec. Trav. chim, Pays-Bas **71**, 925-932, 1952, No. 8).

A description is given of the synthesis of the thio-analogues of ergosterol and 7-dehydrocholesterol from these two compounds.

2069: H. D. Moed, M. Asscher, P. J. A. van Draanen and H. Niewind: Synthesis of β-phenylethylamine derivates, II. Condensation of phenols with amino-acetonitriles (Rec. Trav. chim. Pays-Bas **71**, 933-944, 1952, No. 8).

A description is given of the condensation of amino-acetonitriles with phenols according to the method of Houben-Hoesch. The condensations have led to the preparation of a number of compounds of the type:

$$R_1R_2N\text{-}CH_2\text{-}CO\text{-}(C_6H_3R)\text{-}OR_3,$$

in which the R denote alkyl groups. A number o β-amino-acetophenones have been converted into compounds with sympathomimetic activity. An investigation has been made into the conditions and the mechanism of the synthesis, and the influence of substituents in the benzene ring on the reactivity of the phenol component. A discussion is devoted to the electronic interpretation of the synthesis according to Houben and Hoesch.

2070: B. Verbeek and A. H. W. Aten: Radiations from arsenic 70 (Physica **18**, 974-975, 1952, No. 11).

Description of the preparation of As^{70} by bombarding arsenic-free germanium oxide with 26 MeV deuterons. The half-life is found to be 52 ± 1 min. Positrons are emitted with maximum energy of 2.7 MeV. The upper limit of the ratio K-capture/β^+-emission is 0.2. Apart from the annihilation radiation, on the average 2 gamma quanta per positron are emitted with energies between 1 and 2 MeV.

2071: B. H. Schultz: Regenerators with longitudinal heat conduction (Proc. Gen. Discussion on Heat Transfer 1951, pp. 440-443, Inst. Mech. Engrs. London).

For the efficiency η of a non-subdivided regenerator (e.g. in a hot-air engine) in the case of large

heat capacity, the formula

$$\eta = 1-2(1+\Lambda\lambda)/q$$

is derived, in which Λ is the reduced surface and λ is the ratio of heat conductance of the regenerator to the heat capacity of the gas transferred per sec; q is a function of Λ and λ. It is shown that the same formula probably holds in the case of a subdivided regenerator (consisting of a number of slabs with no thermal contact between adjacent parts), provided the reduced heat capacity Γ (heat capacity of the regenerator divided by that of the amount of gas transferred per cycle) is sufficiently large ($\Gamma>3$).

2072: H. P. J. Wijn: A new method of melting ferromagnetic semiconductors: $BaFe_{18}O_{27}$; a new kind of ferromagnetic crystal with high crystal anisotropy (Nature **170**, 707, 1952).

A method is described for preparing single macroscopic crystals of magnetite (Fe_3O_4) and of various ferrites (MeF_2O_4), in which Me is a divalent metal. The material is melted in a nitrogen atmosphere by high frequency heating in a crucible made of Al_2O_3, covered by a disc of rhodium or indium, and allowed to cool in an atmosphere containing oxygen, the oxygen pressure being kept equal to the equilibrium pressure. Single crystals several millimeters in size were obtained. By adding barium carbonate, crystals of hexagonal structure with high crystal anisotropy were obtained, e.g. crystals of $BaFe_{12}O_{19}$ ($c = 23.2$ Å) and $BaFe_{18}O_{27}$ ($c = 32.8$) Å.

2073: P. Braun: Crystal structure of $BaFe_{18}O_{27}$ (Nature **170**, 708, 1952).

Details are given of measurements of the crystal structure of hexagonal $BaFe_{18}O_{27}$ (see No. **2072**).

2074: J. Volger: Anomalous specific heat of chromium oxide (Cr_2O_3) at the antiferromagnetic Curie temperature (Nature **170**, 1027, 1952).

The susceptibility χ of Cr_2O_3 was measured as a function of T between 80 and 550 °K. There is a discontinuity in the slope of the curve at about 318 °K (45 °C). The molecular specific heat C_p has a sharp peak at 308 °K (35 °C), which is due to the disappearance of the antiferromagnetic ordering at this temperature.

2075: A. L. Stuyts, G. W. Rathenau and E. W. Gorter: Preferred crystal orientation in ferromagnetic crystals (J. appl. Phys. **23**, 1218, 1952).

The compound $BaFe_{12}O_{19}$ is the main ingredient for a sintered permanent magnet material. With crystals oriented at random one obtains $(BH)_{max} =$ (0.8-1.1)10^6 gauss. oersted or (6.5-8.5)10^4 J/m^3. By orienting the crystals in a magnetic field before sintering, values roughly 3 times as large are obtained (see these abstracts No. **2059**).

2076: P. B. Braun: A superstructure in spinels (Nature **170**, 1123, 1952).

It is found that the compound $LiFe_5O_8$ has approximately a spinel structure, but with slight deviations which cause a superstructure. Below 735 °C four lithium ions are distributed in an ordered way over 16 octahedral interstices (with long range order). Above 755 °C the long range order has completely disappeared. By rapid quenching from 950 °C the disordered phase is obtained at room temperature. $LiAl_5O_8$ shows the same superstructure but the transformation region is at higher temperatures. It is pointed out that γ Fe_2O_3 containing water may approach the formula HFe_5O_8 (by analogy with HAl_5O_8). In the $\alpha\rightarrow\gamma$ transformation of Al_2O_3 and Fe_2O_3, protons may play an important part.

2077: H. de Lange: Experiments on flicker and some caculations on an electrical analogue of the foveal systems (Physica **18**, 935-950, 1953, No. 11).

The reaction of the author's right eye to a 2° foveal flicker field has been investigated and measurements of the critical flicker frequency for various illuminations and various brightness-time functions are tabulated. It is shown that in the graph of the ripple ratio (r = amplitude of first Fourier component/average brightness) of the stimulus against the critical frequency f_c, the average brightness being constant, the points observed with various time functions fit into one smooth curve which is monotonic at low intensities. At high intensities the curve shows a minimum, at which r is less than the limiting brightness discrimination ratio, observed for $f = 0$. If f_c is plotted horizontally, the curves obtained suggest an explanation of the mechanism present in the light sensitive organ and the connected nerve centres, by analogy with an electric network containing a lowpass filter, a feed-back circuit and a non-linear element which transforms the physical stimulus into a number of impulses.

R 200: A. Bril and H. A. Klasens: Intrinsic efficiences of phosphors under cathode-ray excitation (Philips Res. Rep. **7**, 401-420, 1952, No. 6).

Intrinsic efficiencies of thick phosphor layers and their voltage and temperature dependence were

measured with cathode-ray excitation in a demountable tube. The light outputs of thin layers of phosphors coated on glass were also determined. The light output on the glass side goes through a maximum with increasing layer thickness. For sulphide phosphors e.g., the maximum output on the glass side was about 30 percent of that on the phosphor side for a thick layer. Further losses of light occur in practice due to fabrication difficulties and current saturation, so that only 25 percent of the energy of the electron beam is radiated from the sulphide screen of a cathode-ray tube. For an ideally reflecting metal coating, emission of almost 100 percent can be obtained theoretically. In practice (aluminium layer) the maximum gain is a factor 2 or at most 2.2, due to imperfect reflections. From the measurements the scattering and absorption coefficients were determined for sulphide and silicate screens, using the formulae of Hamaker.

R 201: A. Bril and H. A. Klasens: New phosphors for flying-spot cathode-ray tubes (Philips Res. Rep. 7, 421-431, 1952, No. 6).

The decay rates of a large number of phosphors under cathode-ray excitation have been determined. Cerium-activated phosphors are the best for flying-spot cathode-ray tubes used for black-and-white television. Of these phosphors $2CaO.Al_2O_3.SiO_2$-Ce was selected as the best phosphor in combination with glasses not blackened by soft X-rays or cathode ray bombardment. Bismuth-activated phosphors are suitable as a red component in flying-spot cathode-ray tubes for colour television. A suitable phosphor for u-v flying-spot microscopy is ZrP_2O_7.

R 202: R. Vermeulen: Dimensional analysis, units and rationalization (Philips Res. Rep. 7, 432-441, 1952, No. 6).

Undiscriminating manipulation of dimensional formulae can lead to contradictions, e.g. that 1 sec = 3.10^{10} cm or that 1 oersted = 1000 A/m. These are eliminated by defining the multiplication of physical quantities, strangely enough never done before, and by preventing the mixing up of multiplications whose physical interpretations are essentially different. The same measures solve difficulties connected with the introduction of the rationalized Giorgi-system.

R 203: F. de Jager: Deltamodulation, a method of P.C.M. transmission using the 1-unit code (Philips Res. Rep. 7, 442-466, 1952, No. 6).

It is known that in a communication system the influence of interferences in the transmission path can be reduced considerably by coding the information signal first and transmitting then a corresponding pulse pattern of 0 and 1 pulses. In well-known systems of pulse-code modulation the n-digit binary code is used. In deltamodulation, however, a "code" comprising only 1 digit is used. Here the reproduced signal is obtained by applying the series of quantized pulses to a linear network. This system enables us to obtain a simplification of both coding and decoding devices. The conversion of the information signal into a quantized pulse pattern is achieved by using a negative-feedback circuit in which the voltage applied to the feedback network is quantized both in amplitude and in time. The network in the feedback loop should be related to the mean spectrum of the information signals. For speech an integrating network may be used as such. It is found, however, that the frequency characteristic of the feedback network for frequencies lying between the highest speech-frequency and the pulse frequency has its influence on the amount of quantizing noise in the reproduced signal. If a combination of single and double integration in this frequency region is used, the ratio between the r.m.s. values of signal and quantizing noise is proportional to the 5/2 power of the pulse frequency.

R 204: A. van Weel: Measurement of group-delay time in networks (Philips Res. Rep. 7, 467-473, 1952, No. 6).

A well-known principle for measuring group-delay time is based on measuring the phase relations of a low-frequency modulation on a high-frequency carrier, before and after passing an unknown network. The article describes a very sensitive phase measuring device where the phase angle to be measured is introduced in an oscillating circuit, thus influencing the frequency of oscillation. Under certain conditions the phase variations are proportional to the resulting frequency variations. Phase variations of 0.01 degree are measured without difficulty, permitting the measurement of group-delay time variations of 10^{-9} sec, using a modulation frequency of 30 kc/s.

R 205: Y. Haven and J. H. van Santen: On pre-exponential factors in formulae for ionic conductivity in solids (Philips Res. Rep. 7, 474-477, 1952, No. 6).

It is pointed out that, in formulae for equilibrium constants and rate constants, if the energy is linearly dependent on temperature, i.e. $E = E_0 - aT$, the pre-exponential factor i.e. coefficient preceding $\exp(-E_0/kT)$, does not contain the factor $\exp(a/k)$.

SUBJECT INDEX, VOLUMES 6–15

Figures in bold type indicate the volume number, and those in ordinary type the page number. Volumes 6 and 7 correspond to the years 1941 and 1942; volume 8 corresponds to 1946, and volume 15 to 1953-54. For articles published in volumes 1-5 the reader should refer to the index published in volume 10 (1948-49).

Aberrations, optical 9,301

Accelerators for elementary particles:
- Betatrons 11, 65
- Cascade generator (1.25 MeV) . 6, 46
- Cosmotron (Ferroxcube core) . 15, 73
- Cyclotron, q.v.
- Linear accelerators, q.v.

Acoustics of buildings:
- "ELA" studio 10,196
- Reverberation measurement 8, 82; 9,371
- Theatre, Utrecht 7, 9
- See also **Music**

Adhesive force of lacquers 8,147

Aerials:
- Sensitivity of — 6,302
- Frame —, "aerial" effect . . . 7, 65
- Frame —, combination with normal — 9, 55

"After-effect" phenomena in dielectrics and ferromagnetics 8, 57

Ageing of iron and steel:
- — after quenching 13,165
- Strain — 14, 60

Air-engine, see **Hot-air engine**

Alloys, see **Metals**

Amplifiers:
- D.C. amplification with magnetron 8,361
- See also: **Cathode-ray oscillographs**, **Frequency modulation**, **Negative feedback**, **Noise**, **Oscillators**, **Studio installations**, **Telephony**, **Television**, **Voltmeter, electronic**

Amplifier valves, see **Radio valves**

Analogue equipments, see **Automatic control of industrial processes** and **Rubber membrane analogue**

Analysis:
- Mercury in air 11, 91
- Lead in air 11,215
- Direct-reading polarograph . . 14,257
- Spectrochemical analysis . . . 12,337

Antennae, see **Aerials** . . . . .

Arc-welding, see **Welding** . . .

Audiometer 6,234

Auditorium acoustics, see **Acoustics of buildings** . . . . . .

Austempering of steel 6,279

Automatic change-over for emergency apparatus 8,310

Automatic control of industrial processes
- General considerations 12,221
- Electro-analogue 12,257,319

Automatic potentiometer (PR 1000, PR 2000) 15,189

Automobile headlamps 12,305

Bacteria, sterilization by U-V . . 12,111

Band-spread 6,265

Barium titanate, see **Titanates**

Barrier-layer photocells 8, 65

Barrier-layer rectifiers (small selenium —) 9,267

Batteries, emergency supply systems with 9,231; 11,253

Beacon, magnetron for 14, 87

Betatrons 11, 65

Bicycle dynamo 6,215

Black-out lighting 6,161

Blocking in triode oscillators . . 7,171

Brake for X-ray equipment . . . 11, 50

Broadcasting:
- Transmitters in Holland . . . 6, 1
- Studio installations, q.v.

Cable testing, see **Power cables**

Camera tubes for television (Image iconoscope) 13,119; 14,327

Candle, new international . . . 10,150

Car headlamps 12,305

Carrier telegraphy 8,206

Carrier telephony, see **Telephony**

Cascade generators:
- Stabilized, for electron microscope 10,135
- H.F. filament supply for — . . 11,123
- 1.25 MeV — for nuclear physics 6, 46
- — for TV receivers 25 kV . 10,125; 50 kV . 14, 21

Casting:
- Modern techniques (lost wax method) 15,133
- Vacuum — 15,114

Cathode-ray oscillographs:
- CRO with two push-pull amplifiers 9,202
- Small (GM 5655) 11,111
- "Stroboscopic" —, for high frequencies 12,52,73
- Electronic switch (GM 4580) . 9,340
- Pre-amplifier (GM 4574) . . . 15,169

Applications:
- Photographing cooling curves of hardening oils 9,147
- Checking of watches 9,317
- Use with polarographs . . . 14,257
- Stresses in Diesel engines . . 6, 94
- Testing of transistors . . . 13,254
- Recording with calibrated coordinates 12,283

Cathode-ray tubes:
- Efficiency of fluorescence . . . 15, 63
- Flying-spot scanner 15,233
- Focusing with Ferroxdure magnets 15,214
- Improvements in — for oscillographs 9,180
- Ion-trap magnet 15,258
- Projection television receiver (MW 6-2) 10, 97
- Large screen television projector (MW 13-16) 15, 27
- Short-length (bent neck) — . . 14,361
- Saturation of fluorescence . . . 12,120
- Steel-cone — 14,281

Cathodes:
- Pulsed operation of thermionic — 13,337
- L-cathode — q.v.
- Mercury —, ignition mechanism 8,346; 9,105
- Oxide —, theory of conduction 11,271
- Thoriated tungsten 14,226

Cavity resonators 8,149; 9, 73

"Centralix" 6,225

Ceramics:
- Manufacture 10,205
- Titanium dioxide 10,231
- Ferroxcube, q.v.
- Ferroxdure, q.v.

Characteristic plotter for receiver valves 12,283

Cinematography:
- Light sources for film studios 7,161
- Projector lamps 8, 72

Clandestine radio sets 8,337

Coal mines, illumination of . . . 10,334

Cold-cathode tubes for telephony (Z 500 T, Z 501 T) 15,265

Colour:
- Colour perception 9, 2
- Colour adaptation of the eye . . 9,257
- Checking colour deviations . . 6,186
- Colour triangle 12,137

"Compactix" 11,193

Condensers:
- Extruded — 13,145
- Ceramic — 10,231
- 10,000 kVA battery of — . . 7,182
- Vibrating —, for electrometer 7, 24; 10,338

Condenser microphones:
- Construction 9,330
- — circuit with low noise level 9,357

Conduction, electrical:
- in semiconductors with large negative temp. coefficient . 9,239
- in non-metals 9, 46
- in oxide cathodes 11,271

Conical disk springs 10, 61

Contact arc welding, see **Welding**

Contact therapy, X-ray apparatus for 8, 8; 13, 75

Contacts, electric:
- Checking of contact pressure . 6, 61
- Vibration of — springs 7,155

Convertor, vibrator, D.C.-A.C. . 6,342

Cooling of transmitting valves:
- Air — 9,171; 14,232
- Water — 10,239; 14,230

Copper, etching of 15,238

Copper wire, iron-free 8,315

Cosmotron, Ferroxcube core . . 15, 73

Counters:
- Decade counting tube (E1T) . 14,313
- **Geiger-Muller tube —**
  - Electron counter 6, 75
  - Battery-operated G-M counter 14,369

Cold-cathode tube — 15,265
See also **Geiger-Muller tubes**
Coupling of high freq. amplifier stages 8,125
Crystal structure, see **X-rays, Diffraction**

Crystals:
Quartz oscillator —, see **Quartz**
Lattice faults in — . . 15,105,246,286
Piëzo-electricity in — 11,151
Structure investigations of —, see **X-rays**

Cyclotron:
Theory of — 12, 65
Iron for electromagnet 10,246
Amsterdam —
General construction . . . 12,241
Oscillator and modulator . 12,247
Electromagnet 12,349
Further constructional details 14,263

Damping:
Internal — in iron and steel . 13,172
— of mechanical vibrations . 9,16,85
Darkroom lighting 11, 53
Daylight, effect of hours of — on Poinsettia and Populus . . . 14,175

Dazzle:
— of lighting fittings 12,200
— in lighted tunnels 10,299
D.C. amplification with magnetron ("magnetic diode") 8,361
Debye-Scherrer camera 10,157
De-ionization of thyratrons . . . 12,178
Delta modulation 13,237

Dielectric constant:
Materials with high — 10,231
Measurement at 3000 Mc/s . . 13, 61
Temperature coefficient of — and tan δ 9, 91
— of ferrites . . . . . 13,192 / 15, 79
Directing instrument for femur neck nailing 8,237
Directional hearing, see **Stereophonic sound**
Discharge tube for voltage stabilization 8,272
Discharge tubes, see **Gas-discharge tubes** and **Thyratrons** . . . . .
Disk springs 10, 61
Dislocations in metals . . 15, 246,286
Distortion in sound film due to non-perpendicular slits 6,110

Domestic lighting:
Fluorescent lamps 8,267
Illumination intensity 8,242
Dosimeter for X-rays 10,338
Drying lamp (infra-red) 9,249
Dynamo for bicycle 6,215
Dynamo pocket torch 8,225
Elastic constants of metals . . . 6,372
Electro-acoustic demonstration studio 10,196
Electro-analogue 12,257,319

Electrometers, vibrating condenser:
pH meter 7, 24
Direct reading — 10,338
Electron, 50 years history 9,225

Electron diffraction:
Investigation with electron microscope 9, 33
Texture investigations by — . 7,178
Electron emission, secondary, see **Secondary emission**
Electron lenses, measurement of field in magnetic — 15, 52

Electron microscope:
— with continuously variable magnification 9, 33
100 kV — . . . . . 9,179 / 12, 33
400 kV — 9,193
Stabilization of accelerating voltage for — 10,135
— investigation of plant diseases 14, 13
Electron paths in a space-charge field 14,336

Electron tubes
Review 14,129
X-radiography of — 12,207
See also: **Radio valves**
Transmitting valves
Output valves
Gas discharge tubes
Thyratrons
Switching tubes
Counters
Velocity modulation tubes
Magnetrons
Cathode-ray tubes
Television

Electronic conductivity of non-metallic materials 9, 46
Electronic switch (GM 4580) . . 9,340
Electrophoresis 12,293
Emergency apparatus, change-over switching 8,310
Epicyclic gearing 9,285
Equalization of telephone cables 7,184
Etching of copper 15,238
Expansion, thermal, of ferromagnetic alloys 10, 87
"Expressor" system for the transmission of music 11,281
Extrusion products, composite . 13,145

Eye:
Colour perception, see **Colour**
Night-blindness 9,211
Perception on the road at low illumination levels 9,149
Photography of the — 15,342
Visibility, q.v.
See also **Lighting**

Facsimile Transmission apparatus:
General considerations 10,189
Transmitter 10,257
Receiver 10,265
Transmissions 10,289
Synchronization 10,325
Femur neck nailing, apparatus for — 8,237
Fencing, electric 13,328

Ferrites:
Dimensional effect in — . . . 13,192 / 15, 79
Ferroxcube, q.v.
Crystal structure of — 9,186
Large negative temp. coefficient 9,239
Gyromagnetic phenomena . . 11,313
Ferro-electricity, see **Titanates**

Ferroxcube:
Development and properties . . 8,353
Magnetic & electric properties . 13,181
Measurement of properties . . 14,245

Applications
General review 13,301
Core for Cosmotron . . . 15, 73
Carrier telephony . . . 9,161 / 10,353
Television pulse generator 10,133
Transformer for aerials . . 9, 62
Deflecting circuits for television 10,307
See also **Ferrites**

Ferroxdure:
Development and properties . 13,194
Application in focussing of television picture-tubes . . . 15,214
Field strength in high-voltage installations 6,270

Filament lamps:
Car headlamps 12,305
Flickering 6,295
— for film projectors 8, 72
— for film studios 7,161
— for photography 6,259
Series connection of — 6,105
Improvement of light output (krypton lamp) 6,334
See also: **Lighting**
Film projector lamps 8, 72
Film studio lamps 7,161

Filters, electric:
— for carrier telephony 7,104
Influence of losses 7,138
Through-supergroup circuit — . 13,223
— with Ferroxcube cores . . 9,169 / 13,301
Filters, magnetic, for oil 6,169

Flash-bulbs, photographic:
Improved types of Photoflux bulbs 12,185
Specific light output of Photoflux bulbs (PF 3) 15,317

Flash-tubes, electronic:
— for stroboscopes 8, 25
— for Wilson cloud chambers . 10,178
— for ophthalmic photography 15,342

Flicker:
— of electric lights 6,295
— of television pictures 13, 55

Fluorescence:
— of solids, theory 6,349
Influence of temperature . . . 12, 6
Halophosphors for lamps . . . 13,346
Efficiency, in cathode-ray tubes 15, 63
Saturation, in TV picture-tubes 12,120
Zinc sulphide phosphors . . . 13, 90

Afterglow
Image converter 14,213
Flying-spot scanner 15,221
Line-converter 15,297

Applications
General review 9,215
Tubular fluorescent lamps . 6, 65
High pressure mercury lamps 13,109
X-ray screens 9,321
Painting with — paints . . . 11, 16

Fluorescent lamps:
Tubular luminescent lamps . . 6, 65
Starter switch for — 10,141
Control of light intensity . . . 12, 83

Use in mines 10,334
Use in homes 8,267
— in series with filament lamps 12,129
Equivalent circuit 15,161
See also **Lighting**
Flying-spot scanner 15,221,233
Focusing of picture-tubes (Ferroxdure) 15,214
Frequency analysis, recording apparatus 7, 50
Frequency-drift meter 12,193
Frequency modulation:
Fundamentals of — 8, 42
Comparison with amplitude modulation 8, 89
— detector tube (φ-detector, EQ 80) 11, 1
U.h.f. radio telephony using —
Transmitter 8,121
Receiver 8,193
Frequency stabilization (IGO) . 14,130

Galvanometer 7,113,128
Gas discharge lamps:
Flicker in — 6,295
Rapid modulation of — (for facsimile telegraphy) 10,265
Equivalent circuit of — . . . 15,161
Mercury vapour/filament lamp 7, 34
See also **Fluorescent lamps**
Gas discharge tubes:
Cold-cathode tubes (Z.500 T, Z.501 T) 15,265
Mercury cathode 8,346; 9,105
Stabilizer tube (85 A 1) . . . 8,272
See also **Thyratrons**
Gases, absorbtion in metals:
—, theory 6,365
—, experiments 7, 74
Gases, detection of poisonous . 8,341
Gearing, epicyclic 9,285
Geiger-Müller counters:
— for electrons 6, 75
— with halogen filling 13,282
— for X-rays 10, 1
— battery operated, as monitor 14,369
Applications
Use in tracer techniques . . 8,334
Spectrochemical analysis . . 11,215
See also **Counters**
Giorgi system 10,55,79
Glass:
Dielectric losses 13,361
Sintered glass 8, 2
Stress investigation in — . . . 9,277; 14,290
Structure of — 8,231; 13,293
"Iron" glass, for sealing to steel 14,281
Glass-metal seals:
Expansion coeffts. of alloys for — 10, 87
Sintered glass for — 8, 2
Radio valves 6,317
Steel picture-tube 14,281
Gramophones, for long-playing records 13,134
Grazing using electric fencing . . 13,328
Greenhouse with artificial light . 12, 1
Group-delay measurement in networks 15,307
Gyrator 11,321
Gyromagnetic phenomena in ferrites 11,313

Hardening of metals:
Isothermal hardening of steel 6,279
Various methods (e.g. oxidic —) 14,203
Investigation of cooling curves of hardening oils 9,147
Hearing aid, electronic 15, 37
High-frequency heating:
Inductive — 11,165
Capacitative — 11,232
High tension:
Electric fields in — equipment 6,270
— generators, see **Cascade generators** and **Television**
Rectifier valves for — 8,199
Hot-air engine:
Theory of — 8,129
Fundamentals of — 9, 87
Construction of a — 9,125
Measurement of temperature fluctuations in a — 13,104

IGO-frequency stabilization . . . 14,130
Illumination, see **Lighting**
Image converter for ultra-rapid photography (ME 1201) 14,213
Image iconoscope 13,119; 14,327
Image intensifier, X-ray 14, 33
Impact strength of iron and steel 11,303
Impedance measurement:
— with Lecher system 8, 16
— non-tuned Lecher system . 8,278
Indentation-meter for paints . . 13,352
Indicator for internal combustion engines (GM 3154, 4300, 4301) . 6, 22
"Infraphil" 8,177
Infra-red rays:
Drying with — 9,249
Medical applications
Irradiation apparatus . . . 6,202
"Infraphil" lamp. 8,177
Instability in triode oscillators . . 7,171
Intensification of X-ray images . 14, 33
Investment casting 15,133
Influenza vaccine, manufacture . 12,273
Ion accelerators, see **Accelerators for elementary particles**
Ion burn:
— in image iconoscope 14,327
— in television picture tubes . . 14,364; 15,258
Ion trap magnet 15,258
Ionosphere:
Radio investigation of — . . . 8,111
Ionispheric sounder 13,145
Iron content of copper wire . . 8,315
Iron for cyclotron magnet . . . 10,246
Isothermal hardening of steel . . 6,279
Isotopes:
Use as tracers 8,296
Tracer techniques 8,330

Jubilee review of research at Philips, 1891-1951 . . . . . 13, 1

Klystrons, see **Velocity modulation tubes**
Krypton, in incandescent lamps 6,334

Laboratory, history of Philips Research — 13, 3
Lacquers, adhesive force 8,147
Lamps:
See: **Filament lamps**
Flash-bulbs
Flash-tubes
Fluorescent lamps
Gas discharge lamps
Mercury-vapour lamps
Lighting
Lattice defects in crystals:
General review 15,105
— and plastic deformation 15,246,286
L-cathode:
Fundamentals and construction 11,341
— under pulsed operation . . . 13,337
Applications:
Image iconoscope 13,119
Magnetrons 14, 56
Lead estimation in air 11,215
Leak-tracing with Philips vacuum gauge (Penning gauge) 11,116
Lecher systems 6,240
Light distribution of projectors . 9,114
Light intensity, unit of 10,150
Light sources, see Lamps
Lighting:
Airfield — 6, 33
Blackout — 9,161
Blended-light lamps 7, 34
Coal mine — 10,334
Comfortable lighting 14, 69
Distribution and flux of projectors 9,114
Domestic lighting 8,267
Efficiencies of installations . . 7, 97
Film studio — 6,161
Fittings, brightness of 12,200
Intensity, unit of light . . . 10,150
Linear light sources
Calculation of installations . 6,147
Light distribution on a plane 12, 60
Measurement of light distribution of lamps 14,200
Office and home lighting . . . 8,242
Road —, physiological considerations 9,149
Theatre lighting, Utrecht . . 6, 1
Tunnel — 10,299
Luminescence, see **Fluorescence** and **Fluorescent lamps**
Linear electron-accelerators:
Fundamentals of — 14, 1
15 MeV — for Harwell 15, 1
Neutron spectrometer, use with — 15,325
Linear light sources, see **Fluorescent lamps** and **Lighting**
Line spectra, light sources for . 11,299
Loading coils with Ferroxcube cores 13,301
Losses:
Dielectric — in glass 13,360
— in networks 7,138
Magnetic — 8,353

Power factor measurements at 3000 Mc/s 13, 61
Variation with temperature coefficient of dielectric constant 9, 91
— in Ferroxcube . . . 13,181,301 14,245
After-effect — 8, 57
Lost-wax casting 15,133
Luminous intensity, unit 10,150

Magnadur, see **Ferroxdure**
Magnet steels 6, 8
"Magnetic diode" 8,361
Magnetic fields, measurement of . 15, 49

Magnetic materials:
Permanent —:
Magnet steels 6, 8
Non metallic — (Ferroxdure) 13,194
Soft — :
— for cyclotron 10,246
Non-metallic —, see **Ferrites** and **Ferroxcube**
Magnetostriction, effect on thermal expansion 10, 87

Magnetrons:
Theory and construction . . . 14, 44
— for beacons ($\lambda = 3$ cm) . . . 14, 87
— for D.C. amplification . . . 8,361 ("magnetic diode")

Magnets, permanent:
Dynamo — 8,225
Focussing — 15,214
Ion trap — 15,258
Manometer, Philips —, new design 11,116

Mechanical stresses:
— in glass 9,277 14,290
Investigation with strain gauges 6, 94 11, 23

Mechanical vibration:
— of contact springs 7,155
Investigation of — with strain gauges 11, 23
Vibration-free mountings . . 9, 16,85
Medical treatment, see **Therapy** and **X-rays**
Melting point, connection with thermal expansion 10, 87
Mercury-cathode rectifier valves 8,346 9,105
Mercury estimation in air 11, 91

Mercury-vapour lamps:
Low pressure fluorescent —, see **Fluorescent lamps**
— for photographic printing . 6,250
—/filament lamp combination - 7, 34
See also **Gas discharge lamps**

Metals:
Measurement of elastic constants 6,372
Etching of copper 15,238
Hardening of —, q.v.
Effect of impurities in iron and steel:
Internal friction 13,172
Impact strength 11,303
Ageing phenomena 13,165
Strain ageing 14, 60
Oxidation of heat resistant alloys 12,213
Plastic deformation and lattice defects 15,246,286
Preparation of very pure metals 11,241
Melting and casting in vacuo . 15,114
Metal detectors 15, 97
Metal leads through glass, see **Glass-metal seals**
Microphones, see **Condenser microphones**
Millivoltmeter, electronic, see **Voltmeter, electronic**
Mines, lighting of 10,334
Mirror galvanometer 7,114,128
Mixer circuit, self-oscillatory, with triodes 8,194
Mixer, diode as a — 6,285
M.K.S. system, see **Units**
Mobile transmitters, valves for . 12,157
Models, radio control of 15,281

Modulation:
Delta — 13,237
Frequency —, q.v.
Pulse —, multiplex — 11,133
See also **Telephony**
Modulators for carrier telephony . 7, 83
Moisture content, determination . 9, 13
Molybdenum, texture of 7,120
Monochromatic light sources . . 11,299
Moving coil meters, iron-free copper wire for 8,315
Multiplex pulse modulation . . 11,133
Multi-reflection klystron 8,257

Music:
Long-playing records 13,134
"Duplication" of concerts . . 10,169
Transmission by "Expressor" system 11,281
Stereophonic music in cinemas 11,129
Calibration of tuning forks . . 12,228

Negative feedback:
— for constant amplification . 9, 25
— in electrometer amplifier . . 10,338
— in telephone cable repeaters . 14,147

Networks:
Filters, q.v.
Influence of losses 7,138
Measurement of group delay in — 15,307
Unit function response 12,233
Neutron spectrometer, time-of-flight 15,325
Nickel-iron strip, texture 7, 45
Night-blindness, testing for . . . 9,211

Noise:
Reduction of — in sound film . 8, 97
Microphone circuit with low — level 9,357
Origins of — 6,129
At very high frequencies:
— in radio receivers 6,178
— in diodes 14,153
— in triodes 14,236
— in X-ray image intensifier 14, 33
Nuclear reactions, see **Accelerators for elementary particles**
Nuclear resonance method for measuring magnetic fields . . . 15, 55

Offices, illumination intensity . . 8,242
Oil filters, magnetic 6,169
Optical aberrations 9,301
"Oralix" 10,221

Oscillators:
— with constant output voltage 14,304
Pulse-synchronised — (IGO) . 14,130
Quartz oscillator crystals, q.v.
Travelling-wave — 11,221
***LC*-oscillators**
Operation 7, 40
Stability 7,171
Mechanism of synchronisation 14,292
Measuring oscillators
0.1-30 Mc/s — (GM 2882) . . 6,154
Television signal generator (GM 2887) 15,205
RC —, 20-250000 c/s (GM 2317) 15,240
Frequency drift of — . . . 12,193
Oscillograph, see **Cathode ray oscillograph**

Output valves:
— for low voltages (DL 41) . 10,346
See also **Radio valves**
Oxidation, acceleration in presence of foreign oxides (e.g. MoO_3) 12,213

Oxide Cathode:
Theory of conduction in — . . 11,271
Pulsed operation of — 13,337
Oxidic hardening of metals . . . 14,203

Painting with fluorescent paints . 11, 16

Paints and Lacquers:
Adhesive force meter 8,147
Identation meter 13,352
Testing for colour variation . . 6,186
Peak voltage meter 7, 20
Pentodes, see **Radio valves** and **Transmitting valves**
Permanent magnets, see **Magnets, and Magnetic materials**
Permeability of metals to gases, see **Gases, permeability of metals**
Phase difference, measurement up to 100 Mc/s 14,102

Philips-Miller sound recording system:
Photographic reproduction of sound track 9,65,289
Stereophonic recording 6, 80
"Philishave" dry-shaver 12, 27
*p*H measurement (GM 4491) . . 7, 24
Phosphorescence, see **Fluorescence**
Phosphors, see **Fluorescence**

Photo-electric cells:
Barrier layer - 8, 65
— with secondary emission, see **Photo-multipliers**
"Photoflux" lamp, see **Flashbulbs**

Photographic reproduction:
— of sound film 9, 65
Metal-diazonium system of — . 9,289
Mercury lamps for — . . . 6,205 13,323

Photography:
Darkroom lighting 11, 53
Flash-tubes, q.v.
Film studio lamps 7,161
Flash-bulbs, q.v.

Incandescent lamps for — . . 6,259
Positioning of light sources . . 12,145
Ultra rapid — (10^{-7} sec) with image converter 14,213

Photometry:
Unit of luminous intensity (new Candle) 10,150
— at low luminances 15,182

Photo-multipliers, applications of:
Facsimile transmitter 10,263
Flying-spot scanner 15,221
X-ray fluorography 10,109

Photo-printing and copying, mercury lamps for — 6,250 13,323

Photosynthesis 14,298

Pick-up tubes for television . 13,119 14,327

Piezo-electricity:
Materials exhibiting — 11,145
Detection of — 11,151

Plastic deformation of metals 15,246,286

Planetary lapping (grinding) of quartz oscillator crystals 12,166

Plant diseases, investigation with electron microscope 14, 13

Plastics:
— in the radio and electrical industries 11, 33
Polyvinyl chloride for wire insulation 12, 97

Poisonous gases, detection of . . 8,341

Polarograph, direct-reaching . . 14,257

Polyvinyl chloride ("Podur") for wire coverings 12, 97

Potatoes, storing in artificially-lighted cellars 10,318

Potential field in mercury cathode tube, plotting with electrolytic tank 8,346

Potentiometer, automatic (PR 1000, PR 2000) 15,189

Potter-Bucky diaphragm for X-rays 8,183

Power Cables:
Testing equipment for — . . . 7, 59
Localization of faults in — . . 7,113

Power valves, see **Output valves** and **Radio valves**

"Practix" 6,225

Precision casting 15,133

Pressure indicating equipment (GM 3154) 6, 22

Projection television, see **Television**

Projector lamps, film 8, 72

Projectors, light distribution of . 9,114

Proton resonance, measurement of magnetic fields with 15, 55

Pulse generator (carrier-wave supply) 8,137

Pulse modulation, multiplex . . . 11,133

Pulse transformer for control of fluorescent lamps 12, 90

Pyrometer, optical, for hardening of metals 6, 30

Quartz:
Piezo-electric property 11,145
— **oscillator crystals**
Cutting of — 11,323
Control of cutting angles by X-ray diffraction 11,351
Lapping and frequency adjustment 12,166
Quenching oils, cooling curves . . 9,147

Radar station at Ijmuiden . . . 14, 95

Radiation monitoring:
Battery-operated Geiger counter 14,369
X-ray dosimeter 10,338
— with Geiger-Müller counters q.v.
See also **Counters**

Radio-controlled models 15,281

Radio interference, sensitivity of aerials to 6,302

Radio-isotopes, see **Isotopes**

Radio propagation:
Investigation of ionosphere by — 8,111 13,152
— through the troposhere 15,148,175

Radio receivers:
Aerials, q.v.
Bandspread — 6,265
Clandestine — during wartime . 8,337
Diversity reception 9, 55
See also **Frequency modulation**
Oscillators
Noise
Telephony, u.h.f. links

Radio valves:
Diode as a frequency changer 6,285
Glass or metal envelopes? . . . 6,317
Characteristics, automatic plotting 12,283
Small — ("Rimlock) 8,289
Secondary emission in output valves (DL 1) 10,346

Special valves
Output valves for battery operation (DL 41) 10,346
Magnetron for D.C. amplification 8,361
Diode voltmeter for 3000 Mc/s (DA 50) 7,124
Frequency changer diode . 6,285
Frequency modulation detector (φ-detector, EQ 80) . . 11, 1
Triodes for decimetre waves (EC 80 and EC 81) . . . 11, 79
Subminiature valves 15,46,170,283

RC-oscillator, 20 c/s - 250 kc/s (GM 2317) 15,240

Receiving sets, see **Radio receivers**

Receiving valves, see **Radio valves**

Rectifiers:
De-ionization time of thyratrons 12,178
High voltage valves for X-ray diagnostics 8,199
Mercury cathode valves with capacitative igniter 8,346 9,105
Selenium —, small 9,267
Stabilization, voltage, q.v.
Voltage peaks in — 9,135
— for small telephone exchanges 6, 39
Thyratrons, q.v.
See also **Cascade generators**

Research Laboratory, history . . 13, 3

Resins, see **Plastics**

Resonant, circuits, u.h.f. 6,217

Reverberation, see **Acoustics of buildings**

Road lighting 9,149

"Rotalix" X-ray tubes 8, 33 13, 71

Roughness meter 14, 80

Rubber membrane analogue:
— for waveguides 11,156
— for electron paths in space-charge fields 14,336

Sampling:
Lot inspection by — 11,176
Sampling inspection plans . . 11,260
Practical application, methods 11,362

Schmidt optical system:
Principle of — 9,301
Correction plate, manufacture 9,349
Testing of spherical mirrors . 10,286
Applications:
Projection television . . . 10, 69 15, 27
X-ray fluorography 13,269

Secondary emission:
Undesirable — in output valves (DL 41) 10,346
— in television camera tubes . . 13,119
Application in switching tubes 13,49,82

Selenium photocells 8, 65

Selenium rectifiers (barrier-layer—) 9,267

Semiconductors:
Electronic conductivity of non-metallic materials 9, 46
— with large neg. temp. coefft 9,239
Controlled valency — 13, 90

Shaving, investigation of electrical 12, 25

Short circuit, safeguarding of permanent magnet dynamos against — 8,229

Short-waves, see **Ultra-high frequencies** 7,147

Short-wave therapy generator . 7,147

Signal generators:
Telephony — (10 kc/s) 8,251
Television — (GM 2887) 15,205
RC —, 20 c/s - 25 kc/s (GM 2317) 15,240
Receiver —, 0.1 - 30 Mc/s (GM 2882) 6,154

Signalling:
Carrier telephony 8,168
Electronic telephone exchanges 15,265

Sintered glass:
Properties and applications . . 8, 2
Application in transmitting valves — 10,273

Soap film models of electrical networks 7,138

Sound:
Audiometer; auditory perception 6,234
Electro-acoustic demonstration studio 10,196
Gramophones, long playing . . 13,134
Hearing aid, electronic 15, 37
Recording frequency analyzer . 7, 50
See also: **Acoustics of buildings**
Amplifiers
Condenser microphones
Music
Philips-Miller system
Stereophonic sound
Tape recorders

Sound film:
Photographic reproduction of — 9, 65
Photographic reproduction by metal-diazonium process . . 9,289
Noise reduction during reproduction 8, 97
Philips-Miller system, q.v.
Distortion due to non-perpendicular slit 6,110
Spectrochemical analysis:
Photographic methods of — . . 12,337
— with Geiger-Müller counter (lead estimation) 11,215
Spinels 9,185,239
Springs:
Disk — 10, 61
Helical — 9, 85
Instability of — 11,245
Stability of triode-oscillator . . . 7,171
Stabilization, voltage:
Battery preservation rectifier . 9,231
Battery-preservation rectifier with electronic voltage stabilization 11,253
Rectifier for telephone exchanges 6, 39
Stabilizer tube 8,272
300 V stabilized rectifier . . . 6, 54
5 kV stabilized rectifier 14,190
150 kV stabilized rectifier . . . 10,135
Starter switch for fluorescent lamps 10,141
Steel:
Magnet — 6, 8
— for cyclotron 10,246
See also **Metals** and **Hardening of metals**
Step-function response of amplifiers 6,193
Step-function response of networks 12,233
Stereophonic sound:
Stereophonic images 8, 51
Condenser microphone for — . 9,330
Perception of "depth" 6,359
— recording on Philips-Miller film 6, 80
Apparent elevation 9, 8
Applications:
Reproduction of concerts . . 10,169
Cinema interval music . . . 11,129
Strain gauges, measurements with . . . . . 6, 94 / 11, 23
Street lighting, see **Lighting**
Stress investigations, see **Mechanical Stresses**
Stroboscope (GM 5500) 8, 25
Structure investigations, X-ray, see **X-rays, diffraction**
Studio installations:
New principles of construction of — 6,139
Peak voltage meter for — . . . 7, 20
Electro-acoustic demonstration studio 10,196
Switching tubes:
Cold-cathode gas-filled — (Z.500T, Z.501T) 15,285
Secondary emission —:
Contact (make-break) . . . 13, 49
Multi-pole switches 13, 82
Switch-over, automatic, for emergency apparatus 8,310
Synchrocyclotron, see **Cyclotron**
Synchronization:
Facsimile equipment 10,325
Pulse — (IGO) 14,130
LC-oscillators 14,292
Television receivers . . . . 10,364 / 13,312 / 15,205

Tape recorders:
Description of apparatus . . . 14,181
Recording and reproducing process 15, 84
Telegraphy, carrier 8,206
Telephony:
Carrier-wave telegraphy . . . 8,206
Carrier-wave telephony:
General 6,324
Carrier supply (17 channel system) 8,137
Equalization of cables (17 channel system) 7,184
Filters (17 channel system) 7,104
Modulators 7, 83
Maintenance testing 8,249
Selenium rectifiers 9,267
Signalling in — 8,168
48-channel system:
Modulation system 9,161
Construction 10,353
— over-co-axial cables:
General problems, . . . 14,141
Through-supergroup filters 13,223
Delta-modulation 13,237
Electronic telephone exchange 15,265
Emergency apparatus, automatic change-over 8,310
Emergency supply systems . . 9,231 / 11,253
Ferroxcube cored coils 13,303
Rectifiers for small exchanges . 6, 39
Wiring for telephone exchanges 6, 85
U.h.f. radio links:
Eindhoven-Tilburg, with frequency modulation . . . 8,121 / 8,193
Installation for tropics . . . 6,120
Mobile installations 12,157
Multiplex pulse modulation 11,133
Television:
Camera tubes (image iconoscope) . . . . . 13,119 / 14,327
Flicker in — pictures 13, 55
Flywheel synchronization . . 13,312
Flying-spot scanner:
Design and construction . . 15,221
Cathode-ray tube 15,233
Group-delay measurement . . 15,307
Line-converter 15,297
Picture-tubes
Efficiency of fluorescence . . 15, 63
Focusing with Ferroxdure magnets 15,214
Bent-neck — 14,361
Ion-trap magnet 15,258
Projection — MW 6-2 . . 10, 97 / MW 13-16 . . 15, 28
— for flying spot scanner . . 15,233
Steel — 14,281
Saturation of fluorescence . . 12,120
Projection receiver (domestic)
Optical system 10, 69
Cathode ray tube 10, 97
High tension unit 10,125
Deflection circuits 10,307
Synchronization 10,364
Examination of mirror surfaces 10,286
Projection receiver (large screen)
General description 15, 27
High-tension generators . . 14, 21
Transmitters
9 kW — 7,129
100 kW output stage 14,345
Testing of TV receivers 15,205
West-European — network . . 14,358
Temperature measurement:
Pyrometer 6, 30
Rapidly fluctuating 13,104
Automatic potentiometer . . . 15,189
Texture of metals:
Aluminium mirrors 7,178
Cross-rolled molybdenum . . . 7,178
General discussion 7, 13
Nickel-iron strip 7, 45
Theatre, Utrecht:
Acoustics 7, 9
Lighting 7, 1
Therapy:
Infra-red — 6, 2
Infra-red — with Infraphil . . 8, 1
Short wave — 7, 1
X-ray —, q.v. . . . . .
Thyratrons:
— for control of light flux from fluorescent lamps 12, 83
De-ionization time of — . . . 12,178
PL. 10 13,331
Titanates:
Piezo-electric materials 11,145
Ferro-electricity 11,183
Titanium dioxide as dielectric . . 10,231
Tone generator for carrier telephony 8,251
Tracers, see **Isotopes**
Transient response of amplifiers . 6,193
Transistors, apparatus for testing 13,254
Transmitting valves:
Air-cooling of — 9,171
Discharge effect in — (Rocky Point effect) 6,208
Thoriated tungsten cathode — (TBL 12/100, TBW 12/100) . 14,226
Turbulent water-cooling . . . 10,239
U.h.f. —
100 Mc/s — (TB 2.5/300 and QB 2.5/250) 10,273
Mobile equipment — (QQE 06/40, QQC 04/15) 12,157
Technical problems of — . . 6,253
Velocity modulation valves, q.v.
Magnetrons, q.v.
Transmutations, see **Accelerators for elementary particles**
Travelling-wave oscillator . . . 11,221
Triode oscillator, see **Oscillators**
Triodes, see **Radio valves and Transmitting valves**
Troposphere, radio propagation through — . . . . . 15,148 / 15,175
Tuberculosis detection, see **X-rays, fluorography**
Tubes, see **Electron tubes**
Tulips:
Broken — (virus-diseased) . . 14, 13
Forcing of — 10,282
Tuning forks, calibration of — . . 12,228
Tunnels, illumination 10,299

Ultra-high frequencies:
Resonant circuits 6,217
Lecher systems 6,240
Cavity resonators 8,149; 9, 73
Waveguides (also q.v.) . . . 10,13,46
Noise at u.h.f., q.v.
Telephony, u.h.f. links, q.v.
Impedance measurement, q.v.

Valves for u.h.f.
Diode as mixer for decimetric waves 6,285
Diode voltmeter for 1000 Mc/s 7,124
Klystrons, see **Velocity modulation valves**
Receiver valves (EC 80 & EC 181) 11, 79
Transmitting valves, q.v.
Magnetrons, q.v.

Ultra-violet irradiation for sterilization of bacteria 12,111

Units:
Giorgi system
Origin of — 10, 55
— in electrical engineering . 10, 55
New international candle

Vacuum gauge, Philips 11,116
Vacuum melting apparatus . . 11,241; 15,114
Valency, controlled — 13, 90
Valve voltmeter, see **Voltmeter, electronic**

Valves:
Review 14,129
X-radiography of 12,207
See also: **Radio valves**
Transmitting valves
Output valves
Velocity modulation tubes
Magnetrons
Gas discharge tubes
Thyratrons
Switching tubes

Velocity-modulation tubes:
Principle and theory 8,214
Multi-reflection klystron . . . 8,257
Travelling-wave tube (helix resonator) 11,221
100-1000 watt klystrons . . . 13,209
Vibrating reed electrometer, see **Electrometers, vibrating condenser**
Vibration, see **Mechanical vibrations**
Vibrator 6,342

Viruses:
Investigation of — by electron microscope 14, 13
Manufacture of influenza vaccine 12,273

Visibility:
— when contours are blurred . . 11,333
— in tunnel lighting 10,299
— on roads, at low illumination 9,150
See also **Eye**
Vitamin D preparation 6,154
Voltage ratios, measurement of complex —, up to 100 Mc/s . . 14,102
Voltage regulation, see **Stabilization of voltage**

Voltmeter, electronic:
—, 1 kc/s-30 Mc/s (GM 6006) . 11,206
Calibration of — 14,308
Pre-amplifier for — (GM 4574) 15,169

Watches, timing with oscillograph 9,317

Waveguides:
Electromagnetic waves in — 10,13,46
— in linear accelerators . . . 14, 1; 15, 1
Rubber sheet model of — . . . 11,156
Measurement of standing wave ratio in — 12, 15

Welding:
Coating of — rods 10,114
Protection against dangerous voltages 15,199

Contact arc-welding
Fundamentals; electrodes types 8,161
Penetration and welding speed 8,304
Automatic — 13,247
Droplet size measurement . . 15,122
Mechanical properties (electrode Ph 55) 6, 97
Porosity of welds due to sulphur 7, 91
Porosity, general causes . . . 11,101
Low hydrogen welding rods . . 14, 96
Oxygen and nitrogen, rôle of 10, 26
Wilson cloud chamber, flash-tube for 10,178

Wire:
— for telephone exchanges . . 6, 85
Investigation into surface cracks 11, 12
"Podur" insulation of — . . . 12, 97
Wiring for telephone exchanges . 6, 85

X-rays
Collimator for — beams 13, 96

Diagnostics:
Fluoroscopy with enlarged image 8,321
Testing with phantoms . . . 11,291
H.T. rectifier valves 8,199
Portable apparatus ("Centralix", "Practix") 6,225
Lung exmination, see below, **Fluorography**
Dental apparatus ("Oralix") 10,221
Apparatus with exposure control (Medio D) 6, 12
Apparatus with exposure control (DX 4) 10, 37

Diffraction
Geiger counter spectrometer 10, 1
Debye-Scherrer camera . . 10,157
Orientation of quartz oscillator crystals 11,351
See also **Texture of metals**
Dosimeter 10,338
Directing instrument for treatment of femur fractures . . 8,237
Image intensifier 14, 33

Fluorography
— with lens-cameras 10,107
— with mirror-cameras . . 13,269
Lung examination 10,105; 11,

Fluorescent screens:
Blurring of image 9,321
Light emission 9,364
Scattered radiation, elimination of 8,183

Therapy
Contact — apparatus 8, 8
"Compactix" (200 kV) . . . 11,193
Midget (KT) apparatus for — 14,165
400 kV installation for — . . 8,105

Tubes
Contact therapy tubes . . 8, 9; 13, 75
Endotherapy tube 13, 77
Fine-focus tube 8,326
Midget (KT) tube 14,165
Oil-immersed hard glass tubes 6,309
Rotating anode — ("Rotalix") 8, 33
"Rotalix" O 55 for A.C. . . 13, 75

AUTHOR INDEX, VOLUMES 6–15

Figures in bold type indicate the volume number and those in ordinary type the page number. It should be noted that Dutch names containing a y may sometimes be found spelt with ij. For articles published in volumes 1-5 the reader should refer to the index published in volume 10 (1948-'49).

Addink, C. C. J.
Precise calibration of tuning forks **12**,228
Addink, N. W. H. and W. de Groot
Spectrochemical analysis **12**,337
Alink, R. J. H., C. J. Dippel and K. J. Keuning
The metal-diazonium system for photographie reproduction **9**,289
Alma, G., G. Diemer and H. Groendijk
A rubber membrane model for tracing electron paths in space charge fields **14**,336
— and F. Prakke
A new series of small radio valves **8**,289
Alons, L.
New developments in the gramophone world . . . **13**,134
Alphen, P. M. van and H. Rinia
The manufacture of correction plates for Schmidt optical systems **9**,349
Projection-television receiver,
I. The optical system for the projection **10**, 69
Amstel, J. J. A. Ploos van
Small selenium rectifiers **9**,267
Arbelet, G., J. C. van der Breggen and H. J. Lindenhovius
A millivoltmeter for the frequency range from 1000 to 30 × 10^6 c/s **11**,206
Arfin, B. and G. A. Espersen
A 3 cm-magnetron for beacons **14**, 87
Aschen, R. and P. Gaillard
A simple ionosphere sounder **13**,152
Aten, A. H. W. and F. A. Heyn
The use of iosotope as tracers **8**,296
The technique of investigations with radioactive and stable isotopes **8**,330
Bakker, C. J.
The causes of voltage and current fluctuations . . . **6**,129
Radio investigation of the ionosphere **8**,111
Bareford, C. F. and M. G. Kelliher
The 15 MeV linear electron accelerator for Harwell **15**, 1
A. Choice of parameters and general design . . . **15**, 2
B. Description of main components and circuits . **15**, 10
C. Testing and performance **15**, 22
Bast, G. H., D. Goedhart and J. F. Schouten
A 48-channel carrier telephone system,
I. The method of modulation **9**,161
II. Mechanical construction **10**,353
Bayfield, W. A., M. Berindei and A. Nemet
A diagnostic X-ray apparatus with exposure technique indication and overload protection **10**, 37
Beek, M. van de
A nine kilowatt experimental television transmitter **7**,129
Beljers, H. G.
A recording apparatus for the analysis of the frequency of rapidly varying sounds **7**, 50
Measurement of magnetic fields by the proton resonance method **15**, 55
— and J. L. Snoek
Gyromagnetic phenomena occurring with ferrites . **11**,313
Berg, J. ter
On the porosity of welds **7**, 91
Bergen, J. A. van, W. P. van den Blink and H. Bienfait
Automatic welding with contact electrodes . . . **13**,247
Bergmans, J. and H. A. E. Keitz
Determining the light distribution and luminous flux of projectors **9**,114
Berindei, M., W. A. Bayfield and A. Nemet
A diagnostic X-ray apparatus with exposure technique indication and overload protection **10**, 37
Bienfait, H., J. A. van Bergen, and W. P. van den Blink
Automatic welding with contact electrodes . . . **13**,247
Bierman, A. and H. A. Oele
Betratons with and without iron yoke **11**, 65
Biermasz, A. L. and H. Hoekstra
The measurement of changes in length with the aid of strain gauges **11**, 23
Blasberg, E. and A. de Groot
Metal-detectors **15**, 97
Bleeksma, J., G. Kloos and H. J. di Giovanni
X-ray spectrometer with Geiger counter for measuring powder diffraction patterns **10**, 1
Blink, W. P. van den, H. Bienfait and J. A. van Bergen
Automatic welding with contact electrodes . . . **13**,247
Bloemen, J. P. de Visser van
Stereophonic music in the cinema **11**,129
Blok, L. and H. J. Kösten
An apparatus for the investigation of sharpness of hearing (audiometer) **6**,234
Blom, P.
An electronic hearing aid **15**, 37
Boeke, J.
Electrical detection of traces of poisonous gases in the atmosphere **8**,341
A new electrical method for determining moisture content **9**, 13
Boelens, W. W.
An instrument for recording the frequency drift of an oscillator **12**,193
Boer, H. J. de and A. van Weel
An instrument for measuring group delay **15**,307
Boer, J. B. de and D. Vermeulen
The permissible brightness of lamp fittings **12**,200
Motorcar headlights **12**,305
Boer, K. de
Stereophonic recording on Philips-Miller film **6**, 80
The formation of stereophonic images **8**, 51
A remarkable phenomenon with stereophonic sound reproduction **9**, 8
— and A. Th. van Urk
Some particulars of directional hearing **6**,359
Boldingh, W. Hondius
The testing of power cables with D.C. voltages . . . **7**, 59
The localization of cable flaws **7**,113
Fluorography with the aid of a mirror system . . . **13**,269
— and W. J. Oosterkamp
A 400 kilovolt installation for X-ray therapy . . . **8**,105
Boosman, H. B. R.
The new broadcasting transmitters in the Netherlands **6**, 1
Botden, P. J. M. and B. Combée
Special X-ray tubes **13**, 71
—, B. Combée and J. Houtman
An experimental X-ray apparatus with midget X-ray tube **14**,165
Bouma, P. J.
The concept of brightness in connection with blackout problems **6**,161
The flickering of sources of electric light **6**,295
Perception on the road when visibility is low . . **9**,149
— and A. A. Kruithof
Colour stimulus and colour sensation **9**, 2
Chromatic adaptation of the eye **9**,257
Bouwers, A. and P. G. Cath
The maximum electric field strength for several simple electrode configurations **6**,270
— and F. A. Heyn
An apparatus for the transmutation of atomic nuclei **6**, 46
A simple apparatus for counting electrons **6**, 75
Breggen, J. C. van der, H. J. Lindenhovius and G. Arbelet
A millivoltmeter for the frequency range from 1000 to 30 × 10^6 c/s **11**,206
Brekoo, H. P. J. and Tj. Douma
Heating the filaments of valves in a cascade generator by means of high-frequency current . . . **11**,123

Bremmer, H.
The troposphere as a medium for the propagation of radio waves
I 15,148
II 15,175
Brey, H. de and H. Rinia
An improved method for the air-cooling of transmitting valves 9,171
— H. Rinia and F. L. van Weenen
Fundamentals for the development of the Philips air engine 9, 97
Bril, A., J. de Gier and H. A. Klasens
A cathode-ray tube for flying-spot scanning 15,233
— and H. A. Klasens
The efficiency of fluorescence in cathode-ray tubes . 15, 63
— and F. A. Kröger
Saturation of fluorescence in television tubes 12,120
Brockman, F. G. and M. W. Louwerse
The Ferroxcube transformer core used in the Brookhaven Cosmotron 15, 73
Bruin, S. L. de
An apparatus for stroboscopic observation . . . 8, 25
Bruining, H. and J. C. Francken
New developments in the image iconoscope 14,327
—, J. C. Francken and P. Schagen
The image iconoscope, a camera tube for television 13, 119
Bruinsma, A. H.
Radio-controlled models 15,281
Bueren, H. G. van
Lattice imperfections and plastic deformation in metals:
I. Nature and characteristics of lattice imperfections, notably dislocations 15,246
II. Behaviour of lattice imperfections during deformation 15,286
Bügel, R. D. and E. J. W. Verwey
Ceramic materials with a high dielectric constant 10,231
Burger, G. C. E.
Phantom tests with X-rays 11,291
—, B. Combée and J. H. van der Tuuk
X-ray fluoroscopy with enlarged image 8,321
Burgerjon, J. J. and F. A. Heyn
The synchrocyclotron at Amsterdam,
IV. Details of construction and ancillary equipment 14,263
Burgt, C. M. van der, M. Gevers and H. P. J. Wijn
Measuring methods for some properties of Ferroxcube materials 14,245
Carpentier, E. E.
A cathode-ray oscillograph with two push-pull amplifiers 9,202
An electronic switch with variable commutating frequency 9,340
An easily portable cathode-ray oscillograph 11,111
Cassee, E.
A preservation rectifier with electronically stabilized charging voltage 11,253
Casimir, H. B. G.
A magnetron for D.C. voltage amplification . . . 8,361
Cath, P. G. and A. Bouwers
The maximum electric field strength for several simple electrode configurations 6,270
Cayzac, J. and P. Périlhou
A stabilized extra-high tension rectifier for 5000 V, 50 mA 14,190
Chabot, J. J. M. Taüdin, T. G. Willemze and H. C. Hamaker
The practical application of sampling inspection plans and tables 11,362
Chippendale, R. A., and J. A. Jenkins
High speed photography by means of the image converter 14,213
Cisney, E. and W. Parrish
An improved X-ray diffraction camera 10,157
Coeterier, F.
The multi-reflection tube, a new oscillator for very short waves 8,257
Combée, B. and P. J. M. Botden
Special X-ray tubes 13, 71
—, G. C. E. Burger and J. H. van der Tuuk
X-ray fluoroscopy with enlarged image 8,321
—, and J. Fransen
An X-ray apparatus for therapeutic treatments . 11,193
—, J. Houtman and P. J. M. Botden
An experimental X-ray apparatus with midget X-ray tube 14,165
Constable, J. M.
An automatic braking device for X-ray apparatus . 11, 50
Cornelius, P.
The sensitivity of aerials to local interferences . . . 6,302
The aerial effect in receiving sets with loop aerials . 7, 65
The rationalized Giorgi system with absolute volt and ampere as applied in electrical engineering . . 10, 79
— and J. van Slooten
Installations for improved broadcast reception . . 9, 55
Custers, J. F. H.
Considerations on the texture of metals 7, 13
The texture of nickel-iron strip 7, 45
The texture of cross-rolled molybdenum 7,120
The investigation of texture with electron rays . . 7,178
Daams, J. and J. Voogd
Inactivation of bacteria by ultra-violet radiation . 12,111
Dammers, B. G., P. D. van der Knaap and A. G. W. Uitjens
The electrical recording of diagrams with a calibrated system of coordinates 12,283
Defize, L. F. and P. C. van der Willigen
The determination of droplet size in arc welding by high-speed cinematography 15,122
Dell, H. A. and C. H. R. Gentry
A direct reading precision polarograph 14,257
Derksen, J. C. and M. Stel
Plastics and their application in the electrotechnical industry 11, 33
Diemer, G., H. Groendijk and G. A. Alma
A rubber membrane model for tracing electron paths in space charge fields 14,336
—, and K. S. Knol
A model for studying electromagnetic waves in rectangular wave guides 11,156
The noise of electronic valves at very high frequencies,
I. The diode 14,153
II. The triode 14,236
Dijkstra, L. J. and J. D. Fast
Internal friction in iron and steel 13,172
Dippel, C. J., R. J. H. Alink and K. J. Keuning
The metal-diazonium system for photographic reproduction 9,289
— and K. J. Keuning
Problems in photographic reproduction, in particular of sound-films 9, 65
Domburg, J. and W. Six
A cold cathode gas-discharge tube as a switching element in automatic telephony 15,265
Dorgelo, E. G.
Several technical problems in the development of a new series of transmitting valves 6,253
Sintered glass 8, 2
Glass transmitting valves of high efficiency in the 100 Mc/s range 10,273
High-power transmitting valves with thoriated tungsten cathodes. 14,226
— and P. Zijlstra
Two transmitting valves for use in mobile installations 12,157
Dorsman, C.
A *p*H-meter with a very high input resistance . . 7, 24
Dorsten, A. C. van
Stabilization of the accelerating voltage in an electron microscope 10,135
—, and A. J. J. Franken
Measurement of the field on the axis of magnetic electron lenses 15, 52
—, H. Nieuwdorp and A. Verhoeff
The Philips 100 kV electron microscope 12, 33
—, W. J. Oosterkamp and J. B. Le Poole
An experimental electron microscope for 400 kV . 9,193
Douma, Tj.
Voltage impulses in rectifiers 9,135
—, and H. P. J. Brekoo
Heating the filaments of valves in a cascade generator by means of high-frequency current 11,123

Druyvesteyn, M. J.
The determination of the elastic constants of metals 6,372
du Pré, F. K. and M. Gevers
A remarkable property of solid dielectrics 9, 91
—, and H. Rinia
Air engines 8,129
—, and J. L. Snoek
Several after-effect phenomena and related losses in alternating fields 8, 57
Elenbaas, W. and T. Holmes
An instant-starting fluorescent lamp in series with an incandescent lamp 12,129
—, and K. R. Labberté
High-pressure mercury vapour lamps for photocopying processes 13,323
—, K. R. Labberté and J. L. Ouweltjes
A new high-pressure mercury lamp with fluorescent bulb 13,109
—, and R. Riemens
Light sources for line spectra 11,299
Embden, H. J. Meerkamp van
Modern casting techniques 15,133
— and B. Jonas
New kinds of steel of high magnetic power 6, 8
— and J. A. M. Smelt, O. L. van Steenis, J. G. W. Mulder, J. de Gier and Th. Hagenberg
A steel picture-tube for television reception 14,281
Ensing, L. and H. J. J. van Eyndhoven
An oscillator with constant output voltage 14,304
—, and J. M. L. Janssen
The electro-analogue, an apparatus for studying regulating systems,
I. Components and functions 12,257
II. The electrical execution 12,319
Espersen, G. A. and B. Arfin
A 3 cm-magnetron for beacons 14, 87
Eyndhoven, H. J. J. van and L. Ensing
An oscillator with constant output voltage 14,304
Fast, J. D.
The permeability of metal walls for gases 6,365
Experiments on the permeation of gases through metal walls 7, 74
The part played by oxygen and nitrogen in arc-welding 10, 26
The function of the coating of welding rods 10,114
Causes of porosity in welds 11,101
Apparatus for preparation of metals with an exactly known content of impurities 11,241
Investigations into the impact strength of iron and steel 11,303
Ageing phenomena in iron and steel after rapid cooling 13,165
Strain ageing in iron and steel 14, 60
Low-hydrogen welding rods 14, 96
—, and L. J. Dijkstra
Internal friction in iron and steel 13,172
—, A. I. Luteijn and E. Overbosch
Preparation and casting of metals and alloys under high vacuum 15,114
Feiner, H. and H. N. Hansen
Coaxial cable as a transmission medium for carrier telephony 14,141
—, and J. Hoekstra
Switchboard wire for telephone installations 6, 85
Francken, J. C. and H. Bruining
New developments in the image iconoscope 14,327
—, P. Schagen and H. Bruining
The image iconoscope, a camera tube for television 13,119
Franken, A. J. J. and A. C. van Dorsten
Measurement of the field on the axis of magnetic electron lenses 15, 52
Fransen, J.
A miniature X-ray apparatus for dentists 10,221
—, and B. Combée
An X-ray apparatus for therapeutic treatments . . 11,193
— and J. M. Ledeboer
Generators for short-wave therapy 7,147
Frederik, W. S.
Testing for nyctalopia (night-blindness) 9,211
Fremery, F. de and J. W. G. Wenke
New principles of construction for the electro-acoustic installation of studios 6,139
The measurement of peak voltages in a studio installation 7, 20
Fry, D. W.
The linear electron accelerator 14, 1
Funke, J. and P. J. Oranje
The development of blended-light lamps 7, 34
Gaillard, P. and R. Aschen
A simple ionosphere sounder 13,152
Geel, W. Ch. van
Blocking-layer photocells 8, 65
Geiss, W.
Improvements in the efficiency of electric incandescent lamps 6,334
Gelder, Z. van and J. L. H. Jonker
New electronic tubes employed as switches in communication engineering
I. Contact tubes 13, 49
II. Switch tubes 13, 82
Gentry, C. H. R. and H. A. Dell
A direct reading precision polarograph 14,257
Gevers, M.
Measuring the dielectric constant and the loss angle of solids at 3000 Mc/s 13, 61
— and F. K. du Pré
A remarkable property of solid dielectrics 9, 91
—, H. P. J. Wijn and C. M. van der Burgt
Measuring methods for some properties of Ferroxcube materials 14,245
Gier, J. de
Projection-television receiver,
II. The cathode-ray tube 10, 97
—, Th. Hagenberg, H. J. Meerkamp van Embden, J. A. M. Smelt, O. L. van Steenis and J. G. W. Mulder
A steel picture-tube for television reception . . . 14,281
—, H. A. Klasens and A. Bril
A cathode-ray tube for flying-spot scanning 15,233
— and A. P. van Rooy
Improvements in the construction of cathode-ray tubes 9,180
Giovanni, H. J. di, J. Bleeksma and G. Kloos
X-ray spectrometer with Geiger counter for measuring powder diffraction patterns 10, 1
—, W. Kes and K. Lowitzsch
A transportable X-ray apparatus for mass chest survey 10,105
Goedhart, D., G. H. Bast and J. F. Schouten
A 48-channel carrier telephone system,
I. The method of modulation 9,161
II. Mechanical construction 10,353
— and G. Hepp
Carrier supply in an installation for carrier telephony 8,137
— and J. de Jong
Carrier-wave telephony 6,325
Gorter, E. W., G. W. Oosterhout, J. J. Went and G. W. Rathenau
Ferroxdure, a class of new permanent magnet materials 13,194
— and J. J. Went
The magnetic and electrical properties of Ferroxcube materials 13,181
Gradstein, S.
Positioning of the sources of light when photographing with artificial light 12,145
— and J. L. H. Jonker
Fluorescent pigments as an artistic medium 11, 16
Greefkes, J. A., J. F. Schouten and F. de Jager
Delta modulation, a new modulating system for telecommunication 13,237
Groendijk, H., G. A. Alma and G. Diemer
A rubber membrane model for tracing electron paths in space charge fields 14,336
Groot, A. de and E. Blasberg
Metal-detectors 15, 97
Groot, F. A. de
Signalling in carrier telephony 8,168
— and P. J. den Haan
Modulators for carrier-telephony 7, 83

Groot, W. de
Fifty years of electrons 9,225
Optical aberrations in lens and mirror systems . . . 9,301
The origin of the Giorgi system of electrical units . 10, 55
The new Candle 10,150
Cyclotron and synchrocyclotron 12, 65
Philips diamond jubilee: Scientific research of Philips' Industries, 1891-1951 . . . . . 13, 1
Photometry at low luminance levels 15,182
— and N. W. H. Addink
Spectrochemical analysis 12,337
— and H. A. Klasens
The light-emission of X-ray screens 9,321
— and F. A. Kröger
The influence of temperature on the fluorescence of solids 12, 6
— and A. A. Kruithof
The colour triangle 12,137
Haan, P. J. den and F. A. de Groot
Modulator for carrier-telephony 7, 83
Haantjes, J.
Judging an amplifier by means of the transient characteristic 6,193
— and F. Kerkhof
Projection-television receiver,
IV. The circuits for deflecting the electron beam . . 10,307
V. The synchronization 10,364
— and C. J. van Loon
A large-screen television projector 15, 27
— and Th. G. Schut
A line converter for the international exchange of television programmes 15,297
— and F. W. de Vrijer
Flicker in television pictures 13, 55
Haaijman, P. W., E. L. Heilman and E. J. W. Verwey
On the crystalline structure of ferrites and analogous metal oxides 9,185
— and J. L. H. Jonker
The wire capacitor and other composite drawn products 13,145
— F. C. Romeijn and E. J. W. Verwey
Semi-conductors with large negative temperature coefficient of resistance 9,239
Hagenberg, Th., H. J. Meerkamp van Embden, J. A. M. Smelt, O. L. van Steenis, J. G. W. Mulder and J. de Gier
A steel picture-tube for television reception 14,281
Hagendoorn, P. J. and M. F. Reynst
The recording of diagrams with the electrical pressure indicator 6, 22
Halbertsma, N. A. and J. A. M. van Liempt
Incandescent filament lamps for series-connection . 6,105
Lamps for use in photography 6,259
Hamaker, H. C.
Lot inspection by sampling 11,176
The theory of sampling inspection plans 11,261
—, J. J. M. Taudin Chabot, and F. G. Willemze
The practical application of sampling inspection plans and tables 11,362
Hammerton, J. C., E. Wolfendale and P. E. Trier
A time-of-flight neutron spectrometer 15,325
Hansen, H. N. and H. Feiner
Coaxial cable as a transmission medium for carrier telephony 14,141
Haringx, J. A.
Magnetic oil filters 6,169
The vibration of contact springs 7,155
Vibration-free mountings with auxiliary mass . . 9, 16
The practical construction of vibration-free mountings with auxiliary mass 9, 85
Conical disc springs 10, 61
Instability of springs 11,245
Hazeu, H. A. G.
Small apparatus for medical X-ray examination . 6,225
— and J. M. Ledeboer
A universal apparatus for X-ray diagnosis 6, 12
— J. M. Ledeboer and J. H. van der Tuuk
An X-ray apparatus for contact therapy 8, 8
Hehenkamp, Th.
A rapid-action starter switch for fluorescent lamps . 10,141
Heilman, E. L., P. W. Haaijman and E. J. W. Verwey
On the crystalline structure of ferrites and analogous metal oxides 9,185
Helmer, A. M. C. and N. Warmoltz
A flash lamp for illuminating vapour tracks in the Wilson cloud chamber 10,178
Hengel, J. van and W. J. Oosterkamp
A direct-reading dynamic electrometer 10,338
Hepp, G.
Automatic change-over to an emergency apparatus in a communication system 8,310
A battery-operated Geiger-Müller counter 14,369
— and D. Goedhart
Carrier supply in an installation for carrier telephony 8,137
Hertzberger, E. and A. J. Klein
The manufacture of virus vaccine against influenza . 12,273
Hess, K. W.
Measuring the de-ionization time of gas-filled diodes and triodes 12,178
— and F. H. de Jong
Controlling the luminous intensity of fluorescent lamps with the aid of relay valves 12, 83
Heyboer, J. P.
A discharge phenomenon in large transmitter valves 6,208
Heyn, F. A.
The synchrocyclotron at Amsterdam,
I. General description of the installation 12,241
II. The oscillator and the modulator 12,241
III. The electromagnet 12,349
— and A. H. W. Aten
The use of isotopes as tracers 8,296
The technique of investigations with radioactive and stable isotopes 8,330
— and A. Bouwers
An apparatus for the transmutation of atomic nuclei 6, 46
A simple apparatus for counting electrons 6, 75
— and J. J. Burgerjon
The synchrocyclotron at Amsterdam,
IV. Details of construction and ancillary equipment 14,263
Hoekstra, J. and A. L. Biermasz
The measurement of changes in length with the aid of strain gauges 11, 23
— and H. Feiner
Switchboard wire for telephone installations 6, 85
— and J. A. W. van Laar
An indentation meter for paint 13,352
Hofweegen, J. M. van
The measurement of impedances particularly on decimeter waves 8, 16
Impedance measurements with a non-tuned Lecher system 8,278
Holmes, T. and W. Elenbaas
An instant-starting fluorescent lamp in series with an incandescent lamp 12,129
Holthoon, J. van and F. G. Peuscher
A pre-amplifier for use with electronic voltmeters and oscilloscopes 15,169
Hondius Boldingh, W. see **Boldingh, W. Hondius**
Houtman, J., P. J. M. Botden and B. Combée
An experimental X-ray apparatus with midget X-ray tube 14,165
Hugenholtz, E. H.
The impulse-governed oscillator, a system for frequency stabilization 14,130
Ingen Schenau, B. W. van, see **Schenau, B. W. van Ingen**
Iperen, B. B. van
The helix as resonator for generating ultra high frequencies 11,221
Velocity-modulation valves for 100-1000 watts continuous output 13,209
Jager, F. de, J. A. Greefkes and J. F. Schouten
Delta modulation, a new modulating system for telecommunication 13,237
Jansen, C. G. J., R. Loosjes and H. J. Vink
Thermionic emitters under pulsed operation . . . 13,337
Jansen, M. J., R. Loosjes and H. J. Lemmens
A new thermionic cathode for heavy loads 11,341
Janssen, J. M. L.
An experimental "stroboscopic" oscilloscope for frequencies up to about 50 Mc/s,
I. Fundamentals 12, 52

— and L. Ensing
The electro-analogue, an apparatus for studying regulating systems,
I. Components and functions 12,257
II. The electrical execution 12,319
— and A. J. Michels
An experimental "stroboscopic" oscilloscope for frequencies up to about 50 Mc/s,
II. Electrical build-up 12, 73
Jenkins, J. A. and R. A. Chippendale
High speed photography by means of the image converter 14,213
Jochems, P. J. W. and F. H. Stieltjes
Apparatus for testing transistors 13,254
Jonas, B. and H. J. Meerkamp van Embden
New kinds of steel of high magnetic power 6, 8
Jong, F. H. and K. W. Hess
Controlling the luminous intensity of fluorescent lamps with the aid of relay valves 12, 83
— and D. W. van Rheenen
An apparatus for protection against dangerous voltages in welding equipment 15,199
— and L. W. Roosendaal
Electric fencing of grazing land 13,328
Jong, J. de
Maintenance measurements on carrier telephony equipment 8,249
— and D. Goedhart
Carrier-wave telephony 6,325
Jong, J. J. de
A remarkable etching of copper 15,238
Jonker, G. H. and J. H. van Santen
The ferro-electricity of titanates 11,183
Jonker, J. L. H.
Secondary emission in output valves 10,346
Electronic tubes 14,117
A short length direct-view picture-tube 14,361
— and Z. van Gelder
New electronic tubes employed as switches in communication engineering,
I. Contact tubes 13, 49
II. Switch tubes 13, 82
— and S. Gradstein
Fluorescent pigments as an artistic medium . . . 11, 16
— and P. W. Haaijman
The wire capacitor and other composite drawn products 13,145
— and A. J. W. M. van Overbeek
The "φ-detector", a detector valve for frequency modulation 11, 1
—, A. J. W. M. van Overbeek and K. Rodenhuis
A decade counter tube for high counting rates . . . 14,313
Jürgens, B. F.
Measurement of magnetic fields based on the generator principle 15, 49
Jurriaanse, T.
A voltage stabilizing tube for very constant voltage 8,272
Kalff, L. C.
The illumination of the new municipal theatre in Utrecht 7, 1
Ways to comfortable lighting 14, 69
— and J. Voogd
Living room lighting with tubular fluorescent lamps 8,267
Keitz, H. A. E.
A method of measuring in the investigation of bicycle dynamos 6,215
— and J. Bergmans
Determing the light distribution and luminous flux of projectors 9,114
Kelliher, M. G. and C. F. Bareford
The 15 MeV linear electron accelerator for Harwell 15, 1
A. Choice of parameters and general design 15, 2
B. Description of main components and circuits . . 15, 10
C. Testing and performance 15, 22
Kerkhof, F. and J. Haantjes
Projection-television receiver,
IV. The circuits for deflecting the electron beam . 10,307
V. The synchronization 10,364
— and G. J. Siezen
Projection-television receiver,
III. The 25 kV anode voltage supply unit . . . 10,125
Kes, W., H. J. di Giovanni and K. Lowitzsch
A transportable X-ray apparatus for mass chest survey 10,105
Keuning, K. J., R. J. H. Alink and C. J. Dippel
The metal-diazonium system for photographic reproduction 9,289
— and C. J. Dippel
Problems in photographic reproduction, in particular of sound-films 9, 65
Klasens, H. A.
The blurring of X-ray images 9,364
— and A. Bril
The efficiency of fluorescence in cathode-ray tubes . 15, 63
— A. Bril and J. de Gier
A cathode-ray tube for flying-spot scanning 15,233
— and W. de Groot
The light-emission of X-ray screens 9,321
Klein, A. J. and E. Hertzberger
The manufacture of virus vaccine against influenza . 12,273
Kleis, D., H. Rinia and M. van Tol
Experimental equipment for high-speed facsimile transmission,
I. General 10,189
— F. C. W. Slooff and J. M. Unk
Experimental equipment for high-speed facsimile transmission,
II. Details of the transmitter 10,257
— and M. van Tol
Experimental equipment for high-speed facsimile transmission:
IV. Transmission of the signals 10,289
V. Synchronization of transmitter and receiver . . 10,325
Klinkhamer, H. A. W.
A rectifier for small telephone exchanges 6, 39
Emergency supply systems with accumulator batteries 9,231
Klinkhamer, J. F.
A through supergroup filter for carrier telephone systems on coaxial cables 13,223
Kloos, G., J. Bleeksma and H. J. di Giovanni
X-ray spectrometer with Geiger counter for measuring powder diffraction patterns 10, 1
Klute, J. W. and J. F. Schouten
The influence of losses on the properties of electrical networks 7,138
Knaap, P. D. van der, A. G. W. Uitjens and B. G. Dammers
The electrical recording of diagrams with a calibrated system of coordinates 12,283
Knol, K. S. and G. Diemer
A model for studying electromagnetic waves in rectangular wave guides 11,156
The noise of electronic valves at very high frequencies,
I. The diode 14,153
II. The triode 14,236
— and M. J. O. Strutt
A diode for the measurement of voltages on decimetre waves 7,124
Koole, P.
Measurement of the adhesive force of lacquers 8,147
— and J. J. Went
A method for the control of colour deviations . . 6,186
Koppius, O. G.
An application of Geiger counter tubes for spectrochemical analysis 11,215
Köster, H. J. and L. Blok
An apparatus for the investigation of sharpness of hearing (audiometer) 6,234
Kröger, F. A.
Luminescent substances 6,349
Applications of luminescent substances 9,215
— and A. Bril
Saturation of fluorescence in television tubes . . . 12,120
— and W. de Groot
The influence of temperature on the fluorescence of solids 12, 6
— and E. J. W. Verwey
New views on oxidic semi-conductors and zinc-sulphide phosphors 13, 90
Kruithof, A. A.
Tubular luminescent lamps for general illumination 6, 65

— and P. J. Bouma
Colour stimulus and colour sensation 9, 2
Chromatic adaptation of the eye 9,257
— and W. de Groot
The colour triangle 12,137
— and A. M. Kruithof
Basic principles for the formulation of illumination standards 10,214
— and H. Zijl
Illumination intensity in offices and homes 8,242
Kruithof, A. M.
On the illumination of traffic tunnels 10,299
Perception of contrasts when the contours of details are blurred 11,333
— and A. A. Kruithof
Basic principles for the formulation of illumination standards 10,214
Kuperus, J.
On the construction of vibrators for radio-sets . . . 6,342
Laar, J. A. W. van and J. Hoekstra
An indentation meter for paint 13,352
Labberté, K. R. and W. Elenbaas
High-pressure mercury vapour lamps for photo-copying processes 13,323
—, J. L. Ouweltjes and W. Elenbaas
A new high-pressure mercury lamp with fluorescent bulb 13,109
Ledeboer, J. M. and J. Fransen
Generators for short-wave therapy 7,147
— and H. A. G. Hazeu
A universal apparatus for X-ray diagnosis 6, 12
— H. A. G. Hazeu and J. H. van der Tuuk
An X-ray apparatus for contact therapy 8, 8
Lely, J. A. and T. W. van Rijssel
Collimating X-rays in beams of very small divergence and high intensity 13, 96
Lemmens, H. J., M. J. Jansen and R. Loosjes
A new thermionic cathode for heavy loads 11,341
Levenbach, G. J. and F. de Fremery
The equalization of telephone cables 7,184
Liempt, J. A. M. van and N. A. Halbertsma
Incandescent filament lamps for series connection 6,105
Lamps for use in photography 6,259
Lindenhovius, H. J., G. Arbelet and J. C. van der Breggen
A millivoltmeter for the frequency range from 1000 to 30 × 10^6 c/s 11,206
— and H. Rinia
A direct current supply apparatus with stabilized voltage 6, 54
Lindern, C. G. A. von
A telephone installation on ultra short waves for the tropics 6,120
— and G. de Vries
Resonance circuits for very high frequencies . . . 6,217
Lecher systems 6,240
Flat cavities as electrical resonators 8,149
Loeb, J.
The distribution of illumination on a plane parallel to a tubular lamp 12, 60
Loon, C. J. van
A simple system of bandspread in short-wave reception 6,265
— and J. Haantjes
A large-screen television projector 15, 27
Loosjes, R., H. J. Lemmens and M. J. Jansen
A new therminonic cathode for heavy loads 11,341
— and H. J. Vink
Conduction processes in the oxide-coated cathode . 11,271
—, H. J. Vink and C. G. J. Jansen
Thermionic emitters under pulsed operation 13,337
Louwerse, M. W. and F. G. Brockman
The Ferroxcube transformer core used in the Brook-haven Cosmotron 15, 73
Lowitzsch, K., H. J. di Giovanni and W. Kes
A transportable X-ray apparatus for mass chest survey 10,105
Luteijn, A. I., E. Overbosch and J. D. Fast
Preparation and casting of metals and alloys under high vacuum 15,114
Manders, Th. J. J. A.
Aerodrome illumination by means of water-cooled mercury lamps 6, 33
Light sources for cinematography 7,161
Incandescent lamps for film projection 8, 72
The "Infraphil", an apparatus for infra-red therapy 8,177
The drying lamp and its most important applications 9,249
Meerkamp van Embden, H. J., see **Embden, H. J.**
Meijering, J. L.
Hardening of metals 14,203
— and G. W. Rathenau
The oxidation of heat-resistant alloys in the presence of foreign oxides 12,213
Michels, A. J. and J. M. L. Jansen
An experimental "stroboscopic" oscilloscope for frequencies up to about 50 Mc/s,
II. Electrical build-up 12, 73
Missel, J. C. B.
Piëzo-electric materials 11,145
Moerel, P. G. and A. Rademakers
Non-ferrous copper wire for moving-coil meters . . 8,315
Mol, E. A. J.
Wire and cable insulating materials with polyvinyl chloride as base 12, 97
Mulder, J. G. W., J. de Gier, Th. Hagenberg, H. J. Meerkamp van Embden, J. A. M. Smelt and O. L. van Steenis
A steel picture-tube for television reception 14,281
Neeteson, P. A.
Flywheel synchronization of saw-tooth generators in television receivers 13,312
Nemet, A., W. A. Bayfield and M. Berindei
A diagnostic X-ray apparatus with exposure technique indication and overload portection 10, 37
Nienhuis, K., and F. M. Penning
Construction and applications of a new design of the Philips vacuum gauge 11,116
Nieuwdorp, H., A. Verhoeff and A. C. van Dorsten
The Philips 100 kV electron microscope 12, 33
Niklas, W. F.
An improved ion-trap magnet 15,258
Oele, H. A. and A. Bierman
Betatrons with and without iron yoke 11, 65
"One out of Many"
Secret production of radio receivers in occupied territory 8,337
Oosterhout, G. W. van, J. J. Went, G. W. Rathenau and E. W. Gorter
Ferroxdure, a class of new permanent magnet materials 13,194
Oosterkamp, W. J.
Eliminating scattered radiation in medical X-ray photographs 8,183
—, A. C. van Dorsten and J. B. Le Poole
An experimental electron microscope for 400 kV 9,193
— and J. van Hengel
A direct-reading dynamic electrometer 10,338
— and W. Hondius Boldingh
A 400 kilovolt installation for X-ray therapy 8,105
Opechowski, W.
Electromagnetic waves in wave guides,
I. General theoretical principles; rectangular wave guides 10, 13
II. Coaxial cables and circular wave guides 10, 46
Oranje, P. J. and J. Funke
The development of blended-light lamps 7, 34
Ouweltjes, J. L.
New phosphors for fluorescent lamps 13,346
—, W. Elenbaas, and K. R. Labberté
A new high-pressure mercury lamp with fluorescent bulb 13,109
Overbeek, A. J. W. M. van and J. L. H. Jonker
The "φ-detector", a detector valve for frequency modulation 11, 1
—, J. L. H. Jonker and K. Rodenhuis
A decade counter tube for high counting rates . . . 14,313
Overbosch, E., J. D. Fast and A. I. Luteijn
Preparation and casting of metals and alloys under high vacuum 15,114

Padmos, A. A. and J. Voogd
Mercury lamps for use in making heliographic prints 6,250
— and J. de Vries
Stresses in glass and their measurement 9,277
Pannenborg, A. E.
A measuring arrangement for wave guides 12, 15
Parrish, W.
The manufacture of quartz oscillator-plates,
I. How the required cuts are obtained 11,323
II. Control of the cutting angles by X-ray diffraction 11,351
III. Lapping and final frequency adjustment of the blanks 12,166
— and E. Cisney
An improved X-ray diffraction camera 10,157
Penning, F. M.
Velocity-modulation valves 8,214
— and K. Nienhuis
Construction and applications of a new design of the Philips vacuum gauge 11,116
Perdok, W. G. and H. van Suchtelen
An arrangement for indicating piëzo-electricity of crystals 11,151
Périlhou, P. and J. Cayzac
A stabilized extra-high tension rectifier for 5000 V, 50 mA 14,190
Peuscher, F. G. and J. van Holthoon
A pre-amplifier for use with electronic voltmeters and oscilloscopes 15,169
Plaats, G. J. van der and A. Verhoeff
A directing instrument for the operative treatment of fractures of the neck of the femur 8,237
Ploos van Amstel, J. J. A., see Amstel, J. J. A. Ploos van
Poel, F. H. J. van der and J. J. P. Valeton
The flying-spot scanner 15,221
Poole, J. B. le
A new electron microscope with continuously variable magnification 9, 33
—, A. C. van Dorsten and W. J. Oosterkamp
An experimental electron microscope for 400 kV. 9,193
Prado, E. and J. D. Veegens
A sinusoidal *RC*-oscillator for measurements in the frequency range 20-250,000 c/s 15,240
Prakke, F. and G. Alma
A new series of small radio valves 8,289
Pré, F. K. du, see **du Pré, F. K.**
Rademakers, A.
A condenser-microphone for stereophony 9,330
— and P. G. Moerel
Non-ferrous copper wire for moving-coil meters . . 8,315
Rathenau, G. W.
Imperfections in matter 15,105
—, E. W. Gorter, G. W. van Oosterhout and J. J. Went
Ferroxdure, a class of new permanent magnet materials 13,194
— and J. L. Meijering
The oxidation of heat-resistant alloys in the presence of foreign oxides 12,213
Reynst, M. F. and P. J. Hagendoorn
The recording of diagrams with the electrical pressure indicator 6, 22
Rheenen, D. W. van and F. H. de Jong
An apparatus for protection against dangerous voltages in welding equipment 15,199
Rieck, G. D.
The illumination of coal-mines and the attendant risk of explosions 10,334
Dark-room lighting 11, 53
— and L. H. Verbeek
The "Photoflux" series of flashbulbs 12,185
Riemens, J.
Temperature measurements with the optical pyrometer in the hardening department 6, 30
— and W. Elenbaas
Light sources for line spectra 11,299
Rijssel, T. W. van and J. A. Lely
Collimating X-rays in beams of very small divergence and high intensity 13, 96
Rinia, H. and P. M. van Alphen
The manufacture of correction plates for Schmidt optical systems 9,349
Projection-television receiver,
I. The optical system for the projection 10, 69
— and H. de Brey
An improved method for the air-cooling of transmitting valves 9,171
—, H. de Brey and F. L. van Weenen
Fundamentals for the development of the Philips air engine 9, 97
— and F. K. du Pré
Air engines 8,129
—, D. Kleis and M. van Tol
Experimental equipment for high-speed facsimile transmission,
I. General 10,189
— and H. J. Lindenhovius
A D.C. supply apparatus with stabilized voltage 6, 54
Rodenhuis, K.
Two triodes for reception of decimetric waves . . . 11, 79
— A. J. W. M. van Overbeek and J. L. H. Jonker
A decade counter tube for high counting rates . . . 14,313
Romeijn, F. C., P. W. Haaijman and E. J. W. Verwey
Semi-conductors with large negative temperature coefficient of resistance 9,239
Roosdorp, H. J.
On the regulation of industrial processes 12,221
An automatic recording potentiometer for industrial use 15,189
Roosendaal, L. W. and F. H. de Jong
Electric fenzing of grazing land 13,328
Rooy, A. P. van and J. de Gier
Improvements in the construction of cathode-ray tubes 9,180
Santen, G. W. van
An electrical roughness tester for the workshop . . 14, 80
— and G. H. Jonker
The ferro-electricity of titanates 11,183
Schagen, P., H. Bruining and J. C. Francken
The image iconoscope, a camera tube for television 13,119
Schenau, B. W. van Ingen
A signal generator for the testing of television receivers 15,205
Schouten, J. F.
Non-linear distortion of sound film with oblique light slit 6,110
—, G. H. Bast and D. Goedhart
A 48-channel carrier telephone system,
I. The method of modulation 9,161
II. Mechanical construction 10,353
—, F. de Jager and J. A. Greefkes
Delta modulation, a new modulating system for telecommunication 13,237
— and J. W. Klute
The influence of losses on the properties of electrical networks 7,138
Schultz, B. H.
Measuring rapidly fluctuating gas tempertures . . . 13,104
Schut, Th. G. and J. Haantjes
A line converter for the international exchange of television programmes 15,297
Siezen, G. J. and F. Kerkhof
Projection-television receiver,
III. The 25 kV anode voltage supply unit . . . 10,125
Six, W.
Some applications of Ferroxcube 13,301
— and J. Domburg
A cold cathode gas-discharge tube as a switching element in automatic telephony 15,265
Slogteren, E. van
The role of the electron microscope in the investigation of plant diseases 14, 13
Slooff, F. C. W., D. Kleis and J. M. Unk
Experimental equipment for high-speed facsimile transmission,
II. Details of the transmitter 10,257
—, M. van Tol and J. M. Unk
Experimental equipment for high-speed facsimile transmission,
III. Details of the receiver 10,265
Slooten, J. van
The functioning of triode oscillators with grid condenser and grid resistance 7, 40

Stability and instability in triode oscillators 7,171
Experimental testing of electrical networks by means of the unit function response 12,233
Mechanism of the synchronization of *LC*-oscillators 14,292
— and P. Cornelius
Installations for improved broadcast reception . . 9, 55
Smelt, J. A. M., O. L. van Steenis, J. G. W. Mulder, J. de Gier, Th. Hagenberg and H. J. Meerkamp van Embden
A steel picture-tube for television reception . . . 14,281
Snel, D. A.
Magnetic sound recording equipment 14,181
Snijders, M. J.
A transmitting valve cooler with increased turbulence of the cooling water 10,239
Snoek, J. L.
Non-metallic magnetic material for high frequencies 8,353
— and H. G. Beljers
Gyromagnetic phenomena occurring with ferrites . 11,313
— and F. K. du Pré
Several after-effect phenomena and related losses in alternating fields 8, 57
Staal, C. J. H. A.
An installation for multiplex-pulse modulation . . 11,133
Steenis, O. L. van, J. G. W. Mulder, J. de Gier, Th. Hagenberg, H. J. Meerkamp van Embden and J. A. M. Smelt
A steel picture-tube for television reception 14,281
Stegwee, J. G. C.
Experiments on the austempering of steel 6,279
Stel, M. and J. C. Derksen
Plastics and their application in the electrotechnical industry 11, 33
— and E. C. Witsenburg
Heating by means of high-frequency fields, II. Capacitive heating 11,232
Stevels, J. M.
The vitreous state 8,231
The structure of glass 13,293
Dielectric losses in glass 13,360
Stieltjes, F. H. and P. J. W. Jochems
Apparatus for testing transistors 13,254
Strutt, M. J. O. and K. S. Knol
A diode for the measurement of voltages on decimetre waves 7,124
— and A. van der Ziel
The noise in receiving sets at very high frequencies 6,178
The diode as a frequency-changing valve especially with decimetre waves 6,285
Suchtelen, H. van
Measuring the rate of watches with a cathode-ray oscillograph 9,317
— and W. G. Perdok
An arrangement for indicating piëzo-electricity of crystals 11,151
—, N. Warmoltz and G. L. Wiggerink
A method for determining the mercury content of air 11, 91
Tak, W.
The measurement of reverberation 8, 82
Measuring reverberation time by the method of exponentially increasing amplification 9,371
Taudin Chabot, J. J. M., see **Chabot, J. J. M. Taudin**
Teves, M. C. and T. Tol
Electronic intensification of fluoroscopic images . . . 14, 33
Thirup, G.
An instrument for measuring complex voltage ratios in the frequency range 1-100 Mc/s 14,102
Tol, M. van and D. Kleis
Experimental equipment for high-speed facsimile transmission,
IV. Transmission of the signals 10,289
V. Synchronization of transmitter and receiver . . 10,325
—, D. Kleis and H. Rinia
Experimental equipment for high-speed facsimile transmission,
I. General 10,189
—, F. C. W. Slooff and J. M. Unk
Experimental equipment for high-speed facsimile transmission,
III. Details of the receiver 10,265
Tol, T. and M. C. Teves
Electronic intensification of fluoroscopic images . . 14, 33
Trier, P. E., J. C. Hammerton and E. Wolfendale
A time-of-flight neutron spectrometer 15,325
Troelstra, S. A.
Applying coatings by electrophoresis 12,293
Tromp, Th. P.
Technical problems in the construction of radio valves 6,317
Tuuk, J. H. van der
Hard glass X-ray tubes in oil 6,309
X-ray tubes with rotating anodes ("Rotalix" tubes) 8, 33
High-voltage rectifier valves for X-ray diagnostics . 8,199
—, G. C. E. Burger and B. Combée
X-ray fluoroscopy with enlarged image 8,321
—, H. A. G. Hazeu and J. M. Ledeboer
An X-ray apparatus for contact therapy 8, 8
Uitjens, A. G. W., B. G. Dammers and P. D. van der Knaap
The electrical recording of diagrams with a calibrated system of coordinates 12,283
Unk, J. M., D. Kleis and F. C. W. Slooff
Experimental equipment for high-speed facsimile transmission,
II. Details of the transmitter 10,257
—, F. C. W. Slooff and M. van Tol
Experimental equipment for high-speed facsimile transmission.
III. Details of the receiver 10,265
Urk, A. Th. van
On electrical shaving 12, 25
— and K. de Boer
Some particulars of directional hearing 6,359
Valeton, J. J. P.
High-tension generators for large-picture projection television 14, 21
— and F. H. J. van der Poel
The flying spot scanner 15,221
Veegens, J. D. and Prado, E.
A sinusoidal *RC*-oscillator for measurements in the frequency range 20-250,000 c/s 15,240
— and M. K. de Vries
A simple high frequency oscillator for the testing of radio receiving sets 6,154
Veen, R. van der
Forcing tulips with artificial light 10,282
Storing seed potatoes in artificially-lighted cellars . 10,318
Influence of light upon plants 11, 43
A small greenhouse with artificial lighting for studying plant growth under reproducible conditions 12, 1
Some of the effects of the relative length of day on Poinsettia and Populus 14,175
Photosynthesis 14,298
Verbeek, L. H.
The specific light output of "Photoflux" bulbs . . . 15,317
— and G. D. Rieck
The "Photoflux" series of flashbulbs 12,185
Verhoef, J. A.
The focusing of television picture-tubes with Ferroxdure magnets 15,214
Verhoeff, A.
Epicyclic gearing for low powers 9,285
—, A. C. van Dorsten and N. Nieuwdorp
The Philips 100 kV electron microscope 12, 33
—, and G. J. van der Plaats
A directing instrument for the operative treatment of fractures of the neck of the femur 8,237
Vermeulen, D. and J. B. de Boer
The permissible brightness of lamp fittings 12,200
Motorcar headlights 12,305
Vermeulen, R.
Duplication of concerts 10,169
— and W. K. Westmijze
The "Expressor" system for transmission of music . 11,281
Verweel, J.
Magnetrons 14, 44
Verwey, E. J. W.
Electronic conductivity of non-metallic materials . 9, 46
— and R. D. Bügel
Ceramic materials with a high dielectric constant . . 10,231

—, P. W. Haaijman and E. L. Heilman
On the crystalline structure of ferrites and analogous metal oxides 9,185
—, P. W. Haaijman and F. C. Romeijn
Semi-conductors with large negative temperature coefficient of resistance 9,239
— and F. A. Kröger
New views on oxidic semi-conductors and zinc-sulphide phosphors 13, 90
Vink, H. J., C. G. J. Jansen and R. Loosjes
Thermionic emitters under pulsed operation 13,337
— and R. Loosjes
Conduction processes in the oxide-coated cathode . 11,271
Visser van Bloemen, J. P. de, see **Bloemen, J. P. de Visser van**
Voogd, J. and J. Daams
Inactivation of bacteria by ultra-violet radiation . 12,111
— and L. C. Kalff
Living room lighting with tubular fluorescent lamps 8,267
— and A. A. Padmos
Mercury lamps for use in making heliographic prints 6,250
Vries, G. de
Electromagnetic cavity resonators 9, 73
— and C. G. A. von Lindern
Resonance circuits for very high frequencies 6,217
Lecher systems 6,240
Flat cavities as electrical resonators 8,149
Vries, J. de and A. A. Padmos
Stresses in glass and their measurement 9,277
Vries, M. K. de and J. D. Veegens
A simple high-frequency oscillator for the testing of radio receiving sets 6,154
Vrijer, F. W. de and J. Haantjes
Flicker in television pictures 13, 55
Warmoltz, N.
Potential distribution at the igniter of a relay valve with mercury cathode 8,346
The ignition mechanism of relay tubes with dielectric igniter 9,105
— and A. M. C. Helmer
A flash lamp for illuminating vapour tracks in the Wilson cloud chamber 10,178
Geiger-Müller counters 13,282
—, G. L. Wiggerink and H. van Suchtelen
A method for determining the mercury content of air 11, 91
— and Winkelman
Photography of the eye with the aid of electronic flash-tubes 15,342
Weel, A. van
An experimental transmitter for u.h.f. radio-telephony with frequency modulation 8,121
An experimental receiver for u.h.f. radio-telephony with frequency modulation 8,193
— and H. J. de Boer
An instrument for measuring group delay 15,307
Weenen, F. L. van
The construction of the Philips air engine 9,125
—, H. de Brey and H. Rinia
Fundamentals for the development of the Philips air engine 9, 97
Weg, H. van de
The equalization of telephone cables 7,184
Wenke, J. W. G. and F. de Fremery
New principles of construction for the electro-acoustic installation of studios 6,139
The measurement of peak voltages in a studio installation 7, 20
Went, J. J.
The effect of the melting point and the volume magnetostriction on the thermal expansion of alloys 10, 87
Soft iron for the electromagnet of a cyclotron . . . 10,246
— and E. W. Gorter
The magnetic and electrical properties of Ferroxcube materials 13,181
— and P. Koole
A method for the control of colour deviations . . . 6,186
—, G. W. Rathenau, E. W. Gorter and G. W. van Oosterhout
Ferroxdure, a class of new permanent magnet materials 13,194
Werner, W.
The control of contact pressures 6, 61
Westmijze, W. K.
A new method of counteracting noise in sound-film reproduction 8, 97
The principle of the magnetic recording and reproduction of sound 15, 84
— and R. Vermeulen
The "Expressor" system for transmission of music 11,281
Weijers, Th. J.
Filters for carrier-wave telephony installations . . 7,104
Frequency modulation 8, 42
Comparison of frequency modulation and amplitude modulation 8, 89
Wiggerink, G. L., H. van Suchtelen and N. Warmoltz
A method for determining the mercury content of air 11, 91
Wijk, A. van
An apparatus for treatment with infra-red radiation 6,202
Wijlen, H. B. van
Radiographic examination of electronic valves . . 12,207
Wijn, H. P. J., C. M. van der Burgt and M. Gevers
Measuring methods of some properties of Ferroxcube materials 14,245
Willemze, F. G., H. C. Hamaker and J. J. M. Taudin Chabot
The practical application of sampling inspection plans and tables 11,362
Willigen, P. C. van der
The mechanical properties of welded joints 6, 97
Contact arc-welding 8,161
Penetration and welding speed in contact arc-welding 8,304
— and L. F. Defize
The determination of droplet size in arc welding by high-speed cinematography 15,122
Winkel, J. te
Carrier telegraphy 8,206
Winkelman, J. E. and N. Warmoltz
Photography of the eye with the aid of electronic flash-tubes 15,342
Wit, L. de
Apparatus for measuring light distribution 14,200
Witsenburg, E. C.
Heating by high-frequency fields,
I. Induction heating 11,165
— and M. Stel
Heating by means of high-frequency fields,
II. Capacitive heating 11,232
Wolfendale, E., P. E. Trier and J. C. Hammerton
A time-of-flight neutron spectrometer 15,325
Ydens, R. A.
Ceramics and their manufacture 10,205
Yzeren, E. A. van
The dynamo pocket torch 8,225
Zaalberg van Zelst, J. J., see **Zelst, J. J. Zaalberg van**
Zaayer, D.
An experimental 100 kW television output stage . 14,345
Zelst, J. J. Zaalberg van
Stabilized amplifiers 9, 25
Constant amplification in spite of changeability of the circuit elements 9,309
Circuit for condenser microphones with low noise level 9,357
Ziel, A. van der and M. J. O. Strutt
The noise in receiving sets at very high frequencies . 6,178
The diode as a frequency changing valve especially with decimetre waves 6,285
Zijl, H.
The calculation of lighting installations with linear sources of light 6,147
Efficiences of lighting installations 7, 97
— and A. A. Kruithof
Illumination intensity in offices and homes 8,242
Zijlstra, P.
An apparatus for determining superficial cracks in wires 11, 12
— and E. G. Dorgelo
Two transmitting valves for use in mobile installations 12,157
Zwikker, C.
The equivalent circuit of a gas discharge lamp . . . 15,161