BIBLIOTHEEK NAT. LAB.
WAALRE
GEB. W.B.

# Philips Research Reports

Volume 7

1952

16. DEC. 1953
Bibl. Nat. Lab.
N. V. PHILIPS' GLOEILAMPENFABRIEKE

PHILIPS

*All rights reserved by N.V. Philips' Gloeilampenfabrieken, Eindhoven, Netherlands. Articles or illustrations reproduced, in whole or in part, must be accompanied by full acknowledgement of the source: Philips Research Reports.*

# CHRONOLOGICAL INDEX

## PHILIPS RESEARCH REPORTS, VOL. 7


| | | | Page |
|---|---|---|---|
| *R 179* | J. L. H. Jonker | On the theory of secondary electron emission . . . . . | 1-20 |
| *R 180* | J. Volger | The specific heat of barium titanate between 110 °K and 410 °K . . . . | 21-27 |
| *R 181* | W. J. van de Lindt | Application of multi-hole coupling to the design of a variable and calibrated waveguide attenuator and impedance . . . . . | 28-35 |
| *R 182* | G. Diemer and K. Rodenhuis | Optimum geometry of microwave amplifier valves . . . . . | 36-44 |
| *R 183* | G. Diemer and H. Dijkgraaf | Langmuir's $\xi$, $\eta$ tables for the exponential region of the $I_a$-$V_a$ characteristic . . . . . | 45-53 |
| *R 184* | J. B. de Boer, V. Onate and A. Oostrijck | Practical methods for measuring and calculating the luminance of road surfaces . . . . . | 54-76 |
| *R 185* | J. L. H. Jonker A. J. W. M. van Overbeek and P. H. de Beurs | A decade counter valve for high counting rates . . . . . | 81-111 |
| *R 186* | J. A. Haringx | Instability of thin-walled cylinders subjected to internal pressure . . | 112-118 |
| *R 187* | P. Schagen, H. Bruining and J. C. Francken | A simple electrostatic electron-optical system with only one voltage. | 119-130 |
| *R 188* | A. E. Pannenborg | On the scattering matrix of symmetrical waveguide junctions . . . . | 131-157 |
| *R 189* | J. M. Stevels | Some experiments and theories on the power factor of glasses as a function of their composition; Part III . . | 161-168 |
| *R 190* | A. E. Pannenborg | On the scattering matrix of symmetrical waveguide junctions . . . . | 169-188 |

| | | | |
|---|---|---|---|
| *R 191* | J. A. Haringx | Instability of bellows subjected to internal pressure | 189-196 |
| *R 192* | Th. P. J. Botden | Transfer and transport of energy by resonance processes in luminescent solids | 197-235 |
| *R 193* | F. A. Kröger, A. Bril and J. A. M. Dikhoff | A single component white luminescent screen for television tubes | 241-250 |
| *R 194* | K. S. Knol and G. Diemer | High-frequency diode admittance with retarding direct-current field | 251-258 |
| *R 195* | B. D. H. Tellegen | A general netwotk theorem with applications | 259-269 |
| | | Additional note | 477 |
| *R 196* | A. E. Pannenborg | On the scattering matrix of symmetrical waveguide junctions | 270-302 |
| *R 197* | A. H. Boerdijk | Some remarks concerning close-packing of equal spheres | 303-313 |
| *R 198* | M. E. Wise | Dense random packing of unequal spheres | 321-343 |
| *R 199* | J. Bakker | Kinetics of the emulsion polymerization of styrene | 344-400 |
| *R 200* | A. Bril and H. A. Klasens | Intrinsic efficiencies of phosphors under cathode-ray excitation | 401-420 |
| *R 201* | A. Bril and H. A. Klasens | New phosphors for flying-spot cathode-ray tubes | 421-431 |
| *R 202* | R. Vermeulen | Dimensional analysis, units and rationalization | 432-441 |
| *R 203* | F. de Jager | Deltamodulation, a method of P.C.M. transmission using the 1-unit code | 442-466 |
| *R 204* | A. van Weel | Measurement of group-delay time in networks | 467-473 |
| *R 205* | Y. Haven and J. H. van Santen | On pre-exponential factors in formulae for ionic conductivity of solids | 474-477 |
| | | Abstracts of papers 77-80, 158-160, 236-240, | 314-320 |
| | | Author index | 478-479 |
| | | Index according to the universal decimal classification system | 480 |

VOL. 7 NR 1 FEBRUARY 1952

# Philips Research Reports

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

*R 179* *Philips Res. Rep. 7, 1-20, 1952*

## ON THE THEORY OF SECONDARY ELECTRON EMISSION *)

by J. L. H. JONKER

537.533.8

**Summary**

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, which is found to be in good accordance with results obtained experimentally.

**Résumé**

En se basant (1) sur la loi de Whiddington concernant la perte de vitesse des électrons pénétrant dans une substance solide, (2) sur la loi d'absorption expérimentale et (3) sur l'hypothèse que la distribution des électrons secondaires dans la matière est isotropique, l'auteur calcule le comportement des électrons secondaires en fonction de différents paramètres; il trouve que ce comportement concorde bien avec les résultats obtenus expérimentalement.

**Zusammenfassung**

Von dem Whiddingtonschen Gesetz (1) über die Geschwindigkeitseinbusse der Elektronen bei ihrem Eindringen in einen festen Körper, dem experimentellen Absorptionsgesetz (2), und der Voraussetzung (3) einer isotropen Verteilung der Sekundärelektronen innerhalb der Materie ausgehend, berechnet der Verfasser das Verhalten der Sekundärelektronen in Abhängigkeit von verschiedenen Parametern, welches, wie es sich zeigt, weitgehend mit den bisherigen Versuchsergebnissen übereinstimmt.

### 1. Introduction

Many attempts have been made to explain secondary emission qualitatively and quantitatively. One of these attempts, that of Bruining [1]), has the merit of giving a simple qualitative explanation of some phenomena, although there is no quantitative agreement with experimental results. The present paper is an extension of Bruining's line of argument and yields better agreement with data from previous experiments and from new experiments specially carried out with this object.

First the behaviour of the primary electrons will be discussed.

*) The theoretical work in this paper was done in 1943.

## 2. The primary electron beam

An electron beam penetrating with a great velocity into a solid substance will lose energy according to Whiddington's law [2]),

$$\varepsilon^2 V^2(x) = \varepsilon^2 V_p{}^2 - ax, \tag{1}$$

$\varepsilon V_p$ being the energy of the primary electrons when striking the surface, $\varepsilon V(x)$ the energy of these electrons after passing through the distance $x$ in the material and $a$ a constant depending on the material (fig. 1, curve I). This law was experimentally verified for high velocities (energies $>$ 10000 eV). The validity for low velocities is not known.

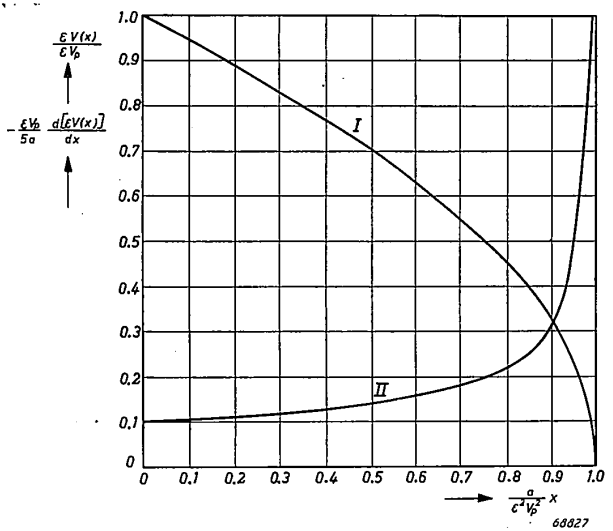



Fig. 1. Electrons penetrating into a solid substance lose energy according to Whiddington's law: $\varepsilon^2 V^2(x) = \varepsilon^2 V_p{}^2 - ax$ (curve I). The production of secondary electrons per unit of length is assumed to be proportional to $\mathrm{d}[\varepsilon V(x)]/\mathrm{d}x$ (curve II).

In accordance with several authors, Bruining [1]) assumes that the production of secondary electrons is proportional to the energy loss in the beam and therefore per unit of length equals

$$-KI_p \frac{\mathrm{d}[\varepsilon V(x)]}{\mathrm{d}x}, \tag{2}$$

supposing the primary beam current $I_p$ to be constant. In this connexion the following may be remarked: (1) the primary electron beam, when giving off energy, will undergo a certain dispersion in the material, and

(2) it will be partially absorbed, so that the intensity of current will decrease according to the absorption law

$$I(x) = I_p e^{-\alpha x}, \tag{3}$$

where $I_p$ is the beam current striking the surface, $I(x)$ the current after passing through a distance $x$ in the substance and $\alpha$ the absorption coefficient. The dispersion will cause a certain widening of the beam, but, as will appear later from experimental data, this has no great influence upon the result of our calculation, so that it will be ignored here. The influence of the absorption can be studied for nickel, the data of which are rather well known ($\alpha = 1{\cdot}5.10^6$ cm$^{-1}$ [3]), $a = 4{\cdot}5.10^{11}$ (eV)$^2$cm$^{-1}$ [4])). The maximum penetration depth $x_{max}$ of the primary electrons follows from putting $\varepsilon V(x) = 0$ in (1):

$$x_{max} = \frac{\varepsilon^2 V_p^2}{a}, \tag{4}$$

so that for primary electrons of, e.g., 250 eV energy

$$I(x_{max}) = I_p \exp\left(-\frac{\alpha}{a}\varepsilon^2 V_p^2\right) \approx I_p e^{-0{\cdot}2} \approx 0{\cdot}82\ I_p.$$

The value of $\alpha$ used was found by Becker [3]) for electrons of 10 eV, but from Terrill's measurements [5]) it is known that for high values of $\varepsilon V_p$ ($> 10^4$ eV), $\alpha v^4$ is practically constant ($v$ = velocity), so that it may be concluded that for higher energies $\alpha$ will probably have a lower value. Moreover, for many substances the ratio $\alpha/a$ is smaller than for nickel [5]). The loss of energy caused by the decrease of the number of electrons in the beam through the absorption will then in most cases be negligible in comparison with the Whiddington loss of energy (fig. 1), so that, to a good approximation, the intensity of current may be supposed to be constant all over the distance covered. According to Palluel [6]) the influence that the primary electrons rediffused have upon the number of true secondary electrons is only perceptible in the case of high primary velocities, which we shall not discuss here.

### 3. The secondary electrons

The secondary electrons will be set free over the whole length of the beam in the material, and owing to their velocity and direction they will be able to leave the surface. Let us examine the effects of this.

(1) It has been shown [1]) that a change in the work function has only a small influence on the secondary-emission factor $\delta$. This can be understood from the fact that the secondary electrons possess relatively great velocities. We shall therefore begin by ignoring this effect.

(2) Whiddington's law no longer holds in the velocity range of the secondary electrons [4]). From Katz's experiments [7]) it appears that electrons with low velocities can travel long distances in films of a thickness of several hundreds of atom layers without changing their direction, which would not be probable if there were a gradual loss of velocity.

In order to obtain further information on this point, with the aid of an apparatus previously described [8]) "countervoltage curves" have been measured for different directions in which the secondary electrons leave the substance. For low velocities of the true secondaries there is little variation in the velocity distribution (fig. 2), although the electrons travel unequal distances in the material. As the loss of velocity would probably depend on the velocity of the electron and on the distance covered, the fact of the velocity distribution being practically constant may be a further indication that Whiddington's law is no longer valid here. For higher velocities where, as may be expected, loss of energy occurs, the countervoltage curves diverge for different emission directions.

It is interesting to note that the velocity distributions of the secondary electrons for very low primary energies (25 eV) are no longer equal for different angles (fig. 3), so that it may be supposed that here we have to do with

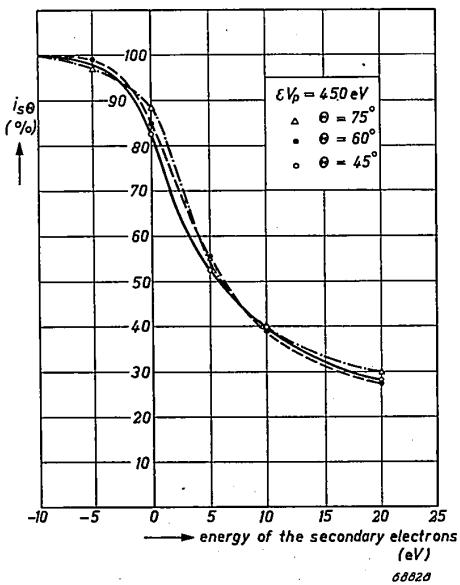



Fig. 2. Countervoltage curves of nickel for secondary electrons emerging under different angles $\Theta$ to the normal under a bombardment voltage of 450 volts. (The curves have not been corrected for the contact potential.) The velocity distribution of the true secondaries proves to be practically independent of the distance covered in the material.

a mixture of true secondary electrons and electrons rediffused with loss of velocity, which groups are indistinguishable here. Later, in the computation, it will be assumed that the true secondaries do not undergo any gradual loss of velocity in the material as is the case with the primary electrons. (3) On their way through the material, however, secondary electrons will be subject to the absorption law

$$i_s = i_0 e^{-\alpha l},$$

where $l$ is the distance the secondaries have to travel to reach the surface, $i_0$ the secondary current set free in the material and $i_s$ the secondary emission observed. This means that the secondaries may suffer such a loss of energy that they can no longer emerge from the surface.

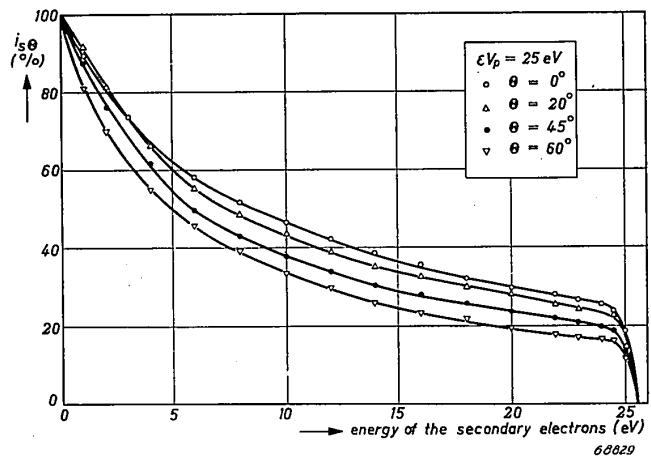



Fig. 3. The same as fig. 2 under a bombardment voltage of 25 volts. Here the electrons emitted consist of a mixture of rediffused primary and true secondary electrons, which are indistinguishable. Owing to the rediffused electrons the velocity distribution is no longer independent of the angle $\Theta$.

(4) The way in which the secondaries are set free in the material is insufficiently known. Some investigators [9]) suppose that in the beginning the secondary electrons have an initial velocity in the direction of the primary electrons and afterwards are scattered in all directions owing to one or more collisions which, according to point (3), must be elastic, so that a part moving towards the surface can be emitted. These elastic collisions, however, cannot occur very frequently along the whole way out, as this is in contradiction with Katz's results (see point (2)). A second argument is that as these collisions may as well take place close to the surface one

would expect an angular distribution more or less in the shape of a semicircle, and this would not result in an angular distribution of the emitted secondary electrons of different velocities as previously measured by us [8]) (fig. 4). As we found practically no different angular distribution of secondary electrons emitted with different velocities (fig. 4) the work function, too, can only have a secondary influence upon the angular distribution, as will be discussed in section 7. It is not probable either that the unevenness of the surface is the cause of the angular distribution, since in the case of oblique incidence of the primary electrons only small deviations in the angular distribution are observed [8]). With polycrystalline material, where the small crystals are orientated at random, a possible influence of the arrangement of the atoms in the crystal upon the angular distribution of the secondary electrons will practically be imperceptible provided the primary beam strikes a sufficiently large number of crystals. In our calculations we shall suppose that the secondaries, after being releasecl, are isotropically distributed in the substance and do not undergo any appreciable change in direction on their way to the surface. Later experiments will confirm the correctness of this assumption.

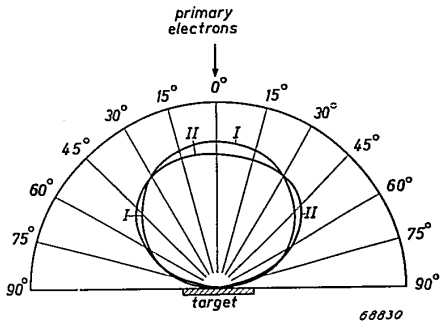



Fig. 4. Angular distribution of secondary electrons under a bombardment voltage of 100 volts. Curve I: secondary electrons with moderate velocity (energy 45 to 55 eV); curve II: slow secondary electrons (energy 5 to 15 eV).

Physically it is only possible that the secondary electrons are dislodged in the backward direction if the impulse of the primary electrons is transferred to electrons more or less bound. The mechanism of the release of the secondaries and the velocity distribution thereby caused will not be discussed further here.

On the grounds of the foregoing, in the following calculation it is assumed that

(1) the primary beam loses energy according to Whiddington's law, and that scattering and absorption may be neglected,

(2) the number of secondaries released in the material is proportional to this energy loss,
(3) the secondaries are absorbed in the material according to the exponential absorption law,
(4) after their dislodgement the secondaries are isotropically distributed in the material.

## 4. Calculation of the secondary emission

Suppose a primary beam falls perpendicularly on a solid substance. If the secondary electrons originating in the material at the point $x$ on the path of the primary electrons are equally distributed in all directions, this can be represented by a sphere with its centre at a distance $x$ from the surface (fig. 5). Secondary electrons travelling a distance $l$ through the material before reaching the surface move along a line making an angle $\varphi$ to the normal on the surface. This part of the total emission of secondaries dislodged at a distance $x$ is (see fig. 5)

$$\frac{2\pi r\,\mathrm{d}s}{4\pi R^2} = \tfrac{1}{2}\frac{r\,\mathrm{d}\varphi}{R} = \tfrac{1}{2}\sin\varphi\,\mathrm{d}\varphi.$$

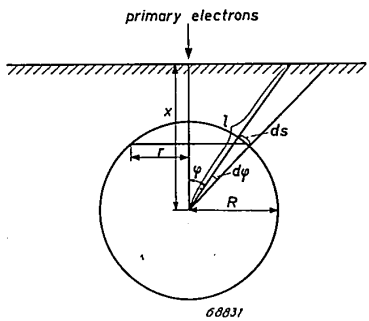



Fig. 5. Calculation of the secondary emission.

A part of these electrons is absorbed, so that only a part

$$\tfrac{1}{2}\sin\varphi\,\mathrm{d}\varphi\,e^{-\alpha l} = \tfrac{1}{2}\exp\left(-\frac{\alpha x}{\cos\varphi}\right)\sin\varphi\mathrm{d}\varphi$$

is emitted. Thus out of the whole of the secondary emission originating at a point $x$ a part

$$\int_0^{\pi/2} \tfrac{1}{2}\exp\left(-\frac{\alpha x}{\cos\varphi}\right)\sin\varphi\,\mathrm{d}\varphi$$

will reach the surface.

The secondary emission originating at a point $x$ (in a range $\mathrm{d}x$) is equal to (see section 2)

$$-KI_p \frac{\mathrm{d}[\varepsilon V(x)]}{\mathrm{d}x}\,\mathrm{d}x = \tfrac{1}{2}\,KaI_p\,(\varepsilon^2 V_p{}^2 - ax)^{-1/2}\,\mathrm{d}x,$$

so that the surface is reached by

$$\mathrm{d}I_s = \tfrac{1}{2}\,KaI_p(\varepsilon^2 V_p{}^2 - ax)^{-1/2}\,\mathrm{d}x \int_0^{\pi/2} \tfrac{1}{2} \exp\left(-\frac{ax}{\cos\varphi}\right) \sin\varphi\,\mathrm{d}\varphi.$$

Integrating from $x = 0$ to $x = x_{\max} = \varepsilon^2 V_p{}^2/a$ gives for the total secondary-emission current:

$$I_s = \tfrac{1}{4}\,KaI_p \int_0^{\varepsilon^2 V_p{}^2/a} \mathrm{d}x \int_0^{\pi/2} \mathrm{d}\varphi \left\{\exp\left(-\frac{ax}{\cos\varphi}\right)\right\} (\varepsilon^2 V_p{}^2 - ax)^{-1/2} \sin\varphi.$$

By substituting $\cos\varphi = z$, $(a/az)\,(\varepsilon^2 V_p{}^2 - ax) = y^2$ and $(a/a)\,\varepsilon^2 V_p{}^2 = A$, this expression changes into

$$\delta = \frac{I_s}{I_p} = \tfrac{1}{2}\,K \sqrt{\frac{a}{a}}\, f(A), \tag{5}$$

where

$$f(A) = \int_0^1 \mathrm{d}z\, e^{-A/z}\, z^{1/2} \int_y^{\sqrt{A/z}} e^{y^2}\,\mathrm{d}y. \tag{6}$$

$\delta$ has a maximum if

$$\frac{\mathrm{d}\delta}{\mathrm{d}(\varepsilon V_p)} \equiv K \int_0^1 \left[\tfrac{1}{2} - e^{-A/z}\sqrt{A/z} \int_0^{\sqrt{A/z}} e^{y^2}\,\mathrm{d}y\right] \mathrm{d}z = 0,$$

thus if

$$\int_0^1 \mathrm{d}z\, e^{-A/z}\sqrt{A/z} \int_y^{\sqrt{A/z}} e^{y^2}\,\mathrm{d}y = \tfrac{1}{2}.$$

From a curve given below (fig. 7) this can be seen to be the case for

$$A \equiv (a/a)\,\varepsilon^2 V_p{}^2 = 0{\cdot}505, \quad \text{or} \quad \sqrt{A} = 0{\cdot}71. \tag{7}$$

This value is less than that found by Bruining [1]) in his calculation ($\sqrt{A} = 0{\cdot}92$). Calculating from $\sqrt{A} = 0{\cdot}71$ the value $\varepsilon V_{p\max}$ of $\varepsilon V_p$ for which the maximum occurs for nickel, by taking for $a$ Becker's value [3]) $a = 1{\cdot}5.10^6$ $\mathrm{cm}^{-1}$ and for $a$ the value [1]) $a = 3{\cdot}5.10^{12}$ $(\mathrm{eV})^2\mathrm{cm}^{-1}$ deduced from Terrill's experiments [10]) for energies of 25 to 50 keV, we find $\varepsilon V_{p\max} = 1086$ eV

(Bruining, l.c., 1420 eV), which lies closer to the experimental value of 500 eV. Conversely, with the aid of the values of $a$ determined for a number of substances (for high energies) [1], [10]) and of the measured values [1]) of $\varepsilon V_{p\max}$, the value of $\alpha$ can be calculated for these substances (table I).

TABLE I

| Metal | $10^{-12}\ a$, $(\mathrm{eV})^2\mathrm{cm}^{-1}$ | $\varepsilon V_{p\max}$, eV(exp.) | $\delta_{\max}$ (exp.) | $10^{-7}\ \alpha$, $\mathrm{cm}^{-1}$ |
|---|---|---|---|---|
| Li | 0·21 | 90 | 0·55 | 1·31 |
| Be | 0·75 | 200 | 0·53 | 0·95 |
| Cs | 0·75 | 400 | 0·69 | 0·24 |
| Al | 1·1 | 300 | 0·97 | 0·62 |
| Ba | 1·5 | 400 | 0·83 | 0·47 |
| Ni | 3·5 | 500 | 1·3 | 0·71 |
| Cu | 3·6 | 600 | 1·3 | 0·5 |
| Mo | 4·1 | 400 | 1·24 | 1·29 |
| Ag | 4·2 | 800 | 1·47 | 0·33 |
| W | 7·5 | 700 | 1·4 | 0·77 |
| Au | 8·9 | 875 | 1·45 | 0·59 |

The values of $a$, however, were determined for much higher energies than those corresponding to the maxima. Assuming the correctness of Becker's $\alpha$ for nickel (l.c.), measured for energies in accordance with those of the secondary electrons (10 eV), we can calculate the corresponding value of $a$ with the aid of the measured value of $\varepsilon V_{p\max}$:

$$a = 1{\cdot}5.10^6 \times 500^2/0{\cdot}505 = 7{\cdot}4.10^{11}\ (\mathrm{eV})^2\mathrm{cm}^{-1}.$$

This value is only 0·21 times that given in table I. Proceeding in the same way for gold, for which Partzsch and Hallwachs [11]) determined $\alpha = 10^6\ \mathrm{cm}^{-1}$, we find $a = 1{\cdot}52.10^{12}\ (\mathrm{eV})^2\mathrm{cm}^{-1}$, which is 0·17 times the value given in table I. Thus it appears that with energies of several hundreds of electron-volts Whiddington's constant is reduced to about one-fifth of the value occurring with energies of several tens of keV. This means, however, that the values of $\alpha$ in table I will be reduced to one-fifth too. To make sure about this point direct measurements of $\alpha$ for different substances at low energies are highly desired.

From (4) and (7) it follows that the maximum penetration depth is

$$x_{\max} = \varepsilon^2 V_p^2/a = 0{\cdot}505/a \approx 33\ \text{Å},$$

which is in agreement with the value $x_m = 30$ Å, found by Bruining in another way, for the mean depth at which the secondaries are dislodged. From the trend of $\mathrm{d}[\varepsilon V(x)]/\mathrm{d}x$ it follows that $x_m$ must lie close to $x_{\max}$ (fig. 1).

The above calculation of $\delta$ can be extended to the case where the primary beam strikes the surface under an angle $\nu$ to the normal (fig. 6). Secondary electrons dislodged at a point $x$ on the path of the primary electrons in the material will then be located at a distance $x\cos\nu$ from the surface, so that in the above calculation $x$ has to be replaced by $x\cos\nu$ and the absorption factor becomes $\exp(-\alpha x\cos\nu/\cos\varphi)$. $A$ and $\delta$ then become

$$A' = (a/a)\ \varepsilon^2 V_p^2\ \cos\nu \tag{8}$$

and

$$\delta = \tfrac{1}{2} K \sqrt{\frac{a}{a}} f(A') \sqrt{\frac{1}{\cos\nu}}. \tag{9}$$

The function $f(A')$ has been determined graphically and is given in fig. 7 as a function of $\sqrt{A'}$.

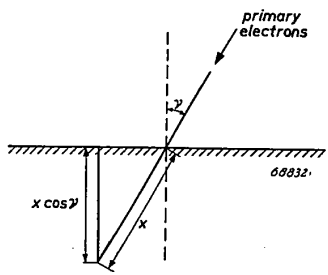



Fig. 6. Calculation of the secondary emission.

## 5. Experiments

From the calculation in section 4 we know that for the maximum in the $\delta$-curve $A' = 0{\cdot}505$, so that

$$\varepsilon V_{p\max} \sqrt{\cos\nu} = \sqrt{0{\cdot}505\ a/a} = \text{constant}, \tag{10}$$

at least for a limited voltage range where $a$ and $a$ are supposed to be constant. For some materials, such as nickel, nickel carbide and lithium, measurements of $\delta$ as a function of $\varepsilon V_p$ with different angles of incidence $\nu$ are available [1]) (from the available measurements of barium the exact value of $\varepsilon V_{p\max}$ could not be determined). Starting from a measured value of the maximum at a certain angle of incidence, we can now calculate the position of the maximum for other angles of incidence. From table II it

appears that for the materials mentioned a very good experimental agreement is found.

TABLE II

| Material | | Angle of incidence to the normal | | | | |
|---|---|---|---|---|---|---|
| | | 0° | 40° | 48° | 60° | 70° |
| Nickel | measured (volts) | 500 | | 600 | 700 | 880 |
| | calculated | **500** | | 615 | 707 | 857 |
| Nickel carbide | measured | 280 | 300 | | 365 | 470 |
| | calculated | **280** | 293 | | 365 | 478 |
| Lithium | measured | 90 | | 140 | 179 | 210 |
| | calculated | 122·5 | | 151 | 175 | **210** |

The values of $\sqrt{0{\cdot}505\ a/\alpha}$ for nickel and nickel carbide were calculated from the position of the maximum at normal incidence, for lithium from the highest potential, since it appears that with lower potentials ($< 150$ volts) there are deviations, probably due to considerable deviations from Whiddingston's law ($a$ no longer being a constant). It is also striking that

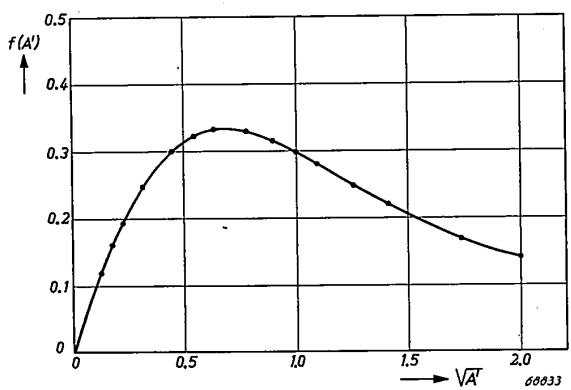



Fig. 7. Theoretical universal curve for $\delta$ as a function of $\varepsilon V_p$, given by

$$f(A') = \int_0^1 \mathrm{d}x\, e^{-A'/x}\, x^{1/2} \int_x^{\sqrt{A'/x}} e^{y^2}\, \mathrm{d}y$$

as a function of $\sqrt{A'} = \sqrt{a/\alpha}\ .\ \varepsilon V_p \sqrt{\cos\nu}$.

Katz, in his above-mentioned experiments [7]), where he obtained an electron-optical image through a silver film, did not obtain any sharp image above 130 volts. Further, it appears that the widening of the primary beam as a result of the dispersion in the material has no influence on these results.

With the aid of (9) the value of $\delta_{max}$ can now also be calculated as a function of the angle of incidence of the primary electrons, starting from the measured value at normal incidence. The results for the three materials mentioned above are given in table III.

TABLE III

| Material | | Angle of incidence to the normal | | | | |
|---|---|---|---|---|---|---|
| | | 0° | 40° | 48° | 60° | 70° |
| Nickel | $\delta_{max}$ measured | 1·225 | | 1·5 | 1·75 | 1·9 |
| | $\delta_{max}$ calculated | 1·225 | | 1·5 | 1·75 | 2·1 |
| Nickel carbide | $\delta_{max}$ measured | 0·93 | 1·06 | | 1·14 | 1·30 |
| | $\delta_{max}$ calculated | 0·93 | 0·99 | | 1·32 | 1·59 |
| Lithium | $\delta_{max}$ measured | 0·47 | | 0·63 | 0·78 | 1 |
| | $\delta_{max}$ calculated | 0·47 | | 0·6 | 0·67 | 0·805 |

Although there is satisfactory agreement, the deviations are somewhat greater than in table II. There is good agreement for nickel, which was measured at the highest potentials. With low potentials the deviation from Whiddington's law again becomes more evident. The curve of $f(A')$ as a function of $\sqrt{A'}$ (fig. 7) proves to be a sort of *universal curve giving the relative course of $\delta$ as a function of $\varepsilon V_p$ and valid for all materials and for all directions of incidence of the primary electrons.*

The good agreement obtained in the above tables induced us to investigate whether such a universal curve exists also experimentally, which would strongly support the various assumptions made. If we plot the measured $\delta$ for one material as a function of $\sqrt{A'} = \text{const.} \times \varepsilon V_p \sqrt{\cos \nu}$, the maxima of the curves measured for different angles of incidence $\nu$ will have the same abscissa. If, now, we change the ordinates in such a manner that the maxima of the different curves coincide, then the curves obtained appear to be congruent, so that each of the materials mentioned above is found to

have a universal curve (figs 8, 9 and 10). Moreover, it appears that these curves for the three materials are mutually congruent, so that, apparently, there exists a universal curve experimentally, independent of the material [12]).

Although there is general agreement in the trend of the experimental and the calculated curves, the maximum of the latter is considerably more pronounced. This, however, can easily be explained. In our calculation we supposed $a$ and $\alpha$ to be constants. In the preceding pages we discussed that $a$ depends on the velocity of the electrons. This is also the case for $\alpha$, as appears from the experiments of Terrill [5]), who found that for very high

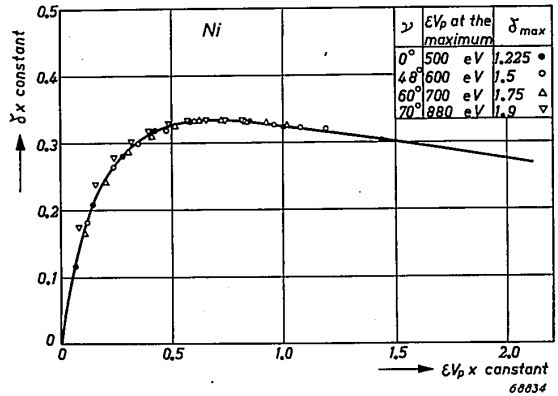



Fig. 8. Experimental universal curve for $\delta$ of nickel, obtained by making the maxima belonging to different angles of incidence $\nu$ of the primary beam coincide.

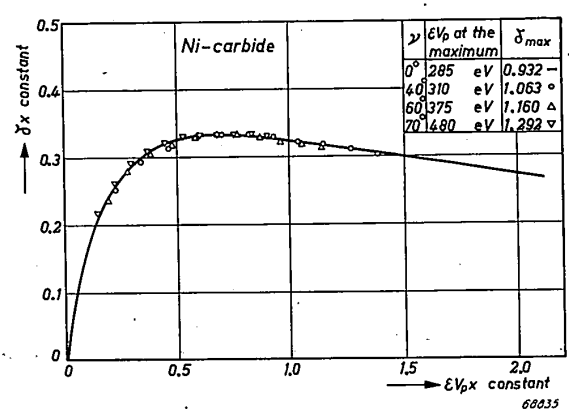



Fig. 9. The same as fig. 8 for nickel carbide.

velocities $av^4 =$ constant ($v$ = velocity). It also follows from the experiments of Katz [7]), who found that for lower velocities the transmissivity increases, and that at velocities of some tens of eV the transmission has a maximum.

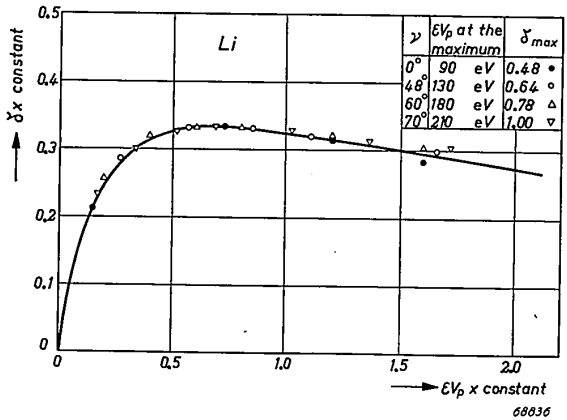



Fig. 10. The same as fig. 8 for lithium.

We can now ascertain the influence of a variation of $a$ on the trend of $\delta$ as a function of $\varepsilon V_p$. If the course of the secondary emission is measured as a function of the energy $\varepsilon V_p$ within a small velocity range of the secondaries, for which it may be supposed that at least $a$ is more constant, then indeed a curve with a sharper maximum is found (fig. 11). Conversely, it can be proved mathematically, by introducing a Maxwellian velocity distribution of the secondaries and Terrill's relation $av^4 =$ constant, that the trend of $f(A')$ as a function of $\sqrt{A'}$ draws nearer to the experimental curve.

## 6. Calculation of the angular distribution of the secondary electrons

It is desirable, perhaps, to define the notion of "angular distribution of the secondary electrons" as it is meant here. The curves which we previously measured [8]) and which will be calculated here give the number of electrons measured in different directions by means of a collector with a fixed solid angle. It makes a difference whether one makes measurements on a radiant plane surface or on a point source of radiation. If in both cases the surface radiates uniformly, the former will give a cosine distribution, the latter a uniform distribution. As in our case we are considering the electron radiation of a uniformly radiating point $dx$ on the path of the primary beam, in principle we have to do with a point source of radiation.

The angular distribution of the emitted electrons can now be calculated with the aid of the calculation of section 4. Let a part $\Delta$ of the secondary emission $di_{sx}$ dislodged in a part $dx$ on the path of the primaries go to the

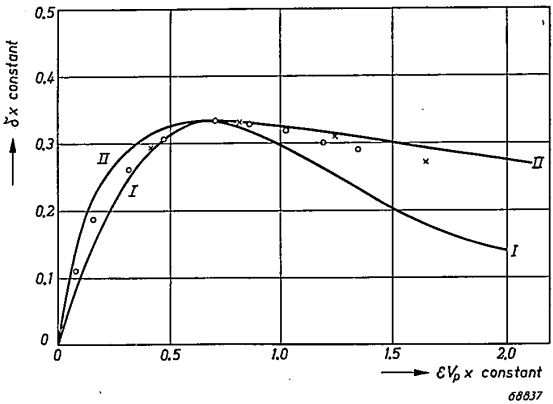



Fig. 11. Comparison of the calculated universal curve of fig. 7 (I) with the experimental universal curve of fig. 8 (II). The points marked o are those obtained by measuring $\delta$ as a function of $\varepsilon V_p$ for secondary electrons within the energy range 5 to 8 eV. The points × are some of those of a universal curve calculated by assuming a Maxwellian velocity distribution and the relation $av^4$ = constant. The difference between curves I and II can be ascribed to the velocity distribution of the secondaries.

surface along the line $l$ under an angle $\Theta$ (fig. 12). As we assumed an isotropic distribution of the secondaries $\Delta$ is independent of $\Theta$. To reach the surface the secondaries must travel a distance $l = x/\cos\Theta$. Owing to the absorption only the part $di_{s\Theta} = \Delta di_{sx} \exp(-\alpha x/\cos\Theta)$ will reach the surface. Now

$$di_{sx} = -KI_p \frac{d[\varepsilon V(x)]}{dx} dx$$

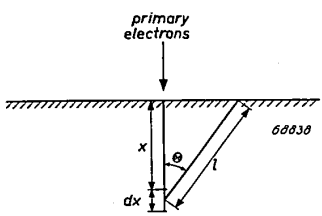



Fig. 12. Calculation of the angular distribution of the secondary emission.

(see section 4), so that with the aid of Whiddington's law we find

$$\mathrm{d}i_{s\Theta} = \tfrac{1}{2}\,\Delta K a I_p \left\{\exp\left(-\frac{ax}{\cos\Theta}\right)\right\}(\varepsilon^2 V_p{}^2 - ax)^{-1/2}\,\mathrm{d}x\,.$$

Integrating from $x = 0$ to $x = x_{\max}$, while keeping $\Theta$ constant, we get the total secondary emission emerging in the direction of $l$ under an angle $\Theta$:

$$i_{s\Theta} = \tfrac{1}{2}\,\Delta K a I_p \int_0^{x_{\max}} \left\{\exp\left(-\frac{ax}{\cos\Theta}\right)\right\}(\varepsilon^2 V_p{}^2 - ax)^{-1/2}\,\mathrm{d}x\,.$$

Let

$$\frac{a}{a\cos\Theta}(\varepsilon^2 V_p{}^2 - ax) = u^2$$

and

$$\varepsilon V_p \sqrt{\frac{a}{a\cos\Theta}} = t\,,$$

then, after some calculation, we get

$$i_{s\Theta} = \Delta K I_p \sqrt{\frac{a\cos\Theta}{a}}\,F(t)\,, \tag{11}$$

where

$$F(t) = e^{-t^2}\int_0^t e^{u^2}\,\mathrm{d}u\,. \tag{12}$$

In fig. 13 we have plotted $F(t)$ as a function of $t$ (maximum for $t = 0{\cdot}92$).

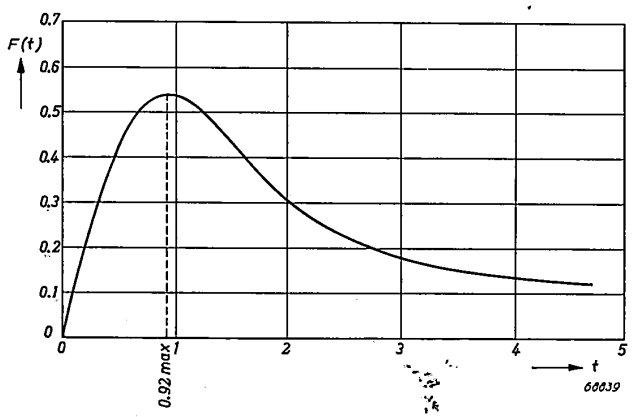



Fig. 13. Graph of the function $F(t) = e^{-t^2}\int_0^t e^{u^2}\,\mathrm{d}u$.

Fig. 14 shows how $i_{s\Theta}$ varies with varying $\Theta$ for $\varepsilon V_p\sqrt{a/a} = 0{\cdot}92$ (curve I) and for $\varepsilon V_p\sqrt{a/a} = 2$ (curve II). The curves obtained show but little deviation from the cosine law (curve III). For the lower value of $\varepsilon V_p\sqrt{a/a}$ (curve I) the circle is flattened, which flattening we actually find with the formerly [8]) measured curves of secondary electrons with low velocities (fig. 4, curve II). This flattened curve, as also found experimentally, is quite independent of the angle of incidence of the primary electrons.

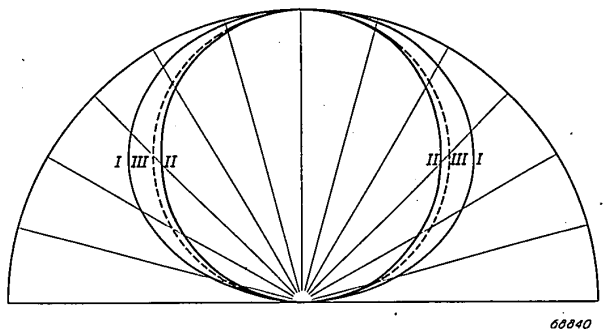



Fig. 14. Calculated angular distribution of the secondary emission with $\varepsilon V_p\sqrt{a/a}$ as parameter. Curve I: distribution when $\varepsilon V_p\sqrt{a/a} = 0{\cdot}92$; curve II: the same when $\varepsilon V_p\sqrt{a/a} = 2$; curve III: cosine distribution.

For "secondary" electrons with a higher energy (e.g. 50 eV) a curve (I in fig. 4) is measured which draws closer to the cosine law. In our calculation this would agree with a greater value of $\varepsilon V_p\sqrt{a/a}$ (see fig. 14), and this can only mean here a greater $a$. According to Terrill [5]), however, for very high velocities $a$ decreases; according to Katz [7]) the permeability shows a maximum, and therefore $a$ a minimum, for low velocities, so that the conclusion is obvious that for intermediate velocities $a$ must somewhere have a maximum. The trend of $a$ in this voltage range, however, is not exactly known. If our calculation were applicable to the rapid secondaries, then the trend of the angular distribution would agree with this variation of $a$ in this range. Further, we found experimentally that the angular distribution is only slightly dependent upon $\varepsilon V_p$, whilst the calculation, with constant $a$ and $a$, would lead one to expect a further flattened distribution, especially for low values of $\varepsilon V_p$. This would point to the fact that in this range Whiddington's law is no longer valid ($a$ no longer constant), which confirms our former statements.

## 7. The influence of the surface barrier

Up to now we have ignored the influence of the surface barrier upon $\delta$ and upon the angular distribution of the secondary emission. This influence will be investigated in this section.

Let an electron inside the material travel under an angle $p$ to the normal and with a velocity $v_i$ and, after being emitted, under an angle $q$ to the normal and with a velocity $v_o$, then we may write

$$\frac{\sin q}{\sin p} = \frac{v_i}{v_o} = \text{const.} = n \tag{13}$$

and differentiating

$$n \cos p \, dp = \cos q \, dq \, . \tag{14}$$

A current $di_s$ in a solid angle $d\Phi$ inside the material will outside the material be located in a solid angle $d\Omega$. For the ratio of the current density $J_o$ outside the material to the density $J_i$ inside we get, using (13) and (14)

$$\frac{J_o}{J_i} = \frac{di_s/d\Omega}{di_s/d\Phi} = \frac{d\Phi}{d\Omega} = \frac{2\pi \sin p \, dp}{2\pi \sin q \, dq} = \frac{1}{n^2} \frac{\cos q}{\cos p}. \tag{15}$$

For $p = 0$ and consequently $q = 0$ (15) gives

$$J_o(0) = J_i(0)/n^2 \, , \tag{16}$$

so that

$$\frac{J_o}{J_i} = \frac{J_o(0)}{J_i(0)} \frac{\cos q}{\cos p}. \tag{17}$$

This means that if inside the material the current arriving before the surface barrier follows a cosine law, i.e. if $J_i = J_i(0) \cos p$, a cosine law will then be followed outside too, and that then *the surface barrier will not change the shape of the angular distribution.*

This does not hold if the current distribution inside the material does not follow a cosine law but, e.g., has an isotropic distribution. In that case we write (15) in the form

$$\frac{J_o}{J_i} = \frac{\cos q}{n^2 \sqrt{1 - \sin^2 p}} = \frac{\cos q}{n^2 \sqrt{1 - \dfrac{\sin^2 q}{n^2}}}, \tag{18}$$

from which it can be seen that outside the material a cosine distribution would occur for large values of $n^2$ only (fig. 15). This could only be the case if the energy of the emitted electrons would be small in comparison with the energy loss in the surface barrier.

Actually an isotropic distribution will be changed in such a way as to reduce the current density outside the material for large values of $q$.

From fig. 15 we see that for secondary electrons with high velocities ($n^2$ small) the distribution curve grows flatter, which is in contradiction with the experiment (fig. 4), so that we conclude that it is probable that the secondary electrons before passing the surface barrier have a cosine distribution and not an isotropic distribution.

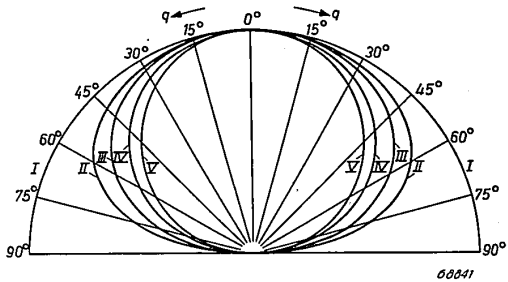



Fig. 15. The function

$$n^2 \frac{J_o}{J_i} \frac{\cos q}{\sqrt{1 - \frac{\sin^2 q}{n^2}}},$$

giving the distribution of the current density outside the material if it is assumed that inside the material the secondary electrons just before passing the surface barrier are isotropically distributed. Curve I: distribution when $n^2 = 1$; curve II: the same when $n^2 = 1{\cdot}2$; curve III: the same when $n^2 = 1{\cdot}5$; curve IV: the same when $n^2 = 3$; curve V: cosine distribution ($n^2 = \infty$).

If we assume a cosine distribution it can be seen from (16) that the current density outside the material changes with $1/n^2$, so that we must include then a constant factor $1/n_m{}^2$ in the expressions (5) and (9) for $\delta$, $n_m$ being a mean value of $n$ for a given velocity distribution. Now $n = v_i/v_o = \sqrt{\varepsilon V_i/\varepsilon V_o}$. Taking for $\varepsilon V_o$ of nickel a mean value of 10 eV (fig. 2), and assuming for the work function and the conduction-band level values of 5 eV and 10 eV, respectively, we assume for $\varepsilon V_i$ a mean value of 20 eV, so that then $n_m{}^2 = 2$. A change in the work function of 50 % from 5 eV to 2·5 eV decreases the value of $n_m{}^2$ to 1·75. This results in an increase of $\delta$ of only 14%, which is in agreement with Bruining's experimental results.

## 8. Conclusions and acknowledgement

Starting from a number of simple assumptions we have found an expression for $\delta$ as a function of $\varepsilon V_p$, from which the following laws may be deduced.

(1) For the maximum in the $\delta$-curve if $\nu$ is the angle of the primary beam to the normal:

$$\delta_{max} \approx \frac{1}{\sqrt{\cos \nu}} \qquad \text{(I)}$$

and

$$\varepsilon V_{p\,max} \approx \frac{1}{\sqrt{\cos \nu}}. \qquad \text{(II)}$$

(2) There exists a universal function independent of $\nu$ and of the material constants $C_m$ and $C_m'$ (similarity law):

$$\delta C_m \sqrt{\cos \nu} = f' \left\{ \varepsilon V_p C_m' \sqrt{\cos \nu} \right\}. \qquad \text{(III)}$$

(3) The angular distribution of the secondary emission is, if $\Theta$ is the angle to the normal, given by

$$\frac{i_{s\Theta}}{I_p} = C \sqrt{\cos \Theta}\, F \left\{ \varepsilon V_p \frac{C_m'}{\sqrt{\cos \Theta}} \right\}. \qquad \text{(IV)}$$

These laws have been found to be in good agreement, in detail as well as in the whole, with the experimental data, although both assumptions on which our calculations were based have only a limited validity.

Finally, the author wishes to express his thanks to Dr H. Bruining for many stimulating discussions on the subject of this article.

*Eindhoven, July 1951*

REFERENCES

1) H. Bruining, Thesis, Leyden, 1938; Die Sekundär-Elektronen-Emission fester Körper, Springer, Berlin, 1942, p. 60 ff.
2) R. Whiddington, Proc. roy. Soc. A **86**, 360-370, 1912.
3) A. Becker, Ann. Physik **2**, 249-263, 1929.
4) G. Diemer and J. L. H. Jonker, Philips Res. Rep. **5**, 161-172, 1950.
5) H. M. Terrill, Phys. Rev. **24**, 616-621, 1924.
6) P. Palluel, C. R. Acad. Sci. Paris **224**, 1492-1494, 1947.
7) H. Katz, Ann. Physik **33**, 160-168; 169-184, 1938.
8) J. L. H. Jonker, Philips Res. Rep. **6**, 370-385, 1951.
9) O. Hachenberg, Ann. Physik **2**, 404-416, 1948.
10) H. M. Terrill, Phys. Rev. **22**, 101-108, 1923.
11) A. Partzsch and W. Hallwachs, Ann. Physik **41**, 247-272,1913.
12) E. M. Baroody, Phys. Rev. **78**, 780-787, 1950.

# THE SPECIFIC HEAT OF BARIUM TITANATE BETWEEN 100 °K AND 410 °K

by J. VOLGER

536.631:548.0.536

**Summary**

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.

**Résumé**

On a mesuré la chaleur spécifique de $BaTiO_3$ polycristalline dans un calorimètre de Nernst. La contribution du réseau ne s'exprime pas exactement par une function de Debye. On a déterminé les anomalies de la chaleur spécifique dans les trois régions de transition.

**Zusammenfassung**

Die spezifische Wärme von polykristallinischem $BaTiO_3$ wurde mit einem Kalorimeter nach Nernst gemessen. Es zeigt sich, daß der Beitrag des Gitters nicht mit einer einzigen Debye-Funktion beschrieben werden kann. Die Anomalien der spezifischen Wärme in den drei Übergangsgebieten wurden bestimmt.

## 1. Introduction

From X-ray and dielectric measurements it is well known that barium titanate exhibits polymorphism [1], [2]) with transition temperatures of about 120 °C, 5 °C and —70 °C, where maxima of the permittivity occur. Coming from high temperatures the crystal shows cubic, tetragonal, orthorhombic and rhombohedral symmetry. Anomalies of the specific heat have been observed in the transition regions at 5 °C and 120 °C [3]), [4]), [5]), [6]).

The purpose of this paper is to offer some new caloric measurements on polycrystalline $BaTiO_3$. The specific heat vs temperature curve of the compound between 100 °K and 410 °K will be given, with special attention to the behaviour in the three transition regions.

## 2. Methods

A vacuum calorimeter of the classical Nernst type was used. The apparatus is shown in fig. 1. Heat exchange between the sample and its surroundings could be reduced by adjusting the temperature of a thick copper shielding. The temperatures of sample and shielding were determined from voltage measurements of silver-constantan thermo-elements on a Diesselhorst compensator. Heat was supplied by passing electric current through heating coils. For low-temperature measurements the whole set was immersed in a bath of liquid nitrogen, thereby providing for an initial cooling of shielding and sample.

Two methods have been followed.

(A) In the first place absolute measurements have been carried out. During an exactly known period (e.g. 2 minutes) heat was supplied to the sample, the total amount of this heat being calculated from the time, current and voltage, whilst the resultant increase of temperature (about 1 or 2 degrees) could be determined from the increase of thermo-e.m.f., taking into account the small drift of temperature before and after the heating period. From these data $C_p$ can be calculated.

(B) Secondly, relative measurements have been taken, in particular in the transition ranges. During a long period (e.g. an hour) the reciprocal drift of temperature (expressed in seconds per 10 $\mu$V increase of thermo-e.m.f.) was determined while supplying a small and practically constant amount of heat, which in any case varied only very slowly. In other experiments a slow cooling was provided for by making the shielding much

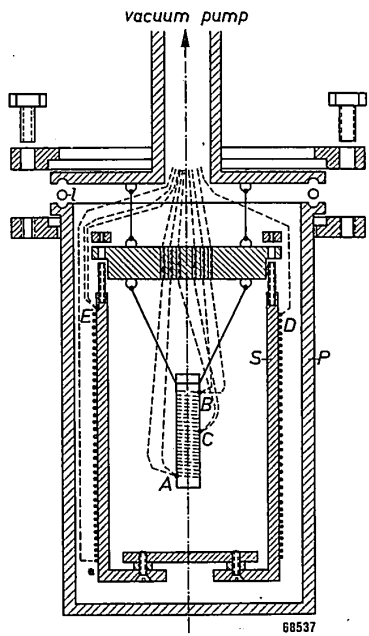



Fig. 1. The Nernst calorimeter. The sample is supported with a nylon thread. The heating coil is made of 0·1 mm thick nickel-chromium wire. At *A* and *B* wires for current and voltage measurements are connected. The temperature of the sample is read with the help of a thermo-element (*C*). The shielding (*S*) has also a heating coil (*D*); its temperature is read from the thermo-e.m.f. at *E*. The whole system is mounted in an iron jar (*P*) which is made vacuum-tight with a lead ring (*l*). The vacuum is provided by a mercury diffusion pump.

colder than the sample. The reciprocal drift is proportional to $C_p$ of the substance, the proportionality factor being a smooth function of time (temperature) during the experiment. This factor can be determined if $C_p$ is known. By extending such a heating or cooling experiment sufficiently far beyond and in advance of the transition region, to temperatures where absolute measurements of $C_p$ had been carried out, and then interpolating the calculated factors, $C_p$ could be calculated also in the transition regions. With the second method more reliable values of the anomalies of $C_p$ can be obtained than with the first one.

### 3. Sample

Barium titanate in the form of a sintered polycrystalline product was prepared by Dr G. H. Jonker of this laboratory. Barium carbonate and titanium dioxide were mixed, milled and prefired at 1000 °C. Afterwards the titanate obtained was milled again, mixed with a suitable organic binder and shaped into a cylindrical rod which was definitely fired at 1360 °C. The rod had a weight of about 54 grams and a diameter $2r = 15$ mm.

An important question is whether the heat applied is distributed fast enough in the sample. This, however, was certainly the case. Putting the diffusitivity of the material at about $5.10^{-3}$ cm$^2$ sec, by a calculation based upon the model of a cylinder in which initially the temperature is everywhere the same whilst from a given moment onwards the surface temperature assumes another, constant value, it is found that within a minute the distribution has practically finished [7]). This is confirmed experimentally. With method (B) one degree of temperature variation took about 2 minutes.

### 4. Results

The results of the absolute measurements (method (A)) are given in table I.

These values have been plotted in fig. 2, where it is seen that they follow a smooth curve. The dashed parts of the curve indicate the three transition regions (I, II and III) where humps in the specific heat occur.

The anomalies of $C_p$ in these regions were determined by method (B). As an example, the observations obtained in the lowest transition region are reproduced in fig. 3. Fig. 4 shows the anomaly at 118 °C as calculated from such observations. The values marked × were measured directly (method (A)).

It appeared that the temperatures where maxima of $C_p$ are found depend on whether the sample is heated or cooled: the substance shows thermal hysteresis at the transitions. Both kinds of critical points ($T_\uparrow$ and $T_\downarrow$) were well reproducible. The anomalous heat has been integrated with the help of a planimeter. This integral, the heat of transition $\Delta E$,

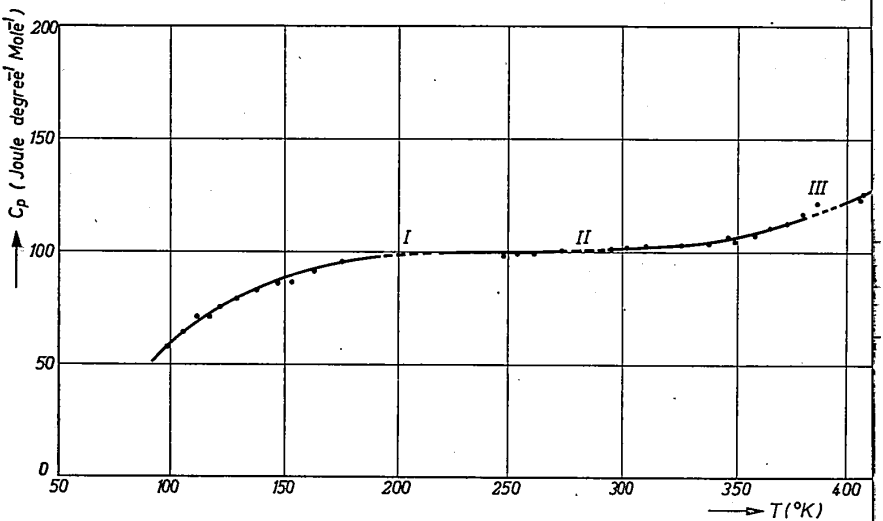



Fig. 2. $C_p$ of $BaTiO_3$. The dashed parts of the curve indicate the transition regions I, II and III.

is given in table II. In this table some other characteristics of the substance investigated are collected.

The quantity $\Delta T$ is the width of a peak, taken as the difference between the temperatures at which the anomalous specific heat has half its top value.

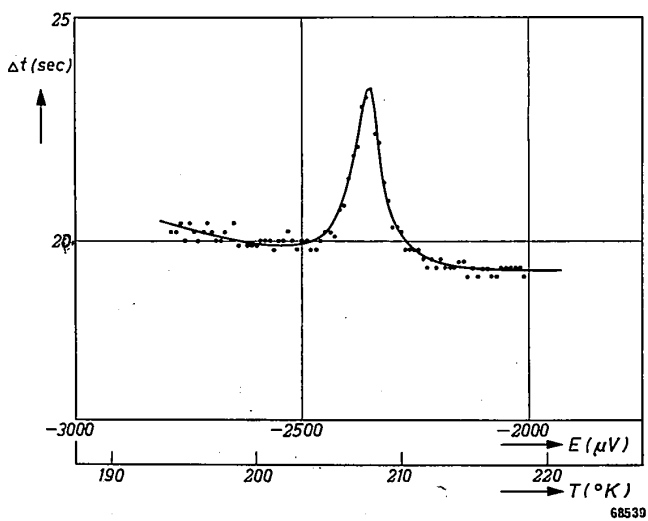



Fig. 3. Example of observations made with the slow heating method (method (B)). Transition I.

TABLE I

$C_p$ of $BaTiO_3$ (polycrystalline) as a function of temperature

| $T$(°K) | $C_p$ (Joule degree$^{-1}$ Mole$^{-1}$) | $T$(°K) | $C_p$(Joule degree$^{-1}$ Mole$^{-1}$) |
|---|---|---|---|
| 99 | 57·5 | 295 | 101·5 |
| 106 | 64 | 302 | 102 |
| 112 | 71 | 311 | 102·5 |
| 117 | 71 | 326·5 | 103 |
| 121·5 | 75·5 | 338 | 103·5 |
| 129 | 79 | 346 | 106·5 |
| 137·5 | 83 | 349 | 104·5 |
| 147 | 86 | 358 | 107 |
| 153·5 | 86·5 | 365 | 110·5 |
| 163·5 | 91·5 | 373 | 112·5 |
| 176 | 95·5 | 380 | 117 |
| 248 | 98 | 387 | 121·5 |
| 254 | 99 | 398 | 121·5 |
| 261 | 99 | 406 | 123·5 |
| 273 | 100·5 | 407 | 125·5 |

$\Delta T$ was the same whether the sample was cooled or heated, with exception of transition III where in a cooling experiment a somewhat broader peak was found (marked (*) in the table) than in a heating experiment. In view of the very-slow drift of temperature in these experiments it seems very plausible to assume that $\Delta T$ is an intrinsic property of the substance under investigation.

TABLE II

Characteristics of critical points of $BaTiO_3$
(For explanation of symbols see text)

| Transition | $T_\uparrow$(°K) heating | $T_\downarrow$(°K) cooling | $\Delta E$ (Joule Mole$^{-1}$) heat of transition | $\Delta T$(°K) half-width |
|---|---|---|---|---|
| I | 207 | 193 | 60 | 3 |
| II | 287 | 280 | 65 | 3 |
| III | 392 | 390·5 | 195 | 1; 2* |

## 5. Remarks

*Specific heat of the lattice.* The curve in fig. 2 shows a saturation of about 12 $R$ at room temperature, which is 3 $R$ less than the Dulong and Petit value. At higher temperatures, however, $C_p$ increases again and reaches 15 $R$. Obviously a single Debye function is not sufficient to describe the whole curve. Probably a sum of a Debye function

$$12\, R\, f_{\mathrm{D}}\left(\frac{\Theta_{\mathrm{D}}}{T}\right)$$

and Einstein functions would fit better. $\Theta_{\mathrm{D}}$ will be of the order of 400 °K. For an exact determination of the characteristic temperatures in the functions, measurements ought to be carried out at lower, respectively higher temperatures than those taken here.

*The $C_p - C_v$ correction.* From thermodynamical reasoning it follows that

$$C_p - C_v = TV.\frac{{a_v}^2}{\beta},$$

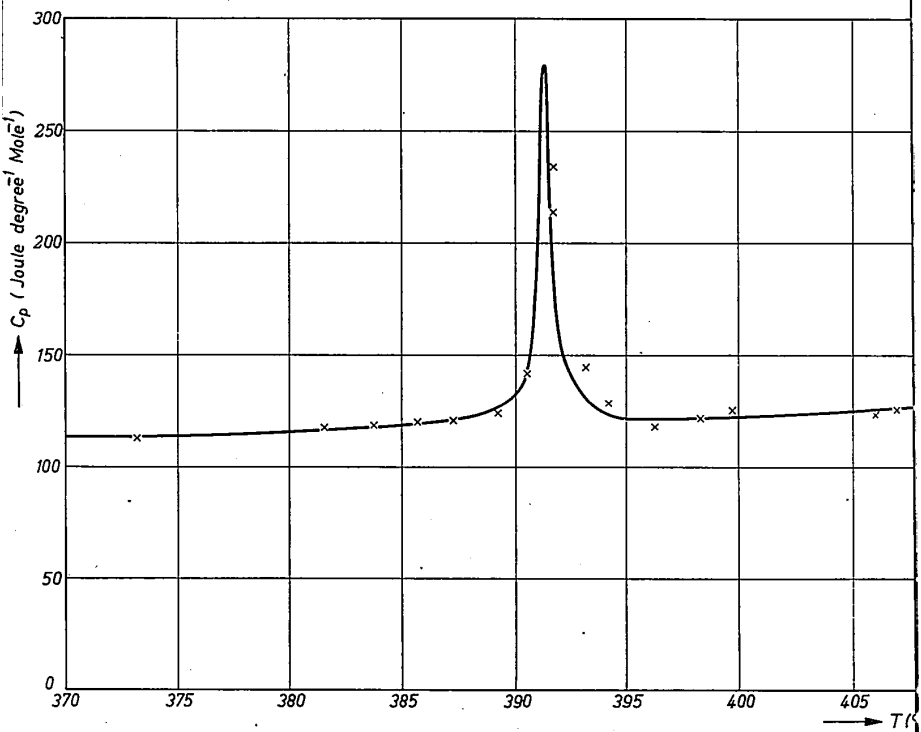



Fig. 4. $C_p$ calculated from the reciprocal drift of temperature by slow heating. Transition III.

where $V$ is the molar volume, $a_v$ the coefficient of volume expansion and $\beta$ the compressibility of the substance. From Mason's [8]) data $\beta$ is readily found to be

$$\beta = 1{\cdot}25.10^{-12} \text{ cm}^2 \text{ dyne}^{-1}.$$

As the molar volume is 42 cm$^3$ and $a_{\text{lin}}$ about $1.10^{-5}$ degree$^{-1}$ [6]),[9]) one gets at room temperature

$$C_p - C_v \approx 1 \text{ Joule degree}^{-1} \text{ Mole}^{-1},$$

which is very small in comparison with the measured value of $C_p$. At higher temperatures the correction will be somewhat larger, but as no exact values of $\beta$ are available in this higher range no calculation has been made.

*The anomalies.* The value $\Delta E$ at the transitions II and III are in accordance with those given by Blattner, Känzig and Merz [5]) (71 respectively 196 Joule Mole$^{-1}$). By heating their sample these authors found in region II a critical temperature of about 5 °C, which value, however, agrees with $T_{\downarrow}$ rather than with $T_{\uparrow}$ of table II. They do not find peaks as sharp as those shown in fig. 3 or fig. 4, but this may be accounted for by the fact that their sample was a powder. Harwood et al. found the Curie point at $T_{\uparrow} = 398$ °K which is 6 degrees higher than that given by Blattner et al. [5]), and that given in this paper. There exists also a considerable difference with the results in the papers cited at [3]) and [6]).

Mr W. Bol of this laboratory carried out some *electrical measurements* at 5 kc/s. When heating the sample he found a maximum of $\varepsilon$ at 393 °K and of tan $\delta$ at 388 °K. Cooling it he found a maximum of $\varepsilon$ at 391·5 °K and of tan $\delta$ at 386·5 °K.

*Acknowledgement.* The author is very grateful to Mr J. H. Aalberts for his valuable help during the measurements and to Dr G. H. Jonker and Dr J. H. van Santen for their interest in this investigation.

*Eindhoven, June 1951*

REFERENCES

[1]) H. F. Kay, H. J. Wellard and P. Vousden, Nature **163**, 636-637, 1949; H. F. Kay and P. Vousden, Phil. Mag. **40**, 1019-1040, 1949.
[2]) W. Merz, Phys. Rev. **75**, 687, 1949.
[3]) B. Vul, J. Phys. U.S.S.R. **10**, 95-106, 1946.
[4]) M. G. Harwood, P. Popper and D. F. Rushman, Nature **160**, 58-59, 1947.
[5]) H. Blattner, W. Känzig and W. Merz, Helv. Phys. Acta **22**, 35-65, 1949.
[6]) S. Sawada and G. Shirane, J. Phys. Soc. Japan **4**, 52-56, 1949.
[7]) See e.g. H. S. Carslaw and J. C. Jaeger, Conduction of heat in solids, Clarendon Press, Oxford, 1947, p. 175.
[8]) W. P. Mason, Phys. Rev. **74**, 1134-1148, 1948.
[9]) H. D. Megaw, Experientia **2**, 183-184, 1946.

# APPLICATION OF MULTI-HOLE COUPLING TO THE DESIGN OF A VARIABLE AND CALIBRATED WAVEGUIDE ATTENUATOR AND IMPEDANCE

by W. J. van de LINDT 621.392.26

**Summary**

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.

**Résumé**

Une brève introduction concernant le comportement de deux guides d'ondes couplés par $n$ éléments directifs identiques équidistants est suivie d'une description d'un nouveau type d'atténuateur variable calibré et d'une impédance variable, à même de modifier indépendamment l'amplitude et la phase du coefficient de réflexion dans toute l'étendue du diagramme de réflexion.

**Zusammenfassung**

In einer kurzen Einleitung wird das Verhalten zweier Hohlleiter behandelt, die durch $n$ äquidistante identische Richtungselemente gekoppelt sind. Es folgt die Beschreibung eines neuen Typs eines geeichten veränderbaren Dämpfers und einer veränderbaren Impedanz, die imstande ist die Amplitude und die Phase des Reflexionskoeffizienten unabhängig von einander zu ändern und die den ganzen Bereich des Reflexionsdiagramms bestreicht.

## 1. Introduction

Consider two infinitely long identical cylindrical waveguides having parallel axes and part of their cross-sectional boundaries coincident. In their common wall a narrow infinitely long slot is cut. Let the initial situation consist of a progressive wave in one of the waveguides; it has been shown[1])[2]) for certain configurations that, in passing down the guide, power is gradually transferred through the slot to the other guide. Once all the power is present as a progressive wave in the other guide the process reverses and power leaks back into the first guide. This phenomenon is sometimes denoted by "spatial beats" [2]).

Whereas in the above situation the interaction between the two guides is achieved by distributed coupling, spatial beats can also be generated by using discrete coupling elements subject to the restriction that each element causes directional coupling.

We will confine our attention to a system made up of two identical waveguides of rectangular cross-section having one of their wide walls

in common. A number of directional coupling elements are located in this common wall. These coupling elements can have a variety of forms such as one or more holes or slots. For weak coupling many examples of the construction of directional coupling elements can be found in the literature[3]). For our purpose it is essential that the elements couple in a forward direction, i.e., the secondary wave caused by an incident wave in the other guide must have the same direction of propagation as the exciting wave.

Without going into detailed theory we derive the relation between the waves in the two guides. To that purpose assume an incident wave in the lower guide of amplitude $E$. In passing the first directional-coupling element this wave will split up in an attenuated wave in the lower guide and a wave in the upper guide *). In the first place conservation of energy requires that the sum of the squares of the amplitudes of the emergent waves equals $|E^2|$, whilst secondly it is clear, by regarding the coupling element as a lumped element, that the vectorial sum of the amplitudes of the emergent waves must equal the amplitude of the incident wave. In this way it is seen that the amplitudes of the three waves must form the sides of a rectangular triangle.

The situation for the next coupling elements is slightly more complex as there are now incident waves in both the upper and lower guides. These two waves bear a definite relationship however. By taking this into account one can construct the vector diagram given in fig. 1. This diagram is valid for systems with identical coupling elements. The phase lag common to waves both in upper and lower guides has been omitted. The amplitude of the waves in lower and upper guides are denoted by $\alpha$ and $\beta$ respectively. These two vectors are joined by a diameter of the circle of radius $\frac{1}{2}E$ making an angle $2na$ with the horizontal diameter, which is given in phase and magnitude by the originally incident wave in the lower guide. Further

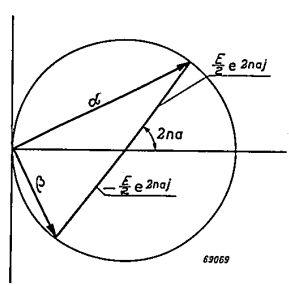



Fig. 1. Diagram showing the relation between various waves in a directional coupler.

*) A directional coupler of the type under discussion is matched, i.e., there is no reflected wave, as can be shown by symmetry considerations. [4])

$n$ is the number of coupling elements passed, and $2a$ is a quantity determined by the coupling strength of one directional element. Starting from the initial condition $\alpha = E$, $\beta = 0$ one can picture, in travelling down the guides, the diameter connecting $\alpha$ and $\beta$ rotating counter-clockwise over an angle $2a$ for each coupling element passed. It is seen in this way that, as the number of elements $n$ passed increases, $\alpha$ diminishes and $\beta$ grows. For $2na = \pi/2$ the waves in the upper and lower guides have equal amplitudes; for $2na = \pi$ all incident power has been transferred to the upper guide, i.e. $\alpha = 0$, $\beta = E$. Still more coupling elements will make $\beta$ decrease in its turn and $\alpha$ increase.

Analytical expressions for $\alpha$ and $\beta$ can be read from the diagram, resulting in

$$\begin{aligned} \alpha &= \tfrac{1}{2}E(1 + e^{2naj}), \\ \beta &= \tfrac{1}{2}E(1 - e^{2naj}). \end{aligned} \tag{1}$$

## 2. Calibrated variable attenuator

A coupler with many directive elements may be used as the basic unit for a calibrated variable attenuator. The number of elements $n_0$ is chosen so that power incident at one terminal is completely transferred to the opposite terminal of the other guide. The condition for this case was found to be

$$2n_0a = \pi . \tag{2}$$

Two terminals are fitted with matched loads and a movable absorbing wedge is introduced in one of the guides. The arrangement is schematically shown in fig. 2.

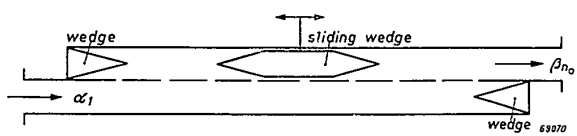



Fig. 2. Schematical cross-section of the attenuator.

To calculate the dependence of the emergent wave on the position of the wedge we will make the simplifying assumption that the absorbing wedge is confined to the space between two coupling elements (fig. 3). To the left of the wedge there are $n$ elements, to the right $(n_0 - n)$.

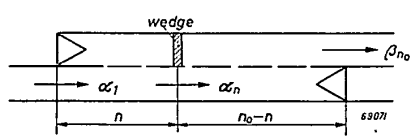



Fig. 3. Simplified diagram illustrating the various quantities relevant in the calculation.

The wave $\beta_{n0}$ orginates from the wave $a_n$ incident on a section of coupler with $(n_0 - n)$ elements.

Application of eqs (1) easily yields the following results:

$$a_n = \tfrac{1}{2} E \left[1 + e^{2jan}\right],$$
$$\beta_{n0} = \tfrac{1}{4} E \left[1 + e^{2jan}\right] \left[1 - e^{2ja(n_0-n)}\right].$$

Using eq. (2), we have

$$\beta_{n0} = \tfrac{1}{4} E \left[1 + e^{2jan}\right] \left[1 + e^{-2jan}\right] = E \cos^2 na ,$$

which gives

$$\text{attenuation} = 40 \log (\cos na) \text{ db} . \qquad (3)$$

An experimental model has been built for 3·3 cm. The coupling elements were constructed according to Riblet and Saad [5]). Each consists of a combination of a transverse and a longitudinal slot as shown in fig. 4. The particular advantage of this system lies in the fact that, for slots small compared to the wavelength, the derivative of the first-order approximations of $a$ with respect to wavelength vanishes, if the dimensions are chosen so as to make the element directional.

If the wedge is as short as has, for simplicity, been assumed for the derivation of eq. (3) the curve relating attenuation to the position of the wedge would be of a discontinuous nature. To avoid this the wedge has been made long enough to extend over several coupling elements. As it was found that a single wedge in the middle of the waveguide still gave discontinuities because of direct interaction with the transverse slots, two flat wedges were mounted on a carriage in a position as indicated in fig. 4.

The instrument was calibrated; the resulting curve is reproduced in fig. 5 where a few points calculated according to (3) also are represented.

A few additional remarks may be of interest. If the movable and terminal wedges are all matched the instrument does not cause any reflection of the incident signal, irrespective of the position of the wedge.

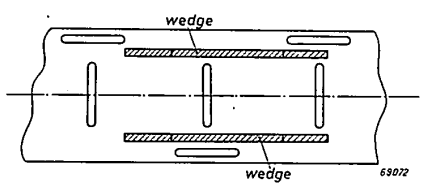



Fig. 4. View of the wall common to the guides, illustrating the position of the absorbing wedges relative to the slots.

The maximum standing-wave ratio measured was below 1·03 with an exception for one position of the wedge, where it rose to 1·08, probably owing to inaccurate construction.

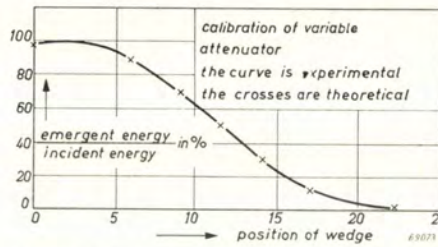



Fig. 5. Calibration curve of the attenuator.

The minimum attenuation is about 0·5 db; the maximum attenuation solely depends on the attenuation of the wedges used. In our case the maximum attenuation was 40 db.

The complete instrument is shown in fig. 6.

### 3. Variable impedance

Consider a system as sketched in fig. 7. Each guide is closed on the right-hand side by a short-circuiting plunger while on the left-hand side the lower guide is terminated by a matched load.

The number of coupling elements is chosen according to the relation

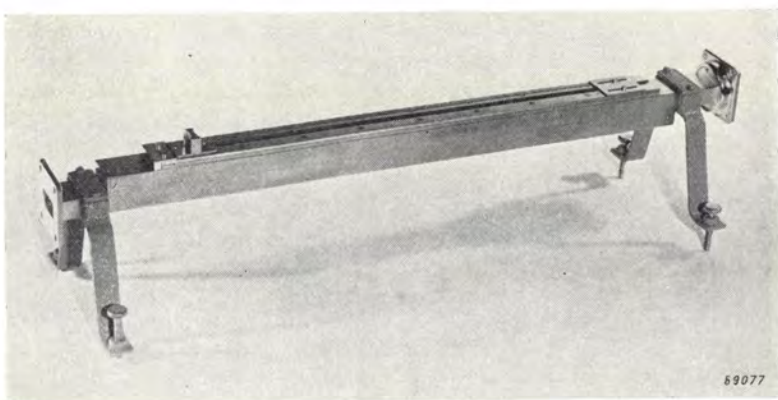



Fig. 6. Calibrated waveguide attenuator.

$2na = \pi/2$, so that at the right-hand side the incident power has been divided equally between upper and lower guides. Furthermore, we assume the pistons to be perfectly reflecting, so that

$$\gamma_n = g_1 a_n ,$$
$$\delta_n = g_2 \beta_n , \; |g_1| = |g_2| = 1 .$$

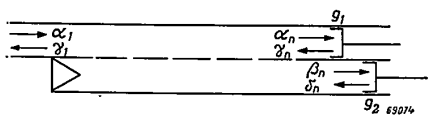



Fig. 6. Schematical cross-section of the variable impedance, showing the various quantities of interest.

Now it is possible to calculate the amplitude of the wave reflected to the left in the upper guide and so the reflection coefficient. Here only the result is given which is

$$f = \frac{\gamma_1}{a_1} = \frac{1}{2j}(g_2 - g_1) = e^{2j\vartheta} \sin 2\xi , \qquad (4)$$

in which $f$ is the reflection coefficient, $\vartheta$ the phase of the plane midway between the two short circuits with respect to a reference plane, and $2\xi$ the angular distance between the two short circuits (see fig. 8).

Equation (4) tells us that by varying $\vartheta$ and $\xi$ we can obtain all possible values of $f$ but, which is far more important, we can vary the absolute value of $f$ and its phase independently, as the former only depends on $\xi$ and the latter on $\vartheta$.

To obtain this in practice we have to provide two controls to the short-circuiting plungers: one which changes their relative position, keeping their midplane constant, and another which can shift the two plungers together without changing their relative distance; a construction that can be quite easily realized.

The general law that a device, capable of transforming a matched load

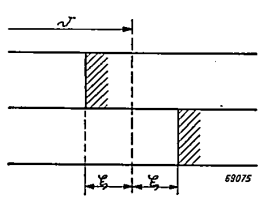



Fig. 8. Diagram showing the coordinates used to fix the plunger positions.

into any impedance, is also capable of transforming any given impedance into any other one, is of course also valid in this case. Therefore by replacing the matched load by an output terminal the instrument can be used as a matching transformer.

In fig. 9 a schematic diagram is drawn of the controls of the plungers used.

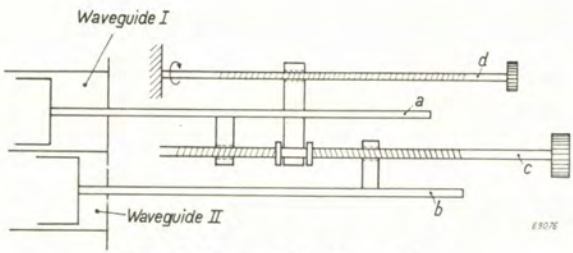



Fig. 9. Outline of the plunger drive.

The two piston rods *a* and *b* can be moved in opposite directions, by means of *c* which is provided with left-hand and right-hand screw-thread. The midplane is unaffected by *c*.

With *d* it is possible to move the rods *a* and *b* in the same direction without changing their relative distance.

As was shown above, with *d* only the phase of the reflection coefficient is changed and with *c* only the amplitude.

The complete instrument is shown in fig. 10.

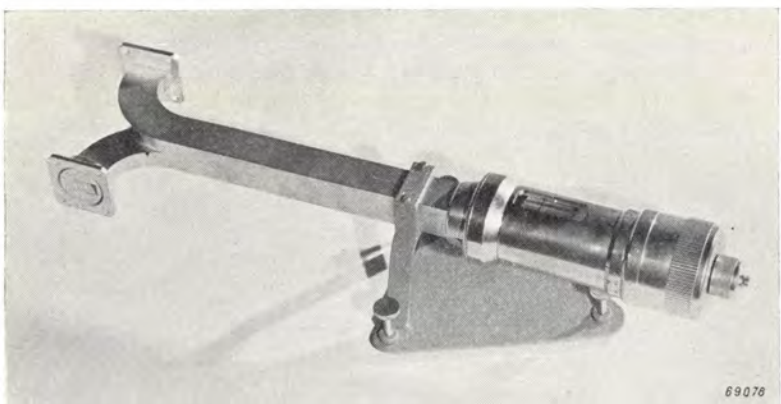



Fig. 10. Variable calibrated impedance.

The author is indebted to Mr Pannenborg of this laboratory for many helpful suggestions concerning the construction of the experimental models.

## REFERENCES

[1]) C. G. Montgomery, Technique of microwave measurements, McGraw-Hill, New York, 1947, pp. 885-890.
[2]) P. E. Krasnushkin and R. B. Khoklov, Spatial beating in coupled wave guides, J. tech. Phys. **19**, 931-942, 1949 (in Russian).
[3]) N. Surdin, Directive couplers in waveguides, J. Inst. elect. Engrs **93**, Part III A, 725-736, 1946.
[4]) C. G. Montgomery, R. H. Dicke and E. M. Purcell, Principles of microwave circuits, McGraw-Hill, New York, 1948, p. 437.
[5]) H. J. Riblet and T. S. Saad, A new type of wave-guide directional coupler, Proc. Inst. Radio Engrs **36**, 61-64, 1948.

*R 182* *Philips Res. Rep. 7, 36-44, 1952*

# OPTIMUM GEOMETRY OF MICROWAVE AMPLIFIER VALVES

by G. DIEMER and K. RODENHUIS

621.385.1.029.64

**Summary**

On the basis of Van der Ziel and Knol's theory of feedback amplifiers it is shown that for u.h.f. amplifier valves a maximum frequency limit for the amplification is reached if the electrode areas are so chosen that the useful capacitance equals the unavoidable parasitic capacitance. For optimum gain-bandwith products the useful capacitance should be somewhat higher.

**Résumé**

En se basant sur la théorie de Van der Ziel et Knol sur les amplificateurs à contre-réaction, on montre que pour les tubes amplificateurs u.h.f. la limite de fréquence maximum pour l'amplification est atteinte lorsque les surfaces des électrodes sont choisies de telle façon que la capacité utile soit égale à la capacité parasite inévitable. Pour le produit gain-largeur de bande optimum, la capacité utile doit être légèrement plus élevée.

**Zusammenfassung**

Auf der Grundlage der Van der Ziel und Knolschen Theorie hinsichtlich der Rückkopplungsverstärker wird gezeigt, daß für UHF-Verstärkerröhren ein Frequenzmaximum für die Verstärkung erreicht wird, wenn die Elektrodenflächen so gewählt sind, daß die nützliche Kapazität gleich ist der unvermeidlichen parasitären Kapazität. Bei Geräten mit optimaler Verstärkungsbandbreite ist eine etwas höhere nützliche Kapazität erforderlich.

## 1. Introduction: The ohmic losses form the decisive factor for the frequency limit of a microwave amplifier

When a microwave amplifier is used at higher frequencies there is ultimately a loss in gain. Rodenhuis [1]) has shown that for a well-constructed microwave triode with pin leads, the frequency limit where the power gain drops below 1 is only slightly influenced by transit-time effects and is mainly determined by the ohmic losses of the leads. This is in fact the main reason why this frequency limit is raised so considerably when disc seals are used. But even with disc-seal triodes it is still the ohmic damping that determines this frequency limit, as is proved by a calculation analogous to that given by Rodenhuis. This is illustrated in fig. 1, where curves are given for the factor $K$, which is a criterion for the frequency limit; $K$ is negative for a gain $G>1$, zero for $G=1$, and positive for $G<1$. The factor $K$ is given as a function of frequency for three cases: (1) a commercial triode with pin leads, (2) a triode with the same electrode system but with ten-times-lower

ohmic losses (this is approximately the improvement to be expected with disc seals), and (3) the same electrode sytem combined with ohmic losses equal to zero.

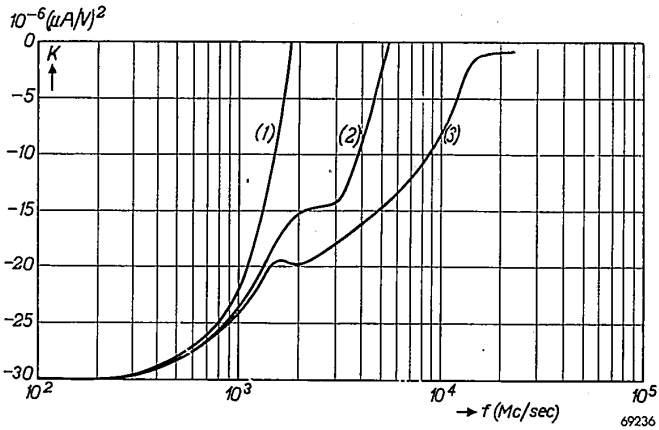



Fig. 1. The $K$-factor of Rodenhuis for an oscillator as a function of frequency, for three different cases:
(1) commercial triode with pin leads and electrode system mounted in mica (EC 81), (2) the same electrode system with the ohmic losses of the leads ten times smaller, (3) the same electrode system with ohmic losses equal to zero. The valve is able to oscillate when $K < 0$.

Morton and Ryder [2]) have analysed a microwave triode amplifier and tried to dimension it, as regards electrode distances and current density, such that the product of power gain and bandwidth is maximum. In doing so, however, they neglected the influence of ohmic losses and did not consider the optimal electrode area. Here we shall investigate how the maximum obtainable gain depends upon the electrode area when taking ohmic losses into account, thereby assuming that certain, favourable, values have been chosen for the clearances between the electrodes and for the current density.

## 2. Qualitative consideration of the optimal electrode area taking ohmic losses into account

At high frequencies the power gain $G$ that can be obtained with optimum matching depends upon the h.f. slope $S$ (in general this has a complex value), the input conductance $g_{11}$, the output conductance $g_{4}$ and the feedback admittance $Y_{14}$, which is mostly mainly reactive and may be capacitive or inductive. Generally, if the feedback is not too great, the power gain will be inversely proportional to the conductances and in direct proportion to the square of the absolute value of the slope. (The influence

of the feedback will be taken into account in the more general theoretical deduction to follow.) This relation can be expressed in the following way:

$$G \backsim \frac{|S|^2}{g_1 g_4}. \tag{1}$$

We shall now consider the influence of the electrode area upon the gain. To keep the discussion simple, this will be done for a conventional u.h.f. valve with a sleeve cathode. The electrode area will be varied by varying the length of the electrodes, keeping all other factors constant. In that case both the electrode capacitances and the slope are linear functions of the length of the electrodes. Part of the length of the electrodes is used for producing the anode current, whilst the remaining part is not active although it has a certain capacitance. It is assumed that when varying the length of the electrodes only the active part is varied, the parasitic part being constant. We suppose the current density to be constant, so that the slope is directly proportional to the active part. If the active part is reduced to zero the slope becomes zero too, whilst there remains a finite capacitance. For the frequencies considered the conductances occurring in eq. (1) depend mainly upon the ohmic losses due to the capacitive currents in the leads. The resulting conductances are, therefore, in direct proportion to the square of the total valve capacitances. Thus from eq. (1) it is clear that when reducing the electrode length $G$ ultimately becomes zero (zero slope, but $g_1$ and $g_2$ positive). On the other hand, for very long electrodes the gain will decrease with increasing length, since $|S|$ is then proportional to the length and both $g_1$ and $g_2$ are proportional to the square of the length.

From this it follows that there is a certain finite length giving optimal gain.

### 3. Quantitative theory for the amplification

To investigate more quantitatively this relationship between gain and the size of the system, let us start from the general formula for the power gain of an active four-pole with feedback in the case of optimum matching and tuning, as given by Van der Ziel and Knol [3]). In a slightly modified form (see Van der Ziel and Knol's formula (22)) this reads:

$$G = G(\delta) = \frac{g_3}{g_2}\delta(1+\beta_2)^2[2-\delta(1-\beta_1\beta_2)+2\sqrt{1-\delta(1-\beta_1\beta_2)-\tfrac{1}{4}\delta^2(\beta_1+\beta_2)^2}]^{-1}, \tag{2}$$

where $g_1$ = input conductance ($=1/R_1$ in Van der Ziel's paper), $g_2$ = feedback conductance ($=1/R_2$ ibid.), $g_3$ = forward conductance ($=1/R_3$ ibid.), $g_4$ = output conductance ($=1/R_4$ ibid.), $\beta_1$ = tangent of the phase angle

of the feedback admittance, $\delta_2$ = tangent of the phase angle of the forward admittance, and

$$\delta = \frac{g_2 g_3}{g_1 g_4}. \tag{3}$$

We now have to investigate how the quantities occurring in the right-hand side of eq. (2) depend upon the electrode area. To this end we distinguish between the active electrode area $O_x$ and the passive electrode area $O_0$. By the active area is meant that portion of the electrodes which takes part in the flow of electron current. For a plan-parallel configuration (to which we shall confine ourselves here) the capacitance associated with the active area amounts to $C_x = \varepsilon_0 O_x/d$, where $d$ is the clearance between the electrodes. The capacitance associated with the passive area will be called the parasitic capacitance $C_0$.

For a well-designed u.h.f. valve one may assume that there is no appreciable amount of resistance or inductance common to the input and the output circuits [4]). In that case $g_2$ and $g_3$ both only depend upon the electron current and thus are proportional to the active cathode area. Therefore the factor $g_3/g_2$ in eq. (2) can be regarded as a constant.

$\beta_1$ and $\beta_2$ are the ratios of the imaginary to the real part of the feedback admittance and the forward admittance respectively. The imaginary part consists of an electronic contribution which is proportional to $O_x$ and a capacitive contribution which in a well-screened valve will also be nearly proportional to $O_x$. It may therefore be assumed that $\beta_1$ and $\beta_2$ are also constant.

Thus according to eq. (2) the gain is a function of $\delta$ only, while $\delta$ is in turn a function of the electrode areas. To find the latter function we consider first the dependency of the valve capacitancies upon the electrode dimensions. It has already been seen that part of the capacitancies varies in direct proportion to the active area, another part is independent of this area and still another part is dependent upon the circumference of the active area (the last part will be called the fringing capacitance).

This can be expressed by the relation

$$C \approx 1 + ax + bx^2 \tag{4}$$

($x$ being the square root of the electrode area), which holds for the following two examples:

(1) Scaling of the electrodes (fig. 2*a*).

In this case $ax$ stands for the fringing capacitance and $bx^2$ for the useful capacitance, whilst 1 stands for the constant part of the parasitic capacitance.

(2) Extending the electrodes in one direction only (fig. 2*b*).

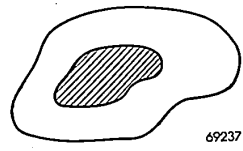



Fig. 2*a*. Scaling of the electrodes.

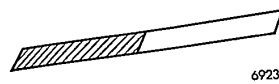



Fig. 2*b*. Varying the useful electrode area by varying the length of the electrode system.

Then the increase of the circumference is proportional to the area; in this case the coefficient $a$ in eq. (4) equals zero.

The conductances $g_1$ and $g_4$ occurring in eq. (3) can be split up into two parts:

(*a*) the part due to the electron currents, which is proportional to $x^2$,

(*b*) the part due to ohmic losses which is proportional to the square of the valve capacitances [1]).

Thus $\delta$ is a function of the following shape:

$$\delta = \frac{g_2}{g_1} \cdot \frac{g_3}{g_4} = \delta_1 \frac{x^2}{(1 + a_1 x + b_1 x^2)^2 + c_1 x^2} \cdot \frac{x^2}{(1 + a_2 x + b_2 x)^2 + c_2 x^2}, \qquad (5)$$

where $\delta_1$ is a constant, independent of $x$. It is assumed that the series resistances do not depend upon the size of the area (see below). Hence $\delta/\delta_1$ is a product of two fractions, each having a maximum as a function of $x$. The maximum of the first fraction, or the minimum of its reciprocal, lies at

$$\frac{\mathrm{d}}{\mathrm{d}x}\left\{\left(\frac{1}{x} + a_1 + b_1 x\right)^2 + c_1\right\} = 0\,,$$

thus at

$$x = x_1 = 1/\sqrt{b_1}\,, \qquad (6a)$$

and that of the second fraction at

$$x = x_2 = 1/\sqrt{b_2}\,. \qquad (6b)$$

When substituting eq. (5) in eq. (2) $G$ can be regarded as a function of $x$. The maximum of this function will occur at

$$\frac{\mathrm{d}G}{\mathrm{d}x} = \frac{\mathrm{d}G}{\mathrm{d}\delta} \cdot \frac{\mathrm{d}\delta}{\mathrm{d}x} = 0\,. \qquad (7)$$

In the appendix it will be shown that $\mathrm{d}G/\mathrm{d}\delta$ is always positive, provided

the phase angles of the forward and the feedback admittances are not so large that instability might occur (see Van der Ziel and Knol, p. 172). Therefore, according to (7), $G$ has only one maximum, which lies at the same value of $x$ where $\delta$ has its maximum.

Returning to eqs (6*a*) and (6*b*), we remark that at the value $x = x_1$ the useful capacitance on the input side equals that portion of the parasitic capacitance which is independent of the size of the area. At $x = x_2$ the same holds for the output side. The maxima of $\delta$ and $G$ are all the sharper the smaller the values of $c_1$ and $c_2$ compared with the other constants in (6*a*) and (6*b*) (i.e. the more ohmic impedances predominate, which is mainly the case at very-high frequencies). If in the latter case the maxima coincide, $G$ is reduced by 25% when $x$ is removed a factor 2 from the optimal value.

Since $\delta$ is a product of two fractions which are both everywhere positive, the maximum of $\delta$ and thus of the power gain $G$ lies generally at a value of $x$ between $x_1$ and $x_2$, and it can be shown that for practical values of the constants occurring in (3) one has to take approximately the geometrical average.

Although, as already stated, the maximum in $G$ is not very pronounced, the microwave-valve constructions commonly employed are often so far removed from this optimal geometry that their power gain is much lower than what would be obtainable with an optimal system. This may be due to the fact that with the small effective electrode areas to which the optimum-geometry considerations lead, formerly only valves of a small power output could be constructed, since no cathode was available which was suitable for high loads. With the new L-cathode [5]), however, it is now possible to construct valves having a very small effective electrode area and still yielding a reasonable microwave output (see section 4). Therefore the condition for optimum gain mentioned above can be realized. (Naturally choosing the valve construction such that the parasitic capacitance is reduced to the lowest possible value).

A further advantage of these small valve capacitances is that they make it easier to minimize the passive feedback arising from a small self-inductance in the grid lead of grounded-grid amplifiers, which at microwaves may be very troublesome [4]).

Something remains to be said about the assumption that the series resistance $r_s$ does not depend upon the electrode area. Such is indeed the case in valve systems where the electrodes are mounted in mica and connected to external leads via pins fused in the glass base. In the case of disc-seal valves $r_s$ might be slightly reduced for larger electrode areas owing to parts of the leads being of larger diameter. In practice it appears that with such valves an important part of the serie resistances is independent of the useful electrode area (e.g. the contact resistances). It can be shown that in this

case the maximum value of $G$ lies at a somewhat larger ratio of the useful capacitance to the parasitic capacitance. This ratio may be at most about equal to 2.

## 4. Bandwidth considerations

Practice has already proved that with valves the electrode area of which is dimensioned for optimal gain a considerable bandwidth can be obtained, sufficient for nearly all practical applications [6]). When, however, one takes a certain reduction of $G$ into the bargain one may try to make the "figure of merit" (bandwidth-quality figure) maximum. For this, in the case of pentodes, mostly the product of voltage gain and bandwidth $B$ is taken, but in the case of grounded-grid triodes it is better [2]) to take the product of power gain $G$ and bandwidth $B$. It can be shown that the maximum of the figure of merit is obtained with a somewhat greater electrode area than gives the maximum gain. For the triode this will be shown from formula (35) of Van der Ziel and Knol's paper, which in our notation reads:

$$BG = \frac{1}{\pi}(1+\beta_2{}^2)\frac{g_3}{g_2}\cdot\frac{g_4}{c_4}\left[1/\delta\left\{\frac{2/\delta-(1-\beta_1\beta_2)}{\sqrt{1/\delta^2-(1/\delta)(1-\beta_1\beta_2)-\frac{1}{4}(\beta_1+\beta_2)^2}}+2\right\}\right]^{-1}, \quad (8$$

where $c_4$ represents the total output capacitance. This formula is valid if the bandwidth $B$ of the network is entirely governed by the bandwidth of the output system as is the case in grounded-grid amplifiers.

It can be proved that for the region of stable operation the expression between square brackets is a monotonically increasing function of $\delta$, which attains its maximum value for the same value of the electrode area as does the gain $G$, i.e., at $x = x_m$ (see section 3). (The mathematical calculations, which are simple but somewhat cumbersome, are omitted here.) Of the other factors in the right-hand side of eq. (8) only $g_4/c_4$ is dependent on the electrode area, thus on $x$. It is clear that $g_4/c_4$, being proportional to $(1 + a_2x + b_2x^2)$, in the neighbourhood of $x=x_m$ is still an increasing function of $x$; so the maximum of $BG$ is shifted towards $x$ values larger than $x_m$. Calculations for practical valves have shown that this shift is not more than about a factor 2.

## 5. Practical examples

The object of the foregoing considerations has only been to find out the order of the optimum electrode area (the limits are not very sharply defined). When dealing with practical examples we shall leave out of consideration the influence of the fringing capacitance, which is often difficult to estimate, and any non-useful capacitance will be regareded as parasitic capacitance.

(1) Experiments with a *push-pull pentode* of the type EFF 50 (having pin leads) showed that at a wavelength of 1 m the gain reaches maximum at a very definite length of the electrode system, which was much smaller than that of the existing EFF 50. For these models the ratio of useful to parasitic capacitance had the following values.

(a) *Normal* EFF 50. Useful length of the cathode 18 mm, output capacitance 5·5 pF, input capacitance 9·4 pF, ratio of useful to parasitic capacitance on the input side 2 and on the output side 4, product 8. Power gain with optimum matching about 2·5.

(b) *Optimum system* 6 EFF. Useful cathode length 6 mm, output capacitance 2·5 pF, input capacitance 5·3 pF, ratio of useful to ballast capacitance on the input side 0·65 and on the output side 1·5, product 0·95. Power gain with optimum matching about 10.

These data agree well with what has been stated in sections 2 and 3. It is to be noted, however, that in these grounded-cathode pentodes a part of the input impedance is caused by the self-inductance $L$ of the cathode lead; this impedance is approximately equal to $\omega^2 L C_x S$ and thus proportional to the square of the useful capacitance; this causes a shift of the maximum in $\delta$ and in $G$ towards somewhat smaller values of $x$.

(2) *An experimental grounded-grid triode* with secondary emission [7]) had an electrode system comparable with that of the 6 EFF as far as cross-section through the system and clearances are concerned, but the parasitic capacitance was reduced by a suitable construction of the leads. The optimum useful length of the electrode system was about 3·5 mm, i.e., indeed somewhat smaller than that of the 6 EFF mentioned above.

(3) *Lighthouse amplifier triode* 49 AA with L-cathode [5]). Useful cathode area 7·07 mm², clearance between cathode and centre of the grid wires 40 μ, $S$ = 13 mA/V, $I_a$ = 30 mA (normal operating point).

The capacitances of the electrode parts inside the bulb which contribute to the microwave tuning are: $C_{ag}$ = 0·57 pF total, useful 0·24 pF, parasitic 0·33 pF; $C_{cg}$ = 2·0 pF total, useful 1·4 pF, parasitic 0·6 pF.

The ratio of useful to parasitic capacitance at the output side is 2·2 and at the input side 0·7 the product being about 1·5, which is fairly close to the optimum value.

This valve is designed for amplification in the microwave region up to some thousands of megacycles/sec. At 7·5 cm (4000 Mc/s) the power gain is about 10. With normal (no wide band) circuits a bandwidth of 50 Mc/s can easily be obtained. The maximum power output at 250 V plate voltage is about 0·5 watt. The passive feedback due to the self-inductance of the grid wires is so low that in this frequency range it practically compensates the feedback due to the static anode-cathode capacitance.

Preliminary experiments show that even at 4 cm wavelength (7500 Mc/s) the power gain is considerably greater than 1, but feedback through the self-inductance of the grid wires plays a considerable part here, so that at optimum matching oscillation easily arises.

## Appendix

We put $F = 1/G$ and $1/\delta = \varepsilon$.
Hence $\mathrm{d}G/\mathrm{d}\delta = (G^2/\delta^2)\ \mathrm{d}F/\mathrm{d}\varepsilon$.
From eq. (2) we have

$$F(\varepsilon) = 2\varepsilon - (1-\beta_1\beta_2) + 2\sqrt{\varepsilon^2 - (1-\beta_1\beta_2) - \tfrac{1}{4}(\beta_1+\beta_2)^2}.$$

Differentiation yields

$$\frac{\mathrm{d}F}{\mathrm{d}\varepsilon} = 2 + \frac{2\varepsilon - (1-\beta_1\beta_2)}{\sqrt{\varepsilon^2 - \varepsilon(1-\beta_1\beta_2) - \tfrac{1}{4}(\beta_1+\beta_2)^2}} = \frac{F(\varepsilon)}{\sqrt{\varepsilon^2 - (1-\beta_1\beta_2) - \tfrac{1}{4}(\beta_1+\beta_2)^2}}.$$

$F(\varepsilon)$ is always positive, and in the region of stable operation, the square root of the denominator is real and positive too (see Van der Ziel and Knol's eq. (18)). Thus $\mathrm{d}F/\mathrm{d}\varepsilon > 0$ and therefore $\mathrm{d}G/\mathrm{d}\delta > 0$, which is what had to be proved.

*Eindhoven, September 1951*

## REFERENCES

1) K. Rodenhuis, Philips Res. Rep. **5**, 46-77, 1950.
2) J. A. Morton and R. M. Ryder, Bell System tech. J. **29**, 496-530, 1950.
3) A. van der Ziel and K. S. Knol, Philips Res. Rep. **4**, 168-178, 1949.
4) G. Diemer, Philips Res. Rep. **5**, 432-434, 1950.
5) H. J. Lemmens, M. J. Jansen and R. Loosjes, Philips tech. Rev. **11**, 341-350, 1950.
6) G. Diemer and K. S. Knol, Philips Res. Rep. **5**, 153-154, 1950.
7) G. Diemer and J. L. H. Jonker, Wireless Engr **27**, 137-143, 1950.

# LANGMUIR'S $\xi, \eta$ TABLES FOR THE EXPONENTIAL REGION OF THE $I_a$-$V_a$ CHARACTERISTIC

by G. DIEMER and H. DIJKGRAAF 537.543(083.5)

**Summary**

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 $\xi$, $\eta$ tables with the voltage gradient at the anode as parameter is given from which the potential distribution in such cases can be derived.

**Résumé**

Les distances entre les électrodes des diodes et triodes modernes pour ondes micrométriques sont souvent si faible que le point de fonctionnement normal se trouve dans la partie exponentielle de la caractéristique. On donne une série de tables $\xi$, $\eta$ avec le gradient de potentiel à l'anode comme paramètre, permettant dans de tels cas de déduire la distribution du potentiel.

**Zusammenfassung**

Die Abstände zwischen den Elektroden moderner Kurzwellen-Dioden und -Trioden sind oft so klein, daß der normale Arbeitspunkt im exponentiellen Teil der Kennlinie liegt. Es werden eine Reihe von $\xi$, $\eta$-Tabellen mit dem Spannungsgradienten an der Anode als Parameter gegeben, aus denen die Potentialverteilung in solchen Fällen abgeleitet werden kann.

For the exponential region of a diode characteristic, when space charge is taken into account, the potential distribution cannot be derived from Langmuir's $\xi, \eta$ tables as supplemented by Kleijnen [1]) and by Van der Ziel [2]). $\xi$ and $\eta$ are dimensionless quantities that describe this potential distribution in the following way:

$$\eta = V/V_T, \tag{1}$$

$$\xi = 9{\cdot}186.10^5\, T^{-3/4}\, I_a^{1/2}\, x\,, \tag{2}$$

where $V$ = potential between anode and cathode, $x$ = distance to anode, $V_T = kT/e = T/11605$, $k$ = Boltzmann's constant, $e$ = electronic charge, $T$ = temperature of the cathode (rationalized Georgi units are used, if not indicated otherwise).

Since hitherto the main emphasis lay on the calculation of the $I_a$-$V_a$ characteristic, there was no need of the $\xi$, $\eta$ tables for the exponential region (anode voltage negative), because in this region the anode current $I_a$ is influenced only by the saturation current $I_s$ and the anode voltage $V_a$, and not by the field distribution between the electrodes:

$$I_a = I_s \exp\,(V_a/V_t)\,. \tag{3}$$

However, for calculating the transit-time functions, high-frequency damping and noise, it is necessary to know the potential distribution between the electrodes in the presence of space charge.

First we want to draw attention to a difficulty which arose when attempting to apply an approximated calculation, as follows: Suppose we have a diode with certain inter-electrode distance $d$, saturation current $I_s$ and cathode temperature $T_c$, which is operated in the space-charge-limited region. The potential distribution betwecn cathode and anode is calculated with the aid of the well-known $\xi$, $\eta$ tables [1]), which are valid as long as a potential minimum is present. As the negative anode voltage increases, the potential minimum on the anode disappears and the calculation no longer holds. We then assume the anode to be at a greater distance, while keeping the applied space-charge-free field the same, such as to investigate whether at a greater distance there still exists a potential minimum and thus be able to calculate the potential distribution from the existing $\xi$, $\eta$ tables. It is thereby assumed that the current, which in the actual case is intercepted by the anode, is so small as to have practically no effect upon the space charge.

In practice, however, it proves that this reasoning hardly ever holds. Fig. 1. gives the maximum negative applied field strength, $F = V_a/d$, at which the potential minimum just falls on the anode, as a function of the distance $d$. For greater negative field strengths there is nowhere a

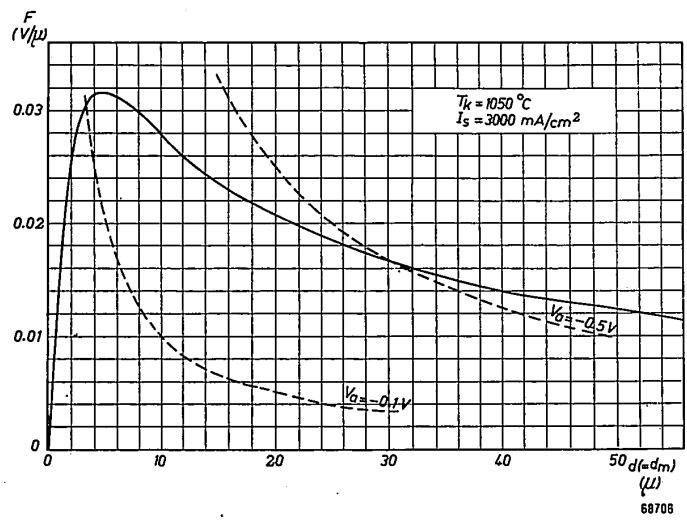



Fig. 1. Maximum negative applied field strength $F$ for the occurrence of a potential minimum, as a function of the distance $d$ between anode and cathode. The cathode temperature $T_c$ and the saturation current $I_s$ have values normal for the L-cathode.

potential minimum. This graph applies for a saturation current $I_s = 3000$ mA/cm$^2$ and $T_c = 1050$ °C (an L-cathode). For further exemplification two other curves are given in fig. 1, showing the relationship between $F$ and the inter-electrode distance $d$ for anode potentials of —0·1 V and —0·5 V. It is seen that with very small and large distances the potential minimum disappears already for small field strengths.

Freeman [3]) has calculated the potential distribution for a plan-parallel diode in a number of cases by means of numerical integration. Briefly, this is carried out in the following way. The Poisson equation can be reduced to a soluble differential equation of the first order for the gradient $\mathrm{d}V/\mathrm{d}x$ of the potential $V$. Given the gradient at one point (if there is a potential minimum this is the value at the minimum, which is equal to zero), then by simple integration it is possible to derive a first-order differential equation for $V$ which, after numerical integration, gives the required relationship between $V$ and $x$. By introducing the variables $\eta$ and $\xi$ mentioned above, in this way a universal relationship can be derived between $\eta$ and $\xi$ from which the potential distribution can be calculated for all special cases. In the exponential region, however, where $\mathrm{d}V/\mathrm{d}x$ is nowhere known, the relation found between $\xi$ and $\eta$ has a parameter that as yet is unknown. Freeman chose for this parameter $(\mathrm{d}\eta/\mathrm{d}\xi)_{\mathrm{an}} \equiv \gamma$, which is a measure for the gradient at the anode in the presence of space charge. He gave a set of $\xi$, $\eta$ curves for different values of $\gamma^2$. If for a particular valve the anode voltage and inter-electrode distance are given, then there is one $\xi$, $\eta$ point known, namely, that corresponding to the cathode. It can then be decided which curve is to be chosen from the set given, so that the $\xi$, $\eta$ relation and thus the potential distribution for that particular case are known.

Unfortunately Freeman's numerical material has not been extended far enough for application to modern microwave diodes and triodes, where it is desired to know the potential distribution for inter-electrode distances of the order of 10 $\mu$ and for negative voltages of 1 or 2 volts on the anode or the grid. We have therefore extended the numerical calculations to $\gamma^2 = 10^5$ and $\eta = 10^2$, the results being compiled in the accompanying tables and graphs, where Freeman's values for $\gamma^2 \leqslant 20$ and $\eta \leqslant 20$ have been included.

With reference to the graphs, it is to be noted that for strongly negative fields the potential distribution does indeed become more and more linear. There is still a bend in the $\xi$, $\eta$ curves, but it is displaced towards higher values of $\eta$; in fig. 3 the area of the bends (framed in fig. 2) is plotted on an enlarged scale.

The region of the bend nearly satisfies the equation

$$\eta = 3\,\xi^{-1/4} \qquad (5)$$

or substituting eqs (1), (2) and (3) in (5):

$$3{\cdot}2.10^2\, T^{-3/4}\, I_s\left(-\frac{V_a}{V_T}\right)\left[\exp\left(\frac{V_a}{4V_T}\right)\right] x^{1/2} = 1\,. \qquad (6)$$

So if we choose the anode voltage $V_a$ and the distance $d$ such that

$$3{\cdot}2.10^2\, T^{-3/4}\, I_s\left(-\frac{V_a}{V_T}\right)\left[\exp\left(\frac{V_a}{4V_T}\right)\right] d^{1/2} < 1\,,$$

we may assume the potential distribution to be fairly well linear. This gives support to the supposition made when calculating the microwave diode conductance in the exponential region of the characteristic, where for negative fields of 1 to 2 volts over distances of 10 to 30 μ the potential distribution was taken to be linear [5]).

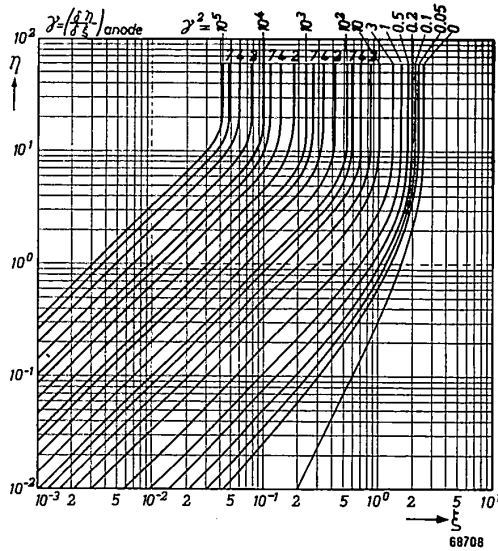



Fig. 2. Langmuir's $\xi$, $\eta$ relationship for a plan-parallel diode with retarding field ($V_a < 0$). The parameter $\gamma = (d\eta/d\xi)_{an}$ is a measure for the field strength at the anode.

*Examples of application*

A plane diode has an inter-electrode distance $d = 10$ μ; $T_c = 1050$ °C and $I_s = 3000$ mA/cm² (the data of an L-cathode [4])).

(a) For the first example let $V_a = -1{\cdot}0$ volt, and let us first calculate $\eta$ and $\xi$ at the cathode:

$$\eta_c = V_a/V_T = 8{\cdot}77\,,$$

$\xi_c = 9{\cdot}186.10^5\ T^{-3/4}\ I_a^{1/2}\ d = 9{\cdot}186.10^5 \times (1323)^{-3/4} \times (0{\cdot}46.10^{-3})^{1/2} \times 10^{-3} = 8{\cdot}99.10^{-2}$. Here $V_T = kT/e$ and $I_a = I_s \exp(V_a/V_T)$.

From fig. 3 it now follows that $\gamma^2 = 8{\cdot}5.10^3$. The potential distribution between cathode and anode is almost linear, as is to be seen on the graph ($\eta$ and $\xi$ decrease towards the anode).

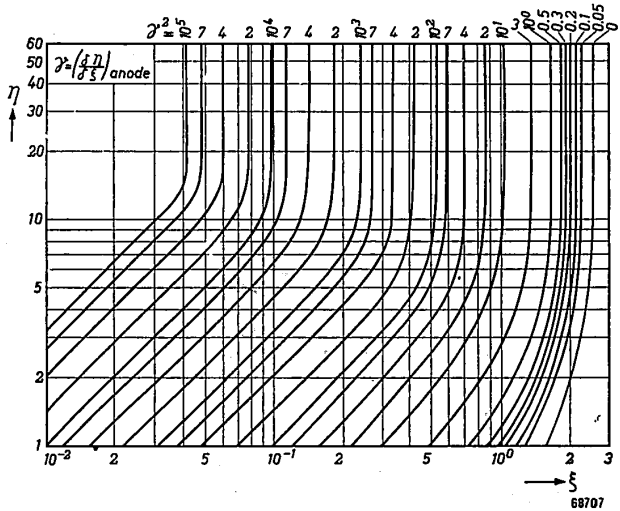



Fig. 3. Enlarged plot of the framed area in fig. 2.

The potential distribution is found by calculating $\xi$ for a certain distance $x$ from the anode and then reading the value of $\eta$ from the corresponding curve in fig. 3, from which the potential at that point is obtained.

For instance, let $x = 8\mu$. Then

$$\xi_x = 9{\cdot}186.10^5 \times (1323)^{-3/4}\ (0{\cdot}46.10^{-3})^{1/2} \times 8.10^{-7} = 7{\cdot}18.10^{-2};$$

$$\eta_x = 6{\cdot}5;$$

$V_x = V_T\ \eta_x = 0{\cdot}74$ V with respect to the anode, or $V_x = 0{\cdot}74 - 1{\cdot}0 = -0{\cdot}26$ V with respect to the cathode. For the field without space charge it would follow that $V_x = -0{\cdot}20$ volt.

(b) For the second example let $V_a = -0{\cdot}3$ volt. Then

$$\eta_c = 1{\cdot}64,$$
$$\xi_c = 9{\cdot}186.10^5 \times (1323)^{-3/4} \times (0{\cdot}125)^{1/2} \times 10^{-3} = 1{\cdot}94.$$

From fig. 3 it follows that $\gamma^2 = 0{\cdot}04$.

The potential is again calculated at $8\mu$ from the anode:

$$\xi_x = \xi_c \, x/d = 1{\cdot}55 \, .$$

From fig. 3 it follows that $\eta = 1{\cdot}37$.
$V_x = 0{\cdot}114 \times 1{\cdot}37 - 0{\cdot}3 = -0{\cdot}114$ V with respect to cathode. For the field in the absence of space charge we find $V_x = -0{\cdot}06$ V.

*Eindhoven, September 1951*

REFERENCES

1) P. H. A. Kleijnen, Philips Res. Rep. **1**, 81-96, 1946.
2) A. van der Ziel, Philips Res. Rep. **1**, 97-118, 1946.
3) J. J. Freeman, J. Res. Bur. Stand. **42**, 75-88, 1949.
4) H. J. Lemmens, M. J. Jansen and R. Loosjes, Philips tech. Rev. **11**, 341-350, 1950.
5) G. Diemer, Philips Res. Rep. **6**, 211-223, 1951.

| $\eta$ \ $\gamma^2$ | $10^5$ | $7.10^4$ | $4.10^4$ | $2.10^4$ | $10^4$ | $7.10^3$ | $4.10^3$ | $2.10^3$ |
|---|---|---|---|---|---|---|---|---|
| $10^{-2}$ | | | | | | | | |
| 2 | | | | | | | | |
| 4 | | | | | | | | |
| 6 | | | | | | | | 0·00133 |
| 8 | | | | | | | 0·00125 | 0·00176 |
| $10^{-1}$ | | | | | 0·00100 | 0·00119 | 0·00156 | 0·00220 |
| 2 | | | 0·00100 | 0·00143 | 0·00200 | 0·00239 | 0·00312 | 0·00440 |
| 4 | 0·00127 | 0·00154 | 0·00200 | 0·00286 | 0·00400 | 0·00480 | 0·00624 | 0·00880 |
| 6 | 0·00189 | 0·00230 | 0·00299 | 0·00430 | 0·00600 | 0·00720 | 0·00934 | 0·0312 |
| 8 | 0·00253 | 0·00308 | 0·00398 | 0·00570 | 0·00800 | 0·00960 | 0·0125 | 0·0176 |
| $10^0$ | 0·00316 | 0·00380 | 0·00495 | 0·00710 | 0·0100 | 0·0120 | 0·0156 | 0·0220 |
| 1·5 | 0·00474 | | | | | | | |
| 2 | 0·00632 | 0·00760 | 0·00990 | 0·0141 | 0·0200 | 0·0240 | 0·0312 | 0·0440 |
| 2·5 | 0·00790 | | | | | | | |
| 3 | 0·00948 | 0·0113 | | | | | | |
| 3·5 | 0·0111 | | | | | | | |
| 4 | 0·0127 | 0·0150 | 0·0197 | 0·0280 | 0·0400 | 0·0479 | 0·0624 | 0·0880 |
| 4·5 | 0·0142 | | | | | | | |
| 5 | 0·0155 | | | | | | | 0·111 |
| 5·5 | 0·0170 | | | | | | | |
| 6 | 0·0185 | 0·0225 | 0·0296 | 0·0421 | 0·0600 | 0·0719 | 0·0936 | 0·130 |
| 7 | | | | | 0·0695 | 0·0840 | 0·1085 | 0·145 |
| 8 | 0·0245 | 0·0300 | 0·0395 | 0·0552 | 0·0780 | 0·0930 | 0·118 | 0·157 |
| 9 | 0·0276 | 0·0340 | 0·0440 | 0·0612 | 0·0840 | 0·0922 | 0·125 | 0·165 |
| 10 | 0·0308 | 0·0372 | 0·0485 | 0·0660 | 0·0885 | 0·104 | 0·130 | 0·171 |
| 11 | 0·0335 | 0·0409 | 0·0521 | 0·0700 | 0·0920 | 0·107 | 0·134 | 0·175 |
| 12 | 0·0365 | 0·0435 | 0·0550 | 0·0730 | 0·0945 | 0·110 | 0·136 | 0·178 |
| 13 | 0·0388 | 0·0455 | 0·0570 | 0·0750 | 0·0965 | 0·112 | 0·138 | 0·179 |
| 14 | 0·0405 | 0·0468 | 0·0585 | 0·0765 | 0·0970 | 0·113 | 0·140 | 0·180 |
| 15 | 0·0410 | 0·0476 | 0·0592 | 0·0770 | 0·0975 | 0·113 | 0·142 | |
| 16 | 0·0413 | 0·0478 | 0·0597 | 0·0775 | 0·0976 | 0·113 | 0·142 | 0·183 |
| 17 | 0·0413 | 0·0479 | 0·0598 | 0·0777 | 0·0977 | 0·113 | 0·142 | |
| 18 | 0·0413 | 0·0480 | 0·0599 | 0·0779 | 0·0978 | 0·113 | 0·142 | 0·184 |
| 20 | 0·0413 | 0·0480 | 0·0600 | 0·0780 | 0·0980 | 0·113 | 0·142 | 0·184 |
| 60 | 0·0413 | 0·0481 | 0·0600 | 0·0780 | 0·0980 | 0·113 | 0·142 | 0·184 |

| $\gamma^2$ | $10^3$ | $7.10^2$ | $4.10^2$ | $2.10^2$ | $10^2$ | 70 | 40 | 20 |
|---|---|---|---|---|---|---|---|---|
| $10^{-2}$ | | | | | 0·00102 | 0.00120 | 0·00158 | 0·00210 |
| 2 | | | 0·00100 | 0·00140 | 0·00205 | 0·00240 | 0·00315 | 0·00420 |
| 4 | 0·00126 | 0·00152 | 0·00200 | 0·00280 | 0·00410 | 0·00480 | 0·00632 | 0·00840 |
| 6 | 0·00189 | | 0·00300 | 0·00420 | | 0·00720 | 0·00948 | 0·0127 |
| 8 | | | 0·00400 | | | | | |
| $10^{-1}$ | 0·00315 | 0·00380 | 0·00500 | 0·00700 | 0·0102 | 0·0120 | 0·0158 | 0·0210 |
| 2 | 0·00630 | 0·00760 | 0·0100 | 0·0140 | 0·0205 | 0·0240 | 0·0315 | 0·0420 |
| 4 | 0·0126 | 0·152 | 0·0200 | 0·0280 | 0·0400 | 0·0480 | 0·0632 | 0·0840 |
| 6 | 0·0189 | | 0·0295 | 0·0420 | 0·0600 | 0·0720 | 0·0948 | 0·126 |
| 8 | | | | | | | | |
| $10^0$ | 0·0315 | 0·0380 | 0·490 | 0·0700 | 0·100 | 0·120 | 0·158 | 0·210 |
| 1·5 | | | | | | | | |
| 2 | 0·0630 | 0·0760 | 0·0980 | 0·139 | 0·190 | 0·230 | 0·300 | 0·400 |
| 2·5 | | | | | | | | |
| 3 | | | | | | | 0·430 | 0·565 |
| 3·5 | | | | | | | | |
| 4 | 0·125 | 0·151 | 0·194 | 0·270 | 0·363 | 0·425 | 0·539 | 0·675 |
| 4·5 | | | | | | | | |
| 5 | 0·155 | 0·185 | 0·235 | 0·320 | 0·428 | 0·490 | 0·610 | 0·742 |
| 5·5 | | | | | | | | |
| 6 | 0·180 | 0·213 | 0·271 | 0·355 | 0·469 | 0·528 | 0·650 | 0·794 |
| 7 | 0·200 | 0·234 | 0·298 | 0·379 | 0·490 | 0·550 | 0·678 | 0·820 |
| 8 | 0·214 | 0·250 | 0·312 | 0·390 | 0·508 | 0·565 | 0.690 | 0·835 |
| 9 | 0·223 | 0·259 | 0·322 | 0·400 | 0·518 | 0·575 | 0·696 | 0·845 |
| 10 | 0·230 | 0·262 | 0·330 | 0·406 | 0·520 | 0·580 | 0·700 | 0·850 |
| 11 | 0·234 | 0·267 | | | | | | |
| 12 | 0·237 | 0·268 | 0·331 | 0·410 | 0·522 | 0·582 | | |
| 13 | 0·238 | 0·269 | | | | | | |
| 14 | 0·239 | 0·270 | 0·332 | 0·413 | 0·524 | 0·584 | | |
| 15 | | | | | | | | |
| 16 | 0·239 | | | 0·416 | 0·526 | 0·585 | 0·700 | 0·855 |
| 17 | | | | | | | | |
| 18 | | | | | | | | |
| 19 | | | | | | | | |
| 20 | | | | 0·420 | 0·530 | | 0·700 | 0·860 |
| 60 | 0·239 | 0·270 | 0·332 | 0·420 | 0·530 | 0·585 | 0·700 | 0·860 |

| $\eta$ \ $\gamma^2$ | 10 | 3 | 1· | 0·5 | 0·3 | 0·2 | 0·1 | 0·05 | 0 |
|---|---|---|---|---|---|---|---|---|---|
| $10^{-2}$ | 0·00300 | 0·00570 | 0·0100 | 0·0140 | | 0·0220 | 0·0300 | 0·0435 | 0·196 |
| 2 | 0·00600 | 0·0115 | 0·0196 | 0·0280 | | 0·0440 | 0·0600 | 0·0820 | 0·275 |
| 4 | 0·0120 | 0·0230 | 0·0388 | 0·0560 | | 0·0850 | 0·112 | 0·150 | 0·385 |
| 6 | 0·0180 | 0·0345 | 0·0580 | 0·0840 | | 0·125 | 0·164 | 0·210 | 0·470 |
| 8 | | | | | | | | | |
| $10^{-1}$ | 0·0300 | 0·0570 | 0·0970 | 0·135 | | 0·198 | 0·252 | 0·310 | 0·485 |
| 2 | 0·0600 | 0·113 | 0·190 | 0·226 | | 0·358 | 0·430 | 0·525 | 0·825 |
| 4 | 0·120 | 0·220 | 0·350 | 0·450 | | 0·595 | 0·700 | 0·790 | 1·12 |
| 6 | 0·180 | 0·320 | 0·495 | 0·625 | | 0·770 | 0·900 | 0·995 | 1·33 |
| 8 | | | | | | | | | |
| $10^{0}$ | 0·300 | 0·500 | 0·720 | 0·875 | 0·970 | 1·05 | 1·16 | 1·27 | 1·58 |
| 1·5 | | | | | | | | | |
| 2 | 0·545 | 0·830 | 1·12 | 1·31 | 1·40 | 1·48 | 1·62 | 1·71 | 2·00 |
| 2·5 | | | | | | | | | |
| 3 | 0·720 | 1·04 | 1·33 | 1·52 | 1·62 | 1·71 | 1·82 | 1·94 | 2·23 |
| 3·5 | | | | | | | | | |
| 4 | 0·840 | 1·16 | 1·45 | 1·64 | 1·74 | 1·84 | 1·95 | 2·05 | 2·36 |
| 4·5 | | | | | | | | | |
| 5 | 0·910 | 1·24 | 1·53 | 1·70 | 1·81 | 1·90 | 2·01 | 2·12 | 2·42 |
| 5·5 | | | | | | | | | |
| 6 | 0·960 | 1·28 | 1·57 | 1·76 | 1·85 | 1·95 | 2·07 | 2·18 | 2·48 |
| 7 | 0·980 | 1·31 | 1·60 | 1·79 | 1·88 | 1·97 | 2·10 | 2·20 | 2·50 |
| 8 | 1·00 | 1·32 | 1·62 | 1·80 | 1·90 | 1·99 | 2·11 | 2·21 | 2·51 |
| 9 | 1·01 | 1·33 | 1·63 | 1·81 | 1·905 | 1·995 | 2·12 | 2·22 | 2·52 |
| 10 | 1·02 | 1·34 | 1·64 | 1·82 | 1·91 | 2·00 | 2·13 | 2·23 | 2·53 |
| 11 | | | | | | | | | |
| 12 | | | | | | | | | |
| 13 | | | | | | | | | |
| 14 | | | | | | | | | |
| 15 | | | | | | | | | |
| 16 | 1·03 | | 1·65 | 1·82 | 1·92 | 2·00 | 2·14 | 2·24 | 2·55 |
| 17 | | | | | | | | | |
| 18 | | | | | | | | | |
| 20 | 1·04 | 1·35 | | | | | | | |
| 60 | 1·04 | 1·35 | 1·65 | 1·82 | 1·92 | 2·01 | 2·14 | 2·24 | 2·55 |

# PRACTICAL METHODS FOR MEASURING AND CALCULATING THE LUMINANCE OF ROAD SURFACES

J. B. de BOER, V. OÑATE*) and A. OOSTRIJCK

535.241.44:628.971

**Summary**

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 description of these methods is preceded by a discussion of some publications on this subject. For the computation of luminance values according to the method given in this paper, data regarding the reflection properties of road surfaces are needed. Such data as are given in existing literature are incomplete or of insufficient accuracy. Therefore, 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.

**Résumé**

Dans la définition des qualités d'une installation d'éclairage de routes, la distribution de la luminance de la surface routière constitue un facteur bien plus important que l'éclairement. Toutefois, l'absence de méthodes précises et rapides pour la mesure et la détermination de la luminance explique pourquoi, jusqu'à présent, on s'en est tenu essentiellement à l'emploi des valeurs de l'éclairement. Cet article expose des méthodes qui permettront à l'éclairagiste d'utiliser en pratique la luminance de la surface routière. L'exposé de ces méthodes est précédé d'une discussion de certaines publications traitant de ce sujet. Le calcul des valeurs de la luminance suivant la méthode exposée dans cet article, nécessite des données relatives aux propriétés de réflexion des surfaces routières. Certaines données figurant dans la littérature consacrée à cet objet sont incomplètes ou n'ont pas la précision requise. C'est la raison pour laquelle on a mesuré les propriétés de réflexion d'une surface routière déterminée (asphalte sec). Enfin, un appendice donne à titre d'exemple, un calcul de la luminance.

**Zusammenfassung**

Bei der Bestimmung der Eigenschaften einer Straßenbeleuchtungsanlage stellt die Leuchtdichteverteilung auf der Straßendecke einen wichtigeren Faktor dar als die Beleuchtung. Trotzdem sind in Ermanglung genauer und schneller Methoden zur Messung und Berechnung der Leuchtdichte bisher hauptsächlich Beleuchtungswerte benutzt worden. In diesem Artikel werden Methoden beschrieben, die den Straßenbeleuchtungsingenieur die Möglichkeit geben, die Leuchtdichtewerte auf der Straße praktisch anzuwenden. Der Beschreibung dieser Methoden geht eine Besprechung einiger Veröffentlichungen über diesen Gegenstand voran. Zur Berechnung von Leuchtdichtewerten nach dem in diesem Artikel gegebenen Verfahren

*) At present with Instituto Tecnico de la Construccion y del Cemento, Madrid.

sind Daten hinsichtlich der Reflexionseigenschaften von Straßendecken erforderlich. Weil die in der Literatur erwähnten Daten unvollständig oder nicht genügend genau sind wurden die Reflexionseigenschaften einer bestimmten Straßendecke (trockene Asphaltstraße) gemessen. Im Anhang wird ein Beispiel für die Leuchtdichteberechnung gegeben.

## 1. Introduction

On a lighted road a car-driver should be able to see objects or bends without using headlights. Objects on the road can be seen if the three following conditions are fulfilled:

(1) the luminance of the object differs sufficiently from the background luminance, the latter being normally that of the road surface;
(2) the object is large enough;
(3) the time available for perception is sufficient.

Visibility is therefore determined by contrast sensitivity, visual acuity and reaction time. All these factors depend strongly upon the luminance of the road surface and its distribution. Furthermore the comfort of seeing depends to a high degree upon the presence of glaring light sources as well as, e.g., in the case of a wet road surface, on the luminance and the dimensions of the bright patches on the road. It is again the general luminance level of the road surface which determines in how far this comfort of seeing is disturbed by glaring light sources or by bright patches on the road. The above-mentioned problems of "visibility" and "comfort of seeing" have been studied in detail in an earlier publication[1]).

The luminance and its distribution, as seen by the observer, does not bear any simple relation to the illumination of the road. All road surfaces behave more or less like a mat mirror. The light radiated by the lamps in a direction opposed to that of the traffic will contribute to the road luminance to a much higher degree than the light radiated in the direction of the traffic. This means that, in order to characterize a road surface, a single reflection factor will not suffice. The reflection properties must be given as a function of two parameters, viz. the direction of illumination and of observation. Thus it is also clear that the distribution of the illumination does not constitute a measure for the quality of a street-lighting installation.

Various investigators have stressed the importance of luminance instead of illumination. Waldram [2]), in a paper which appeared in 1934, was one of the first to emphasize the value of the luminance in connection with road-illumination problems The reflection properties of road surfaces have since been studied by several investigators; see e.g., Cohu [3]), Weigel and Schlüsser [4]), Moon and Hunt [5]), Bergmans [6]), Reid and Chanon [7]). These studies have yielded indispensable material for calculating paving

luminance as well as for determining what light distribution and arrangement of the lamps assures the most uniform distribution of luminance.

Meanwhile methods and instruments, intended for practical use, had been developed for measuring road luminance. For instance there is the photographic method by Hopkinson [8]), Waldram [9]) and Bloch [10]), the brightness meter by Luckiesh and Taylor [11]) and other portable instruments such as that manufactured by G.E.C., England [8]).

Nevertheless everyday practice has not reached a point where luminance and reflection properties of the road play a paramount role. Descriptions of installations rarely state luminance values. Even modern regulations or specifications for road illumination [12]), [13]) only mention illumination data. The instruments for measuring paving luminance are either too complicated and time-consuming or lacking in adequate precision for luminance measurements at places on the road surface seen from directions and distances normal to the road user. In the manner in which data on reflection properties are compiled, either lengthy calculations are involved or else factors are ignored which from the point of view of accuracy cannot be overlooked. Furthermore, these data will often be incomplete, valuable informations as to the directions of incident and reflected light having been omitted.

These are some of the factors which up to now have stood in the way of luminance values being used in road illumination engineering, notwithstanding the generally accepted importance of these values.

The same impediments were experienced in the development work of road-illumination research of the Philips Light-technical Laboratory. This has led to a method being developed for measuring the luminance of a road surface which is suitable for use under practical conditions. It has also been found possible to present reflection properties in such a form as to make the computation of road luminance a simple matter.

In view of the above-mentioned importance of the distribution of luminance on road surfaces, the description of appropriate methods of measuring and calculation suited for practical conditions, should contribute towards further development of road illumination engineering.

## 2. Requirements and existing instruments

A luminance meter for street-lighting purposes serves to collect data on the distribution of the luminance in the normal visual field of a car-driver. Thus it is primarily the perspective image of the road as seen by the car-driver that is of interest. The most important part of this image is the road surface against which most objects will be seen. For this reason it is necessary to know how the luminance of the road surface is distributed.

Let us imagine a box 0·20 m high on a road at a distance of 50 m in front

of a car. The driver will see the box silhouetted against a stretch of road the length of which is approximately 7 m long. At a distance of 100 m the box apparently covers 15 m of the road, at 200 m 30 m. It is the average luminance of the image of this particular stretch of road which is of importance for visibility, provided no sudden changes in luminance occur. A luminance meter should therefore be able to measure average luminance values of fractions of the perspective image of the road surface not greater, and if possible smaller, than the above-mentioned stretches. Only details of the distribution of the luminance in this perspective image are of importance. These details are quite different from those recorded in a plan view of the road. It is of little use to represent the luminance distribution, as is often done, by means of iso-luminance curves plotted in a plan of the road, as this leads to a subdivision that is too fine at great distances in front of the observer, and too coarse at short distances. As a matter of fact it suffices to have a coarser subdivision of the road the farther the places to be measured lie ahead and this makes it easier to construct a luminance meter ready for practical use. In fast traffic the driver must be able to survey the part of the road at 100 to 200 m ahead. Consequently the luminance distribution in this stretch of road, as he sees it, is highly important. Shorter distances enter into the picture when bends or branches of the road are to be considered, or in urban traffic. The shorter the distance however, the less are the difficulties in measuring luminance, so that for a start attention can best be concentrated on measuring at 100 to 200 m.

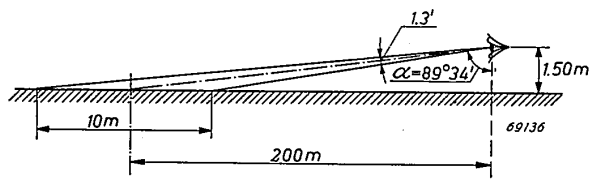



Fig. 1.

Suppose (fig. 1) an observer, whose eye is 1·50 m above road level, is looking 200 m ahead. A stretch of the road, 10 m long at that distance, subtends an angle of 1·3′ at the observer's eye, as is shown in fig. 1. The direction of viewing is given by $\alpha = 89° 34'$. A variation of one minute in $\alpha$ will result in a shift of 8 m in the stretch of road observed. From this it is clear that if it is required to measure the luminance at a "point" of the road surface 200 m ahead the vertical dimension of the field of comparison of the luminance meter has to be very small. Furthermore, the direction of observation will have to be fixed with the utmost accuracy. These demands

may be met by the photographic method mentioned, the photographic density of a picture, taken under well-defined conditions of exposure and development, being measured with a micro-photometer. This, however, is essentially a laboratory method, not fit for general use.

As far as we know, the luminance meters normally available do not meet the demands mentioned. The vertical dimensions of the field of comparison are too large and the instruments cannot be adjusted accurately with respect to the direction of viewing.

To elucidate the difficultics a perspective projection of a road is given in fig. 2, the width of which is supposed to be 10 m. If this figure is viewed at a distance of 0·60 m the impression on the eye will be that as seen by the driver of a car travelling along the centre of the right-hand traffic

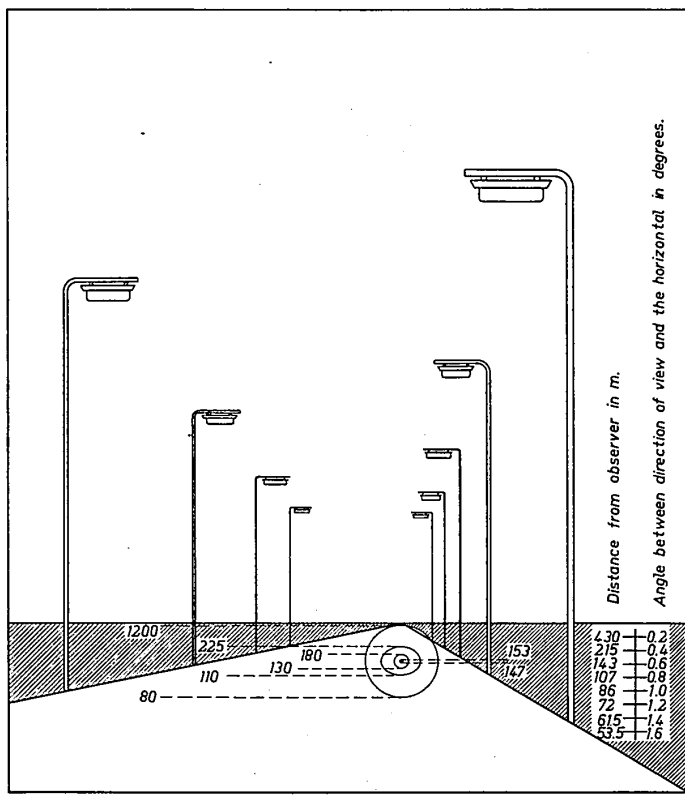



Fig. 2. The part of the road encompassed by a luminance measurement at a point 150 m in front of the observer when using various instruments.

lane. The drawing shows, what part of the road is covered by a luminance measurement at a point 150 m in front of the observer when using various instruments. Most normal photometers have a circular field of comparison subtending an angle of more than 1°. The larger circle in fig. 2 shows the part of the road covered by a circular field of 1°, which proves to be a stretch between 80 and 1200 m in front of the observer. Secondly, fig. 2 shows the oval-shaped field of comparison of a luminance meter intended for taking measurements in interiors and for photographic purposes. This instrument is manufactured by Salford Electrical Instruments Ltd. England. Although the field of comparison of this instrument is strikingly small, the part of the road covered still ranges from 110 to 225 m. The luminance meter for road-lighting purposes as manufactured by the General Electric Co. has a circular comparison field which, if the instrument is held at 0·60 m in front of the eye, covers a stretch of the road from 130 m to 180 m. Even this is too large for obtaining an accurate impression of the luminance distribution in the lengthwise direction of the road. Finally fig. 2 also shows the part of the road, ranging from 147 m to 153 m, which is covered by the field of comparison of the luminance meters developed by the authors.

The accuracy necessary for adjustment of the direction of viewing is shown in the right-hand scale of fig. 2, where the angle is given between the horizontal plane and the line of view for various distances in front of the observer. For a point at 150 m the angle is 0·6°, and here a difference of 0·2° upwards corresponds to a shift of the observed point of about 70 m.

### 3. A new method for measuring the luminance of road surfaces

First a luminance meter was developed in which the demands of a small field of comparison and accurate adjustment of the line of view were fulfilled by separating the field of comparison from the rest of the instrument as indicated in fig. 3. The line of view between the eye $O$ and the point $P$ is defined by two holes in screens 10 m apart. The size of the opening near the eye is approximately that of the largest possible eye aperture. The side of the screen $S$ facing the observer is illuminated by the lamp $L$. The

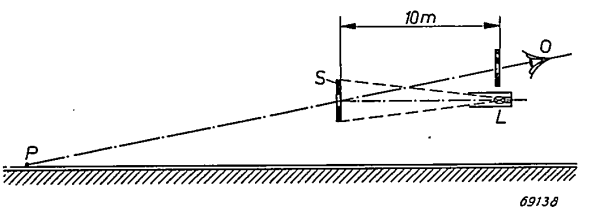



Fig. 3. Principle of a luminance meter where the field of comparison is separated from the rest of the instrument.

illumination can be adjusted by the observer to match the brightness of the part of the road seen through the opening in $S$ with that of the surrounding part of the screen. The latter has been calibrated. This method would be very time-consuming if the screen $S$ had to be readjusted for every new measuring point to be measured. Now, as already stated, the aim in measuring the luminance of the road surface is to obtain an impression of the luminance distribution in the perspective image of the road. This will be achieved if the measurements are carried out for a sufficient number of points of this perspective image. To this end a large number of holes are made in the screen $S$ distributed in such a way that, as indicated in fig. 4,

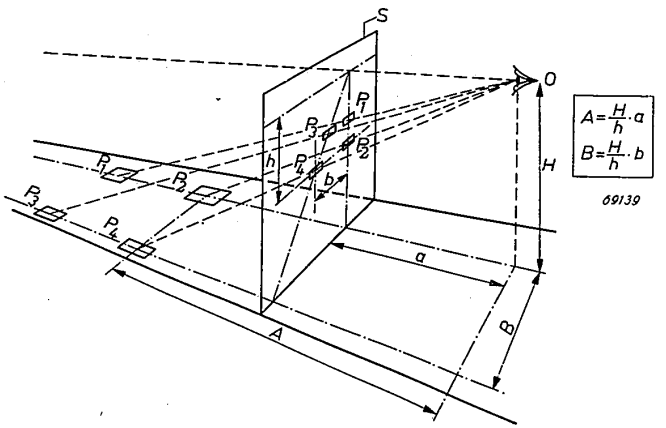



Fig. 4. In a luminance meter according to fig. 3 a large number of holes, corresponding to regularly distributed points on the road surface can be made in the measuring screen. Then the luminance meter will not have to be readjusted for each measuring point.

the observer sees the points to be measured in a regular pattern on the road surface. The place of these holes in the screen $S$ is determined by the perspective projection of this pattern on the screen $S$ with the observer's eye as the centre of projection. Provided the screen $S$ is perpendicular to the axis of the road, it is now only necessary to adjust the apparatus correctly for two holes, the whole pattern of points thereby being fixed with respect to the observer. Thus a complete picture of the luminance distribution of the road in front of the observer will be obtained.

The rectangular holes in the screen are 4 mm high and 20 mm wide, corresponding to a sufficiently small vertical angular dimension of the observed part of the road. The screen is observed through a telescope with an angular magnification of about 12. In this way a sufficient contrast sensitivity

can be obtained notwithstanding the small angular dimension of the field of comparison. Fig. 5 shows a picture of the screen S. *)

As, owing to its large dimensions, the apparatus just described does not lend itself well for general use, a solution was sought in the development of an apparatus of smaller dimensions satisfying the same requirements as to the height of the field of comparison and the directional adjustment and offering a wider field of application in road-illumination practice.




Fig. 5. Actual set-up of the screen of a luminance meter based on the principle of fig. 3.

Since reduction in size involves a smaller field of comparison, a technique was developed for making small mirrors on a glass surface **). Fig. 6 shows the principle of the apparatus. A convex mirror $H$ with a focal length of 0·50 m and a plane mirror $F$ produce an image of the road in a vertical plane through the small mirror $M$ of 0·3 mm diameter made by evaporating aluminium upon a sheet of transparent glass. The apparatus is set upon the road at such a level that the convex mirror $H$ is at the normal height of the eyes of a driver. The observer looks through the Huygens ocular at the image of the road. The angular magnification of the ocular is 6 times. In the small mirorr $M$ a frosted glass FG is seen, which covers a ring-shaped box with a small incandescent lamp (2·5 V - 0·2 A) inside. As the luminance of the frosted glass decreases with increasing distance from the lamp, the comparison luminance seen in the small mirror $M$ can be varied by turning the ring-shaped box. The angle between the plane of the mirror $M$ and the optical axis is about 45°, so that the vertical dimension of the field of comparison is about 0·2 mm. This covers a stretch of road of about 10 m at 200 m ahead of the observer, which is small enough for accurate measurements. At smaller distances the stretch observed becomes proportionally shorter.

*) This luminance meter has been described in detail elsewhere [15]). It has been used i.a. for characterizing the paving of the test road for the international visibility tests with automobile headlights carried out by the I.C.I. at Zandvoort, Holland, in October 1949.

**)The authors acknowledge the valuable assistance of Dr A. M. Kruithof of Philips Research Laboratory in making these mirrors.

The direction of viewing is accurately fixed in the following way. The plane mirror $F$ in the cardan joints $C$ can be turned by means of adjusting screws in two directions at right angles to each other. Thus the image of any desired point on the road can be focused on the small mirror $M$. A transparant glass $CG$, approximately in the focal plane of the convex mirror $H$, is rigidly connected with the plane mirror $F$ by means of the rod $R$. The positions of $F$ required for fixing the points to be measured can thus be marked by the places on $CG$ covered by $M$ in these positions. Thus the pattern of measuring points on the road corresponds to a pattern of indicated points on $CG$. When the mirror $F$ is turned over a certain angle the optical image of the road shifts over twice that angle, which means that the pattern on $CG$ has about half the dimensions of the image of the road.

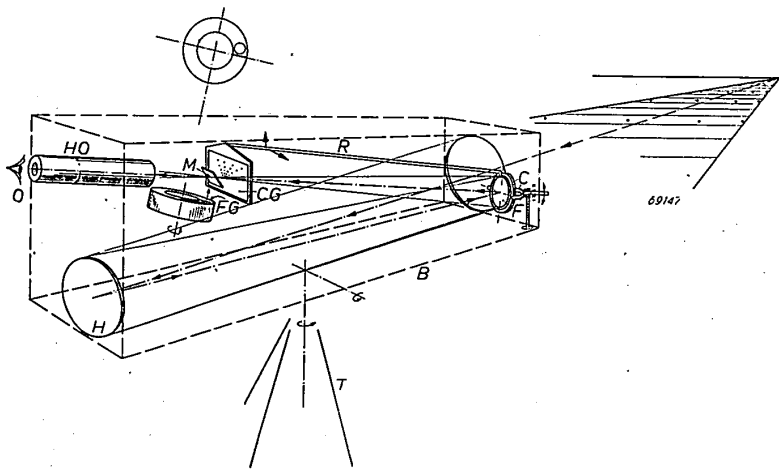



Fig. 6. Diagrammatic representation of a luminance meter with an extremely small mirror as field of comparison; $M$ = small comparison mirror, $H$ = convex mirror with focal plane in $M$ (via $F$), $F$ = plane mirror, $CG$ = transparent glass with adjusting pattern, $R$ = rod connecting $CG$ rigidly to $F$, $C$ = cardan joints allowing rotation of $CG$ and $F$ around the centre of $F$, $HO$ = Huygens ocular, $O$ = observer's eye, $FG$ = frosted glass with decreasing luminance, $B$ = case, $T$ = tripod.

An advantage of this method is that in order to make an adjustment all that is needed is to move a mirror. Furthermore, the range of the angle of measuring is twice that of the adjustment, so that measurements can be made at points relatively close to the observer without the instrument having to be made too wide.

A further advantage as compared with the luminance meter, the screen of which is photographed in fig. 5, is the possibility to measure the luminance

of points on the road not belonging to the pattern of measuring points indicated in the meter. In the screen according to fig. 5 this was impossible because the points to be measured were seen through fixed holes.

It is clear that the instrument has to be so adjusted that the pattern on *CG* corresponds to the pattern chosen for the regularly distributed points on the road. This will be realized if one point on *CG* and a line through it cover a corresponding point and line on the road.

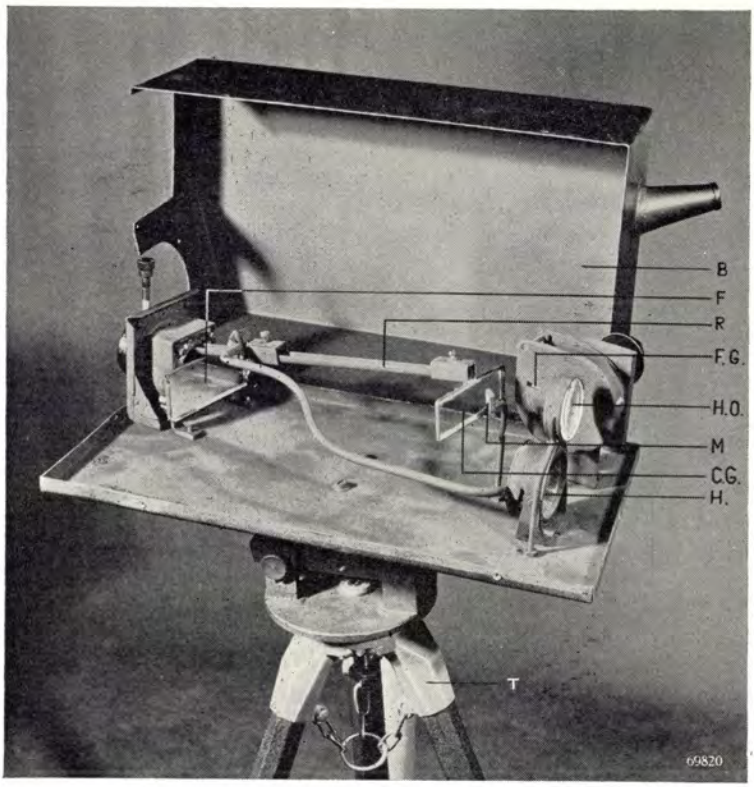



Fig. 7. Photograph of the luminance meter drawn in fig. 6. For the meaning of the letters, see fig. 6.

It lies to hand to choose the vanishing point of the road and the horizon for this purpose. In practice this principle is applied by placing two lamps on the road at a distance of, say, 50 m in front of the observer at the level

of $H$ (1·50 m), one at the same distance from the right-hand border of the road as the observer and the other at the left-hand border.

The mirror $F$ and the box $B$ must be adjusted relative to the tripod $T$ in such a way that the first lamp (as seen through the ocular) is covered by the mirror $M$ on that spot of the screen $CG$ which corresponds to the vanishing point, whilst the image of the other lamp is seen on the line on $CG$ representing the horizon. The plane of $CG$ should be approximately perpendicular to the road; small deviations are of no consequence.

The points selected lie at 30, 40, 50, 60, 70, 90, 120, 150, 180 and 210 m, 6 of them on every line at distances of 1, 2·5, 4, 5, 7 and 9 m from the right-hand side of a 10 m wide road.

The luminance of $CG$, seen via $M$, varies from 2·5 to 35 cd/m$^2$. The voltage of the lamp (2·5 V - 0·2 A) in the ringshaped box is regulated by a potentiometer and read on a voltmeter. The lamp is fed from a dry cell (torch battery) within the apparatus. Neutral grey filters serve to attenuate either the luminance of the observed point or that of the field of comparison by a factor 10, 100 or 1000, thus extending the range of measurement. Hence luminances from 0·0025 cd/m$^2$ to 35000 cd/m$^2$ can be measured. Since an average luminance of 0·3 cd/m$^2$ is accepted as a normal value, a reasonable accuracy in the region of 0·01 cd/m$^2$ will be necessary. With the help of the data of Richard Blackwell [14]) the accuracy of the instrument described here was computed to be about 15 % at 0·01 cd/m$^2$, which is satisfactory.

The complete measurement of the distribution of the luminance in the above-mentioned pattern of points takes about one hour, even without much experience.

The overall size of the apparatus is 0·1 × 0·15 × 0·4 m. Fig. 7 shows a photograph of the instrument.

## 4. A method of calculating the luminance of road surfaces

The reflection properties of road surfaces as given in the literature cannot be used in a simple way or with sufficient accuracy for calculations of the expected luminance distribution resulting from a given lighting installation.

Some investigators, a.o. Bergmans [6]) and Reid and Chanon [7]) gave their results in graphs for the sake of easier reference. Bergmans considers angles of incidence up to 80° only, which for fixtures at 8 meters above road level, corresponds to a horizontal distance from the lamp of approximately 45 m. These data can therefore only be used with fixtures of which the light distribution curve is cut off at an angle of 75°-80° to the vertical. Modern fixtures, e.g. with fluorescent tubes, even radiate above the horizontal plane, and reflectivity data are too incomplete for such installations. Reid and Chanon [7]) have made measurements up to 89° (corresponding to a

horizontal distance of 460 m at a height of the fixture of 8 m). In the region of large angles of incidence, however, these data require rather rough interpolations. Further, particularly the work done by Cohu [3]) shows that many measurements have been made with combinations of illumination- and viewing angles seldom occurring in practice.

In the following pages will be discussed:

(a) the methods for representing the reflection properties of a road surface;
(b) the form in which the results of some earlier work have been given, and the disadvantages thereof;
(c) a method avoiding these drawbacks and its application in computing the luminance distribution for a given installation, planned or existing.

The appendix gives an example of the calculation for a dry asphalt road surface, the reflection coefficients of which were measured as no adequate data could be found in the literature.

(a) *The representation of the reflection properties of a road surface*

For fixed directions of illumination and of observation the luminance of the surface will be proportional to the illumination. With a *horizontal* illumination of 1 lux due to light coming from one direction, let the luminanceas seen in another direction be $q$ cd/m$^2$. In accordance with Cohu[3]) and Bergmans [6]), this factor $q$ will be called the "luminance coefficient" for the corresponding directions of incidence and reflection. According to fig. 8 these directions are defined by the angle of incidence of the light $\gamma$, the viewing angle $\alpha$, and the angle $\beta$ between the horizontal projections of $OP$ and $LP$. If the road surface shows a special texture, as, with brick roads and roads where the traffic has worn the surface, the direction of this texture relative to the geometric diagram of fig. 8 would influence the value of $q$. According to Bergmans this influence is of little consequence, and it will therefore be ignored.

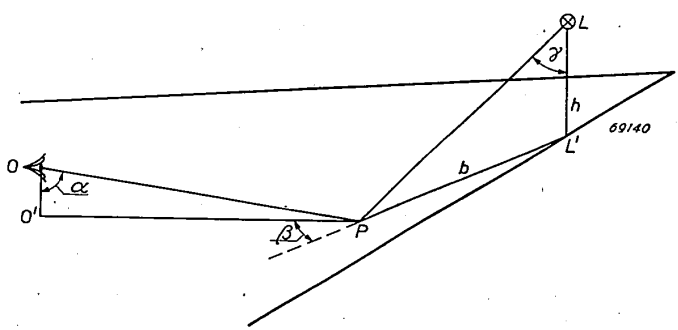



Fig. 8. Geometrical figure with angles $\alpha$, $\beta$ and $\gamma$ determining the luminance coefficient $q$; $L$ = light source, $O$ = observer's eye, $P$ = point under observation.

Theoretically $q = f(\alpha, \beta, \gamma)$ could be given in tabular form for practical ranges of $\alpha$, $\beta$ and $\gamma$ and a given paving. As this involves laborious calculations, a graphical way of presenting the results seems preferable, and has the advantage of being illustrative. Since a two-dimensional graph can only give the relation between three quantities, at least a series of graphs each valid for some value of the fourth quantity, will be required.

In fast traffic the driver's attention will mostly be concentrated on the part of the road more than 50 m, as a rule more than 100 m, ahead. The range of $\alpha$ covering this part of the road is between about 88° and 89·5°. Within this range, and for constant values of $\beta$ and $\gamma$, $q$ is approximately independent of $\alpha$, a fact observed by various investigators, among whom Moon and Hunt [5]) and Bergmans [6]), and confirmed by our own measurements. It is therefore sufficient to give a graph of $q = f(\beta, \gamma)$ at $\alpha = 89°$. For a fixed height of the lamp above the road surface this function can be written as $q = f(\beta, b)$, the angle $\gamma$ having been replaced by the distance $b$ (see fig. 8). The advantage of this new variable $b$ is that it lends itself better for representation in a horizontal plane.

(b) *Earlier research*

Cohu's method of presenting the reflection properties of the road involves too many calculations for practical use. Though Mainie [16]) has given nomographs which considerably simplify these calculations, this method is still laborious. Bergmans' method will be considered with reference to fig. 9, showing a plan of the road. Let the observer's eye be at a height of 1·50 m above the point $O'$, and the lamp at, say, 8 m above $L'$. The distance between the observer and the lamp is, e.g., 150 m. In the diagram lines have been drawn each representing the locus of a number of points on the road (seen by the observer) with a luminance coefficient as indicated at each line (iso-$q$ lines).

Obviously all lines must be symmetrical with respect to the line $O'L'$. Actually this graph is only valid if the distance between observer and the light source has a fixed value, e.g., 150 m. It has been seen that variations of $\alpha$ within the range 88°-89·5° do not affect the value of $q$ if the two other parameters are constant. If the observer moves along the line $O'L'$, whilst

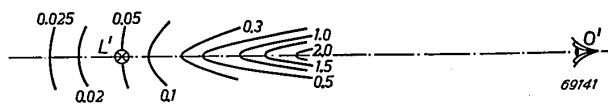



Fig. 9. Representation of the reflection properties of a road surface according to Bergmans [6]). The plane of drawing represents the road surface. The observer is above point $O'$, the lantern above point $L'$. Points on the surface with equal luminance coefficients are connected by curves. Bergmans applies the diagram for distances between $L'$ and $O' > 60$ m.

observing one fixed point on the road, $\beta$ will generally vary, unless the observed point on the road lies on the line $O'L'$ ($\beta = 0$). For all other points a movement of the observer along $O'L'$ involves a slight change of $\beta$. Bergmans and Vervest [6]) assume these variations of $\beta$ to be negligible with respect to the influence upon $q$ and use the same iso-$q$ graph for distances of viewing of 60 m and more. For their purpose, viz. the designing of road-lighting fixtures, for which they introduced the conception of an "average road surface", this is permissible. When, however, it is our aim to carry out luminance calculations for a particular lighting installation and for a particular road, it is not always permissible to disregard these variations of $\beta$, as, especially for wet surfaces, $q$ varies strongly with small variations of $\beta$. This error may be as high as 30 % for dry surfaces and 100% for wet surfaces.

Similar objections can be raised against Reid and Chanon's method, which will be discussed with reference to fig. 10. Again a plane of the road is drawn, with the lamp over point $L'$. The observer moves along the dotted line with his eyes kept on points on the road at a fixed distance (200 ft). These points are situated on a circle with a radius of 200 ft, with the observer in the centre. The lines in the diagram represent the loci of the observed points having the same luminance coefficient for an observer moving as indicated.

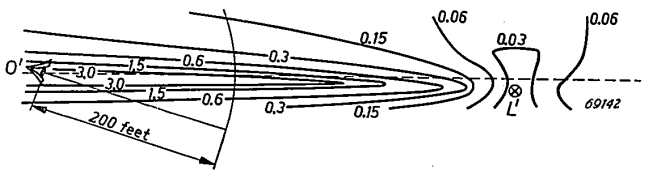



Fig. 10. Representation of the reflection properties of a road surface according to Reid and Chanon [7]). The plane of drawing represents the road surface. The lamp is over point $L'$. The observer can be at any point above the dotted line, looking at points on the road surface 200 feet ahead. The curves of the figure connect points on the road surface with equal luminance coefficients for the observer in question. Reid and Chanon extend the use of the diagram to viewing distances between 200 and 400 feet.

Reid and Chanon assume the same diagram to be valid for viewing distances from 200 to 400 ft, ignoring variations of $a$ which is permissible, as explained above. But, as with Bergmans's method, it is not permissible to ignore variations of $\beta$, the more so since Reid and Chanon are considering wider roads.

(c) *A different method of representing the reflection properties*

The method of representing reflection properties used by us will be discussed with reference to fig. 11 which again shows a plan of the road.

The observers' eye is at 1·50 m perpendicularly above $O'$. The observed point $P$ is, say, 150 m ahead of the observer.

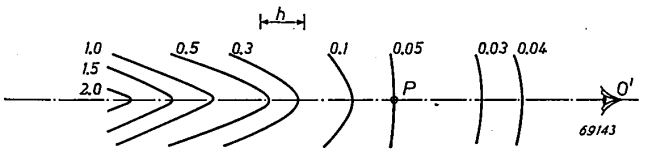



Fig. 11. Diagram of the reflection properties of a road surface in which the curves represent projections of lamp positions for which point $P$ has the same luminance coefficient for the observer above $O'$. The plane of drawing represents the road surface. The diagram can be used for distances between $O'$ and $P$ greater than 40× the eye level of the observer.

Suppose the position and the candle power of a lamp with fixed height is varied, always taking care that a certain constant horizontal illumination at $P$ results. In a diagram like fig. 11 the positions of the light source giving rise to prescribed luminance at $P$ in the direction $PO$ can be interconnected by a curve, provided that at each point of the curve the candle power of the fixture is adjusted so as to ensure a constant horizontal illumination at $P$. In this way a curve is obtained which represents the locus of all those lamp positions giving one and the same value of the luminance coefficient $q$ at $P$ in the direction $PO$. By reverting to the quantity $q$ the requirement that the illumination at $P$ must necessarily be kept constant can be dropped, since, as fully explained sub (a) the luminance coefficient $q$ is independent of variations in illumination. Fig. 11, which thus acquires a more universal significance, shows a combination of the loci of the fixture positions for various values of $q$.

It is of course desirable to remove the restriction still imposed on the use of fig. 11: the height $h$ of the fixture. Up to now nothing has been said about the dimensions of the plane of the road. All difficulties are avoided by drawing the diagram on an arbitrary scale with the fixture height $h$ as a unit of length, indicated on the diagram by a short line. Corresponding to each point of the curves, i.e. at each lamp position, there is then a definite value of $\gamma$ and a definite value of $\beta$. The effect of the variation in $\alpha$ on the value of $q$ is negligible for distances $OP > 60$ m.

From these remarks it may be concluded that the diagram of fig. 11 is of universal use for all distances $OP > 60$ m. It is obvious that the curves in fig. 11 are symmetrical with respect to the line $O'P$ if the road surface does not show any directional effect in texture. As is the case with the diagrams of Bergmans and of Reid and Chanon, however, this diagram does not give a quick impression of the contributions of the various fixtures of an installation to the luminance of $P$ in the observer's direction. From

fig. 11, however, another diagram can be derived which renders this possible. For lamp positions far beyond the observed point the $q$ value will be high. The horizontal illumination due to these lamps at the observed point $P$, however, is very low. That is why the contribution of remote lamps to the total luminance of $P$ in the direction $PO$ will be low inspite of their high $q$ value.

The horizontal illumination caused by a light source at a height $h$ is $(I/h^2)\cos^3\gamma$, where $I$ = the candle power of the lightsource in the direction of the observed point. The observed luminance at $P$ due to this fixture is $q(I/h^2)\cos^3\gamma$. The value of $q$ can be read in a diagram like that of fig. 11. It has already been stated that each point of fig. 11 represents one definite value of $\gamma$. This suggests drawing a new diagram of the loci of those lamp positions which give a constant value of $q\cos^3\gamma$ at the observed point $P$ in the direction $PO$ (fig. 12). Fig. 12 also requires an indication of the dimension of $h$. For a fixture where $I$ is independent of the direction, the diagram of fig. 12 immediately shows the luminance contribution by lamps at various places. For arbitrary fixtures $I$ will nearly always be known as a function of the direction. As the luminous intensity in the direction of different parts of the road does not by any means change so much as the illumination does, one is generally much better informed about the luminous intensity in the direction of the different parts of the road than about the illumination in these parts, and consequently a diagram according to fig. 12 — contrary to the diagram of fig. 11 — gives an impression of the contribution to the luminance obtained from each lamp. A combination of the light distribution pattern of a fixture, preferably in the form of an isocandela diagram plotted in a plan of the road, with the diagram of fig. 12 will afford an easy and quick means of computing luminance values.

The diagram of fig. 12 will be referred to as the "E.P. diagram", "E.P." standing for "equivalent positions", i.e. lamp positions that are equivalent

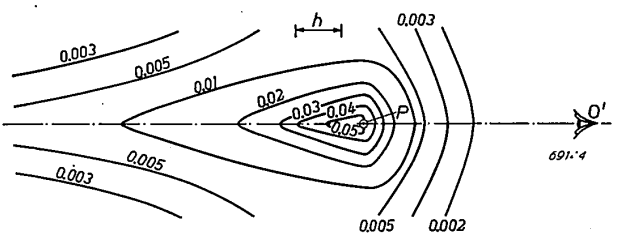



Fig. 12. Representation of the reflection properties of a road surface by means of an E.P. diagram. The observer is over point $O'$ and looking at point $P$ on the road surface. The curves represent projections of lamp positions for which the product $q\cos^3\gamma$ for point $P$ as seen by the observer has the value indicated at each curve. The diagram can be used for distances $O'P$ greater than $40\times$ the eye level of the observer.

with respect to the value of $q \cos^3\gamma$ at $P$. It might be useful at this stage to sum up the analogies and differences between the method described here (fig. 12) and those of Bergmans and of Reid and Chanon. Fig. 13 shows $O$, the observer's eye, the lamp $L$, the point $P$ under consideration, and the projections $O'$ and $L'$ of $O$ and $L$ on the road surface.

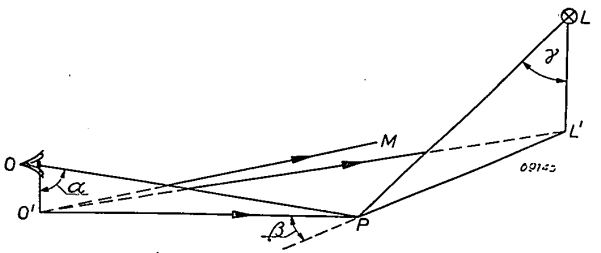



Fig. 13. When calculating the luminance at an arbitrary point of the road surface according to the various methods the observer will be moving:
(*a*) along the line $O'L'$ when using the Bergmans' method of representation (see fig. 9). so that the angles $\alpha$ and $\beta$ are changing;
(*b*) along the line $O'M$ when using the Reid and Chanon method (see fig. 10), the angles $\alpha$ and $\beta$ then likewise changing;
(*c*) along the line $O'P$ when using the E.P. diagram, in which case only the angle $\alpha$ changes.

When Bergman's diagram is used for determining contributions of fixtures to the luminance of a point $P$ of the road at a viewing distance different from that for which the diagram is made, the observer moves along $O'L'$. With Reid and Chanon's diagram the observer will move along a line $O'M$ ($O'M$ runs in the plane $O'PL'$, at a certain distance from $L'$, and corresponds to the dotted line in fig. 10). With the E.P. diagram the observer will follow the line $O'P$. In all these three cases the variations of $\alpha$ proved to be tolerable as regards their effect on the value of $q$. The variation of $\beta$, however, when the observer moves along $O'L'$ or $O'M$ in Bergmans' and in Reid and Chanon's diagrams, has an effect on the luminance value (as already mentioned sub b), which may certainly not be ignored. Only in the E.P. diagram are variations of $\beta$ avoided.

Fig. 14 illustrates how luminance values are calculated for selected points of a paving more than 60 m away from the observer with the aid of an E.P. diagram. A plan of the road with a lighting installation (lamps $L_1$ to $L_5$) is drawn on a sheet of transparent paper, expressing the dimensions of the road in the same unit of length $h$ as indicated in the E.P. diagram. Let $P_1$ in the plan be one of the points on the road to be observed by the observer $O$. Cover the E.P. diagram with the transparent road plan in such a way that $P_1$ coincides with $P$ in the E.P. diagram and the direction $P_1O'$ with the axis of symmetry of the E.P. diagram. Fig. 14*a* shows the situation.

For the lamp positions $L_1$ to $L_5$ the values of $q \cos^3 \gamma$ are then read and compiled in table I (second column). Then the candle power in the direction of $P_1$ has to be determined for each lamp. The simplest way to do this is to apply an iso-candela diagram plotted in a plan of the road and showing for one lamp, by means of iso-candela lines, all points on the road surface towards which the lamp radiates the same luminous intensity.

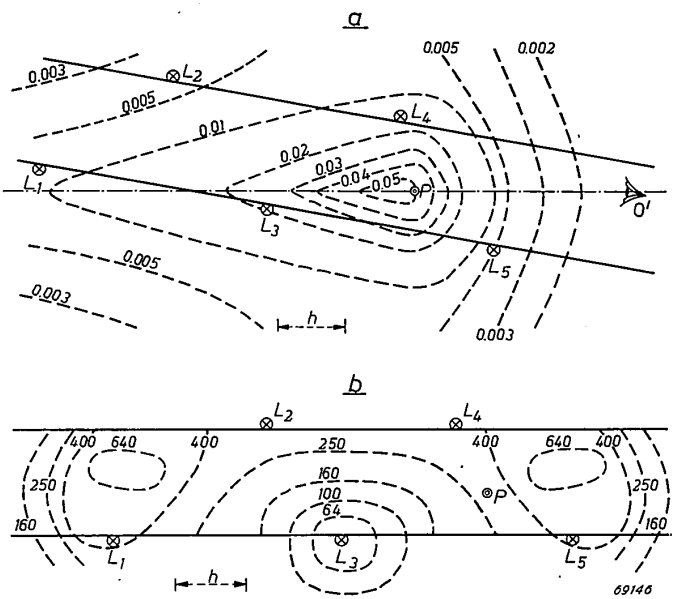



Fig. 14. Explaining the use of the E.P. diagram.

TABLE I

| Lamp | $q\cos^3\gamma$ in $P$ | $I \rightarrow P$ | $h^2 . \Delta B$ |
|---|---|---|---|
| $L_1$ | 0·008 | 40 cd | 0·32 cd |
| $L_2$ | 0·0045 | 640 cd | 2·88 cd |
| $L_3$ | 0·02 | 325 cd | 6·5 cd |
| $L_4$ | 0·015 | 205 cd | 3·08 cd |
| $L_5$ | 0·005 | 190 cd | 0·95 cd |
| | | $\Sigma h^2 . \Delta B = 13{\cdot}73$ cd | |
| | $h = 8$ m | $B = 0{\cdot}214$ cd/m$^2$ | |

Cover this iso-candela diagram with the transparent road plan in such a way that $L_1$ to $L_5$ of the plan coincide successively with $L$ of the iso-candela diagram, making sure that in the case of a non-axial-symmetrical light distribution the iso-candela diagram is in the correct position with respect to the road. In this way the values of the luminous intensity in the direction of $P_1$ can be obtained for each lamp.

For lamp $L_3$ the procedure is shown in fig. 14*b*. The third column of table I gives the various $I$ values. It remains to multiply the numbers in column two by those in column three for each row. The result entered in column four contains the contributions of each lamp to the luminance of the observed point $P_1$ multiplied by $h^2$. Adding the values in column four and dividing the sum by $h^2$ gives the total luminance of the paving at $P_1$ as seen by the observer $O$.

The same operation has to be applied to each point $P_2$, $P_3$, etc. Once this has been done for a sufficiently large number of points iso-limunance lines can be drawn for one particular observer $O$ on a road with a number of fixtures $L_1$, ..., $L_n$.

## Appendix

For a numerical example of this method of calculation a real E.P. diagram is needed. Such a diagram will cover a sufficiently extended area of the road if the value of $q \cos^3\gamma$ at the greatest distances to the point $P$ is smaller than 1% of the maximum value. This maximum value will prevail in the immediate vicinity of the point $P$, since $\cos^3 \gamma$ decreases faster than $q$ increases with increasing distances between the lamp and the point $P$. The only data in the literature sufficiently complete for conversion into an E.P. diagram are those of Reid and Chanon for the two pavings investigated by these scientists: dry concrete and dry asphalt (see figs 15*a* and 16*b*). Owing to the rather coarse division of their diagram, especially for lamps at long distances, this conversion cannot be performed with great accuracy. For one type of paving, i.e. dry asphalt, measurements were therefore carried out on a scale 1 : 4 with the aid of the luminance meter the screen of which is shown in fig. 5.

The result of these measurements is given in the E.P. diagram of fig. 16, which will now be used to calculate the luminance distribution for a road illuminated by fluorescent tubes. Let the road be 10 m wide and paved with dry asphalt. The lighting fixtures are 9 m high, placed 1 m from the side of the road and staggered, the distance between two lamps at one side of the road being 60 m. As fixture was chosen the Philips type UB for four 40-W fluorescent tubes placed horizontally one under another in a vertical plane perpendicular to the longitudinal axis of the road. Fig. 17 shows the iso-candela diagram of this fixture in the horizontal

TABLE II

| $L$ | $P_1$ | | |
|---|---|---|---|
| | $q \cos^3\gamma$ | $I$ in cd | $h^2 . \Delta B$ |
| 1 | 0·0003 | 106 | 0·01 |
| 2 | 0·0003 | 112 | 0·03 |
| 3 | 0·0009 | 107 | 0·10 |
| 4 | 0·027 | 53 | 1·40 |
| 5 | 0·0033 | 107 | 0·35 |
| 6 | 0·016 | 111 | 1·80 |
| 7 | 0·0009 | 106 | 0·09 |
| 8 | 0·0075 | 109 | 0·82 |
| 9 | 0·0004 | 105 | 0·04 |
| 10 | 0·0042 | 108 | 0·46 |
| 11 | 0·0004 | 105 | 0·04 |
| 12 | 0·0028 | 106 | 0·30 |
| 13 | 0·0003 | 104 | 0·03 |
| 14 | 0·002 | 106 | 0·21 |
| 15 | 0·0003 | 104 | 0·03 |

For 1000 lms $\Sigma h^2 . \Delta B = 5{\cdot}71$ cd
For 9000 lumens $B = 0{\cdot}63$ cd/m²

TABLE III

| $P$ | $B$ in cd/m² | $P$ | $B$ in cd/m² |
|---|---|---|---|
| 1 | 0·63 | 14 | 0·64 |
| 2 | 0·59 | 15 | 0·50 |
| 3 | 0·62 | 16 | 0·82 |
| 4 | 0·58 | 17 | 0·63 |
| 5 | 0·39 | 18 | 0·64 |
| 6 | 0·72 | 19 | 0·63 |
| 7 | 0·66 | 20 | 0·45 |
| 8 | 0·69 | 21 | 0·89 |
| 9 | 0·74 | 22 | 0·67 |
| 10 | 0·58 | 23 | 0·67 |
| 11 | 0·72 | 24 | 0·66 |
| 12 | 0·70 | 25 | 0·43 |
| 13 | 0·59 | | |

plane of a road for a total luminous flux of 1000 lumens. The unit of length representing $h$ in this diagram is the same as that in the E.P. diagram of fig. 16.

We draw a plan of the road on transparant paper and of the lighting installation on the basis of the length of $h$ as indicated in the E.P. diagram, taking into consideration that $h$ represents 9 m *). (As the road is 10 m wide, its width in the plan must be 10/9 times the length of $h$ as marked on the E.P. diagram.) In this plan (fig. 18) are marked the position of the observer $O'$, and the points $P_1$ to $P_{25}$ for which the luminance as seen by the observer has to be calculated. The procedure will be described for one point, $P_1$: Cover the E.P. diagram with the transparent road plan, such that $P_1$ of the plan of the road coincides with $P$ of the E.P. diagram and $OP_1$ with the line of symmetry of that diagram. Read the values of $q \cos^3\gamma$ given by each lamp at $P_1$ (second column of table II). Then place the transparent road plan over the iso-candela diagram of fig. 17, covering the lamp position by the various fixture positions on the road in succession, paying special attention to the right directional position of the iso-candela diagram, since in this

*) For practical use the diagrams of figs 15, 16 and 17 are rather small. Diagrams on a larger scale with a plan of the road on transparent paper, may be obtained on request — free of charge — by applying to the writers of this paper.

IPR2025-00217 EXHIBIT 2010 Page 76 of 485

case the fixtures are not axially symmetrical. It is clear that it is advantageous to have the E.P. and iso-candela diagrams with the same length of the unit $h$, as otherwise the road would have to be drawn twice on different scales. Read $I$ in the direction of $P_1$ (and of course at the same time in the direction of $P_2$ to $P_{25}$) for each fixture position (the results for $P_1$ are compiled in column $I$ of table II). Multiply the values of columns two and three of table II for each row (i.e. for each fixture position) and write the result in column four ($h^2.\Delta B$). The total luminance $B_{P_1}$ seen by the observer $O$ at $P_1$ for 1000 lumens of the fixture is the sum of these contributions in column four divided by $h^2$ (i.e. $9^2 = 81$). The resulting quotient has to be multiplied by a factor 9, because the real luminous flux of the fixture is 9000 lumens.

The same procedure has to be applied for each point $P_2$ to $P_{25}$. Table III gives the values of $B$ seen by an observer looking at the points $P_1$ to $P_{25}$ as calculated in this manner.

*Eindhoven, July 1951*
*(Philips Light-technical Laboratory)*

## REFERENCES


1) J. B. de Boer, Fundamental experiments on visibility and admissible glare in road lighting, Paper presented to the I.C.I. congress, Stockholm 1951.
2) J. M. Waldram, Road surface reflection characteristics and their influence on street lighting practice, Proc. Ass. Publ. Lighting Eng. England, Sept. 1934.
3) M. Cohu, Premiers résultats des essais effectués par la Société pour le Perfectionnement de l'Eclairage pour la détermination des charactéristiques réfléchissantes des divers revêtements utilisées dans les voies publiques, Rev. gen. Elect. **37**, 755-767, 1935.
4) R. G. Weigel, P. Schlüsser, Ueber die lichttechnischen Eigenschaften von Strassendecken, Das Licht **8**, 161-166; 237-240, 1935.
5) P. Moon, R. M. Hunt, Reflection characteristics of road surfaces, J. Franklin Inst. **225**, 1-21, 1938.
6) J. Bergmans, Lichtreflectie door wegdekken (in Dutch with an English summary at each chapter), Thesis, Delft, 1938.
J. Bergmans, W. L. Vervest, A new fitting for road lighting, Philips tech. Rev. **5**, 222-230, 1940.
7) K. M. Reid, H. J. Chanon, Evaluation of street lighting, Trans. Ill. Eng. Soc. **34**, 1209-1232, 1939.
8) R. G. Hopkinson, The photographic representation of street lighting installations, G.E.C. Journal **7**, 49-64, 1936.
9) J. M. Waldram, A contouring density comparator, J. sci. Instr. **13**, 352-357, 1936.
10) A. Bloch, The measurement of non-uniform brightness by photographic photometry, J. sci. Instr. **13**, 358-364, 1936.
11) M. Luckiesh, A. H. Taylor, A brightness meter, Trans. Ill. Eng. Soc. **32**, 235-236, 1937.
12) Final report of the ministry of transport departmental committee on street lighting, England, 1937.
13) American standard practice for street and highway lighting, approved Sept. 1947, American Standards Association, Published by Ill. Eng. Soc. (Am.).
14) H. Richard, Blackwell, Contrast thresholds of the human eye, J. opt. Soc. Am. **36**, 624-643, 1946.
15) J. B. de Boer, A. Oostrijck, Een methode voor het bepalen van helderheden en reflectie-eigenschappen van wegdekken (in Dutch only), Electrotechniek **28**, 211-217, 1950.
16) Général H. Mainie, Mode de representation en perspective des brillances d'une chaussé diffusante. Centre de Recherches Scientifiques Industrielles et Maritimes de Marseille, Note Particulière no. 15 du 21 avril 1943.

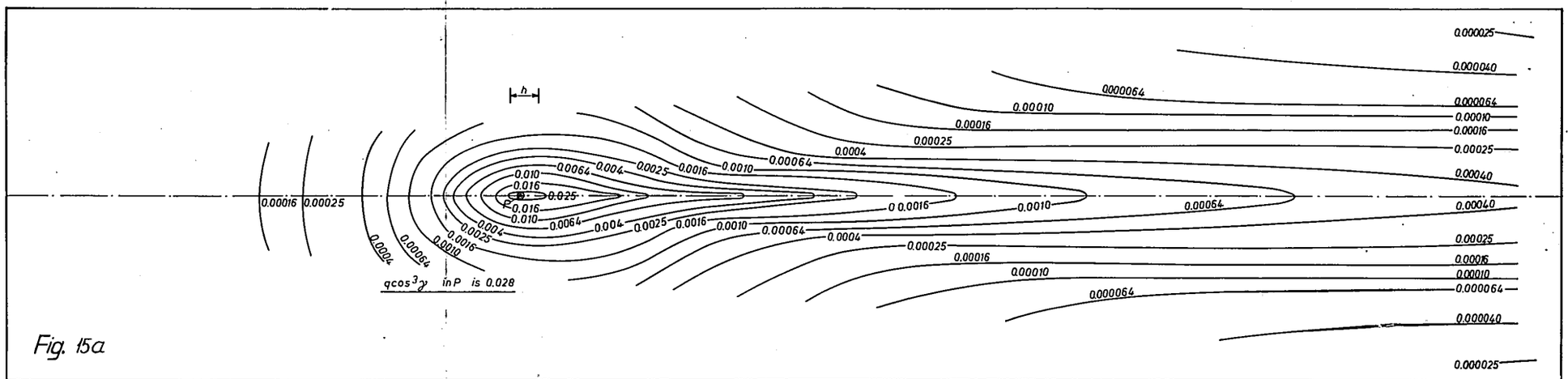



Fig. 15a

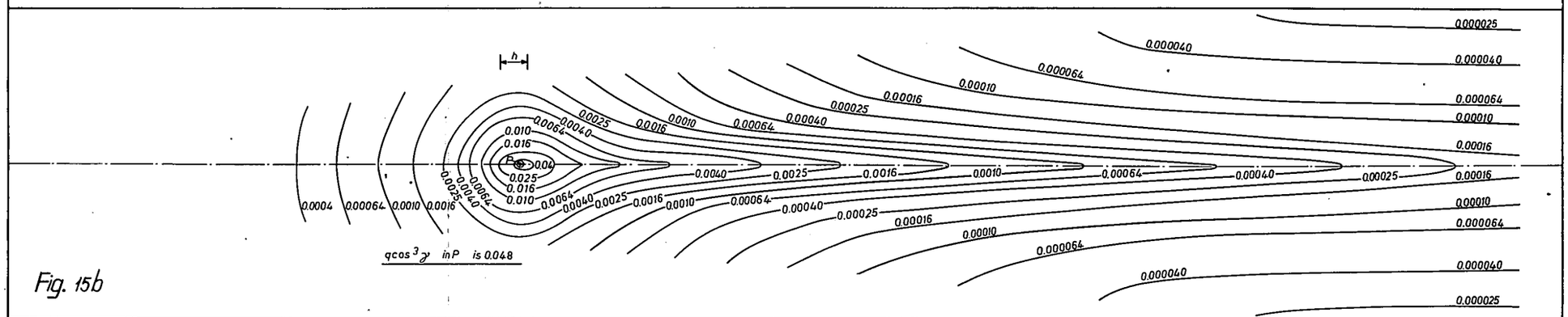



Fig. 15b

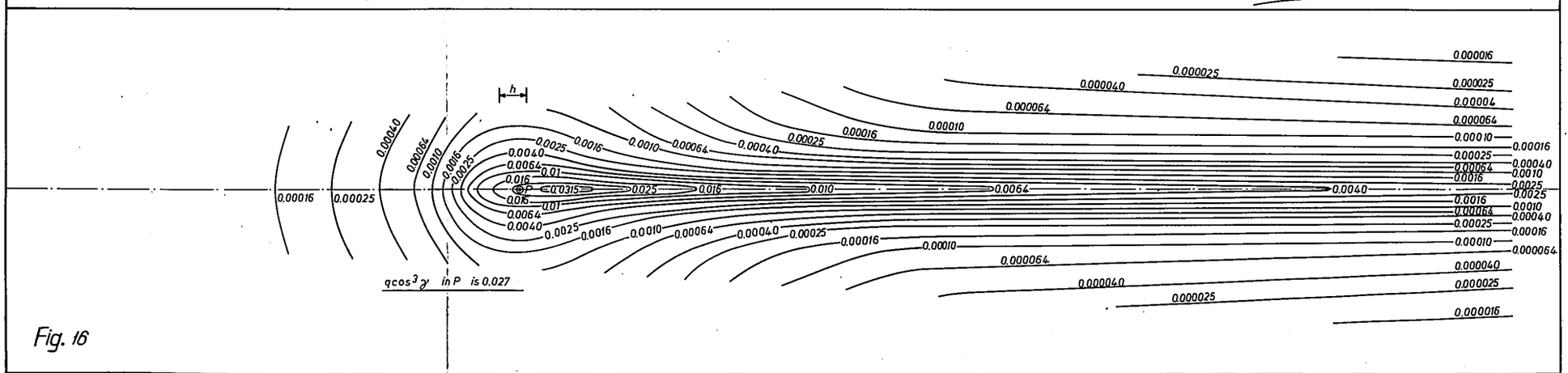



Fig. 16

70557

Fig. 15*a*. E. P. Diagram for dry asphalt (8 years old, total reflection factor 8 per cent, deduced from measurements of Reid and Chanon [7]).

The plane of the drawing represents the road surface. The observer is over the line of symmetry of the diagram to the left of the point $P$ considered, at a distance from $P$ which is at least 40 × the level of his eye. The curves connect projections on the road surface of lamp positions having one value of $q \cos^3\gamma$ (indicated by the figures at the curves). $q$ is the luminance coefficient in cd/m² lux and $\gamma$ is the angle of incidence of the light. The unit of length in the diagram is equal to the height $h$ of the lamp above the road surface (indicated by a short line in the diagram).

*Directions for use*

Plot a plan of the road on a scale such that the distance $h$ of the E.P. diagram equals the fixture height. Indicate in this plan the positions of the lamps and of the observer as well as the points on the road observed for which the luminance has to be calculated. Lay the plan over the E.P. diagram in such a way that one of these points coincides with the point $P$ of the E.P. diagram, while the observer is over the line of symmetry of the E.P. diagram. Read the value of $q \cos^3\gamma$ for each lamp. Determine the luminous intensity radiated by each lamp in the direction of the point considered (e.g., by means of an iso-candela diagram). Multiply this value by the corresponding value of $q \cos^3\gamma$ and divide the product by $h^2$. The result gives the contribution of each lamp to the luminance at the point considered.

Fig. 15*b*. E. P. Diagram for dry concrete (9 years old, total reflection factor 22 per cent, deduced from measurements of Reid and Chanon [7]).

For further description and directions for use, see legend to fig. 15*a*.

Fig. 16. E. P. Diagram for dry asphalt (from the author's own measurements).

For further description and directions for use, see legend to fig. 15*a*.

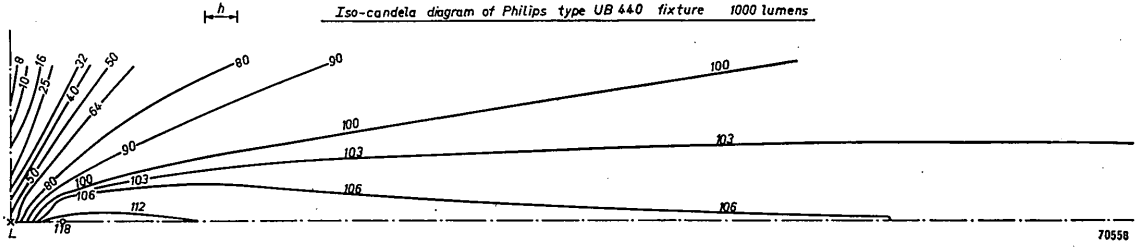



Fig. 17. Iso-candela diagram of U.B. fixture for four 40-W fluorescent tubes. The plane of drawing represents the road surface. The curves connect points towards which the lamp radiates equal luminous intensities.

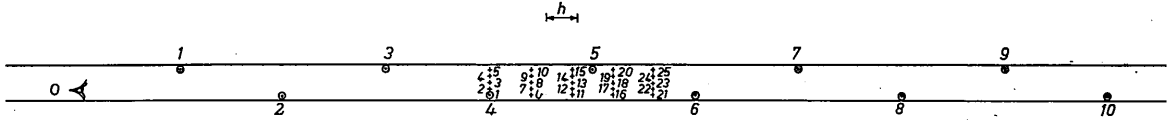



Fig. 18. Plan of the road with lighting installation for which the luminance distribution is calculated in the appendix.

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

Reprints of the majority of these papers can be obtained on application to the Administratiou of the Research Laboratory, Kastanjelaan, Eindhoven, Netherlands. Those papers of which no reprints are available are marked with an asterisk (*).

**1947:** N. W. H. Addink: A possible correlation between the zinc content of liver and blood and the cancer problem (Nature **166**, 693-1950).

The zinc content of various human tissues and blood is determined spectrochemically. In the presence of malignant growth the blood is found to contain $16 \times 10^{-4}$ % of zinc (average of 6 determinations, S.D. of average $6 \times 10^{-4}$ %) as against $7 \times 10^{-4}$ % (6 determinations, S.D. of average $1 \times 10^{-4}$ %) in normal cases and $12 \times 10^{-4}$ % (7 determinations, S.D. of average $1 \times 10^{-4}$ %) in the case of non-tumorous disease, whilst in liver tissue with malignant growth $75 \times 10^{-4}$ % (6 determinations, S.D. of average $12 \times 10^{-4}$ %) is found, as against $53 \times 10^{-4}$ % (7 determinations, S.D. of average $5 \times 10^{-4}$ %) in the case of non-tumorous disease.

**1948*:** W. de Groot and A. A. Kruithof: De kleurendriehoek (Ned. T. Natuurk. **16**, 235-241, 1950). (The colour triangle; in Dutch.)

Description of a device consisting of a system of three slits provided with diffusing colour filters, a triangular screen and a triangular diaphragm covering a plate of opal glass. The colour triangle appears on the opal glass with the different colours in their proper places, when the three slits are illuminated by an incandescent lamp.

**1949:** J. Volger: Note on the Hall potential across an inhomogeneous conductor (Phys. Rev. **79**, 1013-1024, 1950).

The change in the potential function in a homogeneous conductor of rectangular shape due to the Hall effect is calculated to the first approximation. Some important conclusions are drawn for the case of an inhomogeneous conductor.

**1950:** R. van der Veen: Induction phenomena in photosynthesis, III (Physiologia Plantarum **3**, 247-257, 1950).

It is shown that the reaction of Chlorella to sudden illumination is much different from that of higher plants. While higher plants always begin with an initial uptake of $CO_2$, Chlorella does exactly the opposite and begins with a release of $CO_2$ before it starts photo-synthesizing. This

sudden release of $CO_2$ is positively correlated with light intensity, negatively with temperature and positively with the duration of the preceding dark period. Another difference between higher plants and Chlorella is the "de-adaptation" in darkness. In Chlorella there is not much difference in adaptation time after short and after long dark periods. In higher plants, however, after long periods of darkness adaptation takes much longer than after short periods of darkness. A third difference is the behaviour of Chlorella in pure hydrogen without any $CO_2$. In this environment higher plants react to illumination with a sudden release of oxygen, which could not be found when examining Chlorella. Thus Chlorella has a rather different pattern of photosynthesis than higher plants. Protococcus olivaceus Rabenh., which, like Chlorella, belongs to the Protococcales, behaves in the same way as Chlorella. Ulothrix sp., like Hormidium flaccidium and Stichococcus bacillaris, all three belonging to the Ulotrichales, have induction lines like those of higher plants.

**1951:** H. O. Huisman: Investigations on quinones and quinone derivatives. Preparation and antibiotic properties of some substituted p-benzo- and p-toluo-quinones, hydroquinones and hydroquinone esters (Rec. Trav. chim. Pays-Bas **69**, 1133-1156, 1950).

The antibiotic properties of the compounds named in the title are tested in vitro against a yeast, various moulds and Gram-positive and Gram-negative micro-organisms. The toxicity is determined in white mice by oral administration and by interperitoneal injection. The chemotherapeutic activity is investigated in white mice injected with a lethal dose of pneumococci.

**1952*:** P. Cornelius: Electrische eenheden en het Giorgistelsel (Electrotechniek **28**, 454-460, 473-479, 1950). (Electrical units and the Giorgi system; in Dutch.)

Review of the Giorgi system of units.

**1953:** J. M. Stevels: Le fer dans le verre (Verres et Réfractaires **4**, 293-298, 1950). (Iron in glass; in French.)

The behaviour of iron in glass is intimately related to the variable valence of the iron atom, which makes electronic transitions possible. In network-forming positions $Fe^{3+}$-ions cause a brownish colouring, while in network-modifier positions there is practically no colouring. $Fe^{2+}$-ions seldom occur alone and, if so, mostly as network modifiers, but sometimes also as network formers. These ions do not cause a visible colouring. The bluish-green colour of some glasses containing iron is due to the simultaneous action

of $Fe^{2+}$- and $Fe^{3+}$-ions. Indications exist that at least one of these ions (or both) must be present in network-forming positions. The influence of $Ti^{4+}$ is of importance in this respect because this ion shifts the equilibrium between the network-modifier positions of Fe-ions in the latter direction. It may be that decolouring of glass simply means elimination of $TiO_2$. A further study of the equilibrium $Fe^{2+} \rightleftarrows Fe^{3+}$ in both positions is needed.

**1954*:** A. Claassen: Titan (Hb. Anal. Chem., Dritter Teil, Band IVb, Springer, Berlin 1950; pp. 1-169).

Report on quantitative analytical methods for the determination of Ti and the various methods of separation.

**1955*:** A. Claassen: Zirkon und Hafnium (Hb. Anal. Chem., Dritter Teil, Band IVb, Springer, Berlin 1950, pp. 170-287).

Report on quantitative methods for the determination of Zr and Hf and the current methods of separation.

**1956:** J. L. Snoek: Density variations in aluminium (Phil. Mag. **41**, 1188-1192, 1950).

From careful determinations of the density of samples of aluminium in various conditions the conclusion is drawn that grain boundaries in aluminium may have a more spacious structure than is generally assumed. In an addendum the possible influence of gas bubbles ($H_2$) on the density is discussed.

**1957:** C. J. Bouwkamp: On Sommerfeld's surface wave (Phys. Rev. **80**, 294, 1950).

A critical discussion of some recent papers by Kahan and Eckart on Sommerfeld's surface wave in electromagnetic wave propagation over a plane earth.

**1958:** J. M. Stevels: The relation between the dielectric losses and the composition of glass (J. Soc. Glass Technology **34**, 80-100, 1950).

Discussion of the theory of dielectric losses in glass. Losses are expected due to (1) relaxation phenomena caused by the network modifiers, (2) relaxation phenomena caused by the network itself and (3) vibration of the constituents of the network and of the network modifiers. The dependence of each kind of losses on the frequency and the temperature is discussed. The theoretical considerations are compared with the experimental results, and the dependence of the various kinds of losses on the composition is discussed.

**1959:** F. A. Kröger and J. van den Boomgaard: The fluorescence of magnesium germanate activated by manganese (J. electrochem. Soc. **97**, 377-382, 1950).

Magnesium germanate activated by manganese may contain the activator in different valencies. When this valency is higher than two, fluorescence excited by ultraviolet occurs in seven narrow bands decaying with the same time constant. The variation of this decay constant and the intensity of fluorescence with temperature is explained by the assumption that the relative contribution of radiative and non-radiative transitions is determined by a Boltzmann equilibrium regulating the population of various excited states.

**1960:** J. A. Haringx: The rigidity of corrugated diaphragms (Appl. sci. Res. **A 2**, 299-325, 1950).

When the corrugated diaphragm is replaced by a fictitious flat plate of similar properties it is possible to derive a linear differential equation for the deflection. The coefficients of this equation, however, vary in a complicated way and its solution for the pressure-loaded diaphragm is given only for thick and for thin sheets separately. For thick sheets the profile of the corrugation appears to be inessential, whereas for thin sheets it is necessary to distinguish between trapezoidal, triangular and arc-shaped corrugations. By an obvious device the results for thick and for thin sheets are fitted together, so that the deflection can be determined also for the intermediate range of medium sheet thickness. The final results of the present calculation are compared with measurements carried out by others and are found to be in satisfactory agreement with the experiments. It is to be remarked that, compared on the basis of small deflections, the introduction of corrugations into the sheet leads to a considerable increase of rigidity of the diaphragm. The prevailing assertion that the flat plate is more rigid than the corrugated diaphragm holds only for large deflections, because of the non-linearity between the load and the deflection of the flat plate.

**1961:** J. L. Snoek and E. J. Haes: Laboratory apparatus for the production of a steady flow of very pure hydrogen (Appl. sci. Res. **A 2**, 326-238, 1950).

An apparatus is described, containing a diffusor element of twelve thin-walled, nickel tubes in parallel, which, upon being heated to 950 °C and being exposed to an excess pressure of about one atmosphere, yields a steady flow about of one quarter of a litre per minute of the purest hydrogen at atmospheric pressure. The apparatus is rugged and simple and has a lifetime of several hundreds of hours.

VOL. 7 NR 2 APRIL 1952

# Philips Research Reports

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

*R 185* *Philips Res. Rep. 7, 81-111, 1952*

## A DECADE COUNTER VALVE FOR HIGH COUNTING RATES

by J. L. H. JONKER, A. J. W. M. van OVERBEEK and P. H. de BEURS

621.318.572:621.385.832

**Summary**

The Philips decade counter valve is a cathode-ray tube with a ribbon-shaped electron beam and one-dimensional deflection, in which the beam current is distributed over several electrodes of a special shape while, by means of a negative feed-back from one of these electrodes to a deflection electrode, the beam can be fixed in each of ten discrete positions. In each of these positions a part of the beam impinges on a fluorescent part of the wall of the valve, so that the instantaneous position of the beam can be read in luminous figures. The beam moves from one position into another by means of an impulse on one of the deflection electrodes. As these counter valves can be connected in cascade, counting and calculating devices may be constructed with them. The small dimensions, the clear indication and the extremely low power consumption when fed by a voltage supply of 300 volts make the counter valve very attractive for many applications. For pulses occurring at random we measured a mean resolving time of the counter valve in its simplest connexion of about 15 μsec. With the aid of one double-triode valve this could be improved to 6 μsec. With a special circuit incorporating a triode-hexode and a secondary-emission valve it could be reduced to even less than 0·2 μsec. For pulses occurring with equal intervals the resolving time is larger.

**Résumé**

Le tube compteur décimal Philips est un tube à rayons cathodiques avec un faisceau électronique en forme de ruban et une déflexion unidimensionnelle, dans lequel le courant du faisceau est distribué sur plusieurs électrodes de forme spéciale tandis qu'à l'aide d'une contre-réaction négative de l'une de ces électrodes sur une électrode de déflexion, le faisceau peut être fixé dans dix positions différentes. Dans chacune de ces positions, une partie du faisceau tombe sur une partie fluorescente de la paroi du tube, de sorte que la position instantanée du faisceau peut être lue en indication lumincuse. Le faisceau se déplace d'une position à l'autre au moyen d'impulsions sur l'une des électrodes de déflexion. Comme ces tubes compteurs peuvent être reliés en cascade, ils peuvent servir à construire des appareils à compter et calculer. Les faibles dimensions, l'indication nette, et la puissance consommée extrêmement faible lorsque la tension d'alimentation est de 300 volts, rendent ce tube compteur extrêmement intéressant pour de nombreuses applications. Pour des impulsions réparties selon les lois du hasard, nous avons observé un temps de résolution moyen du tube compteur dans son montage

le plus simple d'environ 15 μsec. A l'aide d'une double triode, on peut améliorer ce temps jusqu'à 6 μsec. A l'aide d'un circuit spécial comprenant une triode-hexode et un tube à émission secondaire, il peut même être réduit à moins de 0,2 μsec. Pour des impulsions apparaissant à intervalles réguliers, le temps de résolution est supérieur.

**Zusammenfassung**

Das Philips Dezimal-Zählrohr ist eine Kathodenstrahlröhre mit einem bandförmigen Elektronenstrahl und eindimensionaler Ablenkung. Hierbei ist der Bündelstrom über mehrere Elektroden von besonderer Form verteilt, wobei mittels einer negativen Rückkopplung von einer dieser Elektroden zu einer Ablenkungselektrode der Strahl auf jede einzelne von zehn gesonderten Stellungen gerichtet werden kann. In jeder dieser Stellungen trifft ein Teil des Strahles einen fluoreszierenden Teil der Röhrenwand, so daß die jeweilige Lage des Strahles an Leuchtfiguren abgelesen werden kann. Der Strahl bewegt sich von einer Stellung zu einer anderen durch die Wirkung eines Impulses von einer der Ablenkungselektroden. Wenn diese Zählrohre hintereinandergeschaltet werden können, ist man imstande, mit ihnen Zähl- und Rechnungsbehelfe zu konstruieren. Dank den geringen Abmessungen, der deutlichen Anzeige und dem äußerst niedrigen Kraftverbrauch bei Speisung mit einer Spannung von 300 Volt ist das Zählrohr für viele Anwendungen besonders gut geeignet. Für beliebig auftretende Impulse maßen wir eine durchschnittliche Lösungszeit des Zählrohres in seiner einfachsten Schaltung von ungefähr 15 μsec. Mit Hilfe einer Doppeltriode ließ sich diese Zeit auf 6 μsec herabsetzen. Mit einem Spezialkreis, der eine Triode-Hexode und eine Sekundäremissionsröhre enthält, konnte sie sogar bis auf 0,2 μsec reduziert werden. Bei Impulsen, die in gleichen Intervallen auftreten, ist die Lösungszeit länger.

## 1. Introduction

In the last decade electronic devices for measuring, controlling and regulating have come into use more and more. Examples are: counting devices for number-recording, differentiating and integrating devices.

These devices often possess a large number of common amplifying valves for relatively simple functions. The aim has therefore been to replace these valves by one valve specially constructed for the purpose, and it has proved that one single valve can supersede a complete counting circuit [1]), [2]), [3]).

Such a valve will be described here. Fig. 1 shows a photograph of it. It has small external dimensions, yet its electrode system resembles that of a cathode-ray tube. A screen electrode with ten slots in it is placed in front of the anode, so that a ribbon-shaped electron beam can reach the fluorescent screen only in ten places. A negative feed-back from the anode to one of the deflection electrodes makes it possible to fix the beam in any of these ten places. The beam can be brought from one position into the next by applying a pulse of a special shape to one of the deflection electrodes. The final position of the beam will thus be a measure for the number of pulses received.

## 2. Construction of the decade counter valve

If the voltage $V_{d_1}$ of one of the deflection plates of a cathode-ray tube

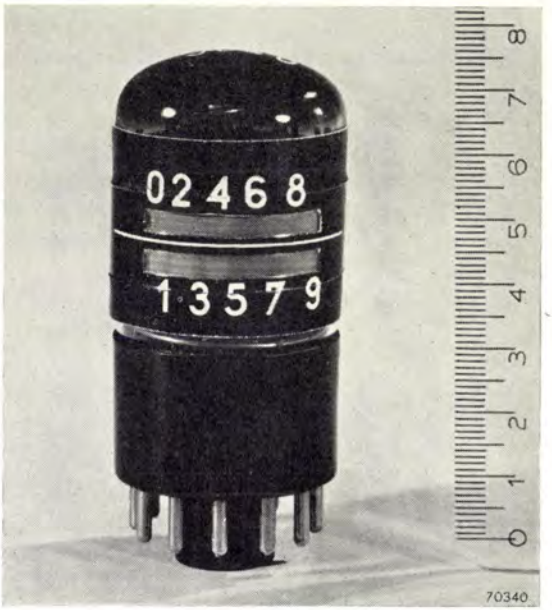



Fig. 1. A Philips decade counter valve.

is kept constant, and the anode current is measured as a function of the voltage $V_{d_2}$ on the other deflection plate, then the characteristic $I_a = f(V_{d_2})$ will be a straight line parallel to the voltage axis: $I_a = I_0$ (fig. 2, line 1). If the anode is connected to the latter deflection plate and both these are connected via a resistance $R_a$ to the battery, then we have in addition

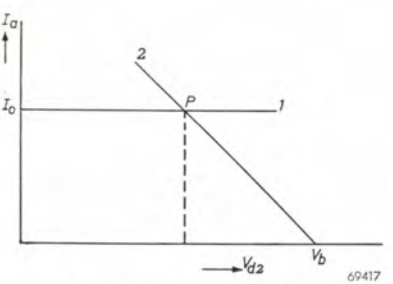



Fig. 2. Anode current as a function of the deflecting voltage $V_{d_2}$ when $V_{d_1}$ is constant, 1 with constant $V_a$, 2 with the anode connected to the deflection plate $d_2$ and fed via a resistance $R$, so that $V_{d_2} = V_b - I_a R$, where $V_b$ = battery voltage. At the point $P$ the electron beam in is a stable position.

$$V_{d_2} = V_b - I_a R_a ,$$

where $V_b$ is the battery voltage (fig. 2, line 2). One stable adjustment is then possible which satisfies both equations for $I_a$. In fig. 2 this adjustment is represented by the point of intersection $P$ of the lines 1 and 2.

For the decade counter valve the function $I_a = f(V_{d_2})$ can be represented by the wave line given in fig. 3. There are 19 points of intersection with the characteristic $V_{d_2} = V_b - I_a R_a$. Ten of these correspond to a stable adjustment, while in the case of the other nine a small change in $V_{d_2}$ will result in a further change in the same direction, so that these nine points are unstable.

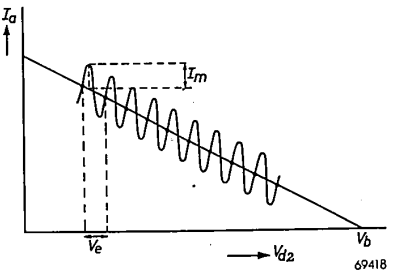



Fig. 3. For the decade counter valve the function $I_a = f(V_{d_2})$ is given by the wave line. There are 19 points of intersection with the characteristic $V_{d_2} = V_b - I_a R$, of which 10 points (marked in the figure) represent the stable positions of the electron beam and 9 points the unstable ones.

Fig. 4 gives the simplified cross-section of a valve construction with which such an undulatory characteristic can be obtained. This figure shows the electrodes for the formation of the ribbon-shaped beam, namely the cathode $c$, the focusing electrodes $p_1$ and $p_2$ [4]), the gun anode $a_0$ and the two

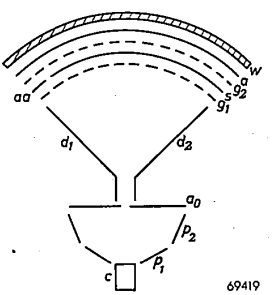



Fig. 4. Simplified cross-section of the counter valve. $c$ cathode, $p_1$ and $p_2$ focusing electrodes, $a_0$ gun anode, $d_1$ and $d_2$ deflection plates, $s$ slotted screen, $a$ anode (slotted), $aa$ auxiliary anode, $w$ glass wall partly provided with a conductive fluorescent layer, $g_1$ and $g_2$ grids to suppress secondary emission.

IPR2025-00217 EXHIBIT 2010 Page 88 of 485

deflection plates $d_1$ and $d_2$. The beam is adjusted so as to have its focus on a screen electrode $s$ provided with slots. A view of this screen is given in fig. 5. By variation of the voltage on the plate $d_2$ the line focus of the beam can be moved across the screen. According to the position of the beam a greater or smaller part of the beam current will be allowed to pass through a vertical slot in the screen, which thus regulates the current to the anode $a$. This causes the undulatory variation of the anode current. The wave line can be made to descend gradually with the aid of horizontal slots partly covered with gauze.

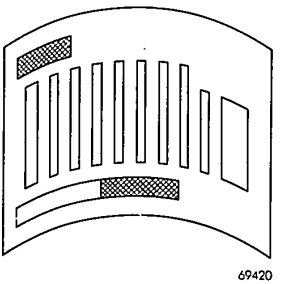



Fig. 5. The screen $s$ of fig. 4 seen from the cathode (auxiliary anode to the left). By varying the deflecting voltage the focus of the beam can be moved across this screen, thus regulating the current to the anode according to fig. 3.

In fig. 4 it can also be seen how the indication is brought about. When the beam is in a stable position part of the beam current will pass through openings in the anode $a$ to a part of the glass wall $w$ covered with a conductive layer on which a fluorescent substance has been sprayed, which wil clearly illuminate a figure corresponding to the position of the beam. The grids $g_1$ and $g_2$ are at cathode potential and serve as secondary-emission suppressors.

### 3. Operation of the decade counter valve

Fig. 6 shows the characteristic $I_a = f(V_{a+d_2})$ measured on a practical valve, the load line $R_a$ (680 kΩ) and the auxiliary-anode current $I_{aa}$. The purpose of the auxiliary anode will be explained later. The supply diagram of the valve is given in fig. 7.

From fig. 6 it is seen that the load line $R_a$ has 19 points of intersection with the characteristic of the counter valve. Ten of these points represent stable adjustments; they have been marked in the diagram. A small displacement of the beam from a stable position to the left or to the right will give, respectively, a smaller or greater anode current. This means an increase or decrease, respectively, in the voltage of the right-hand deflection

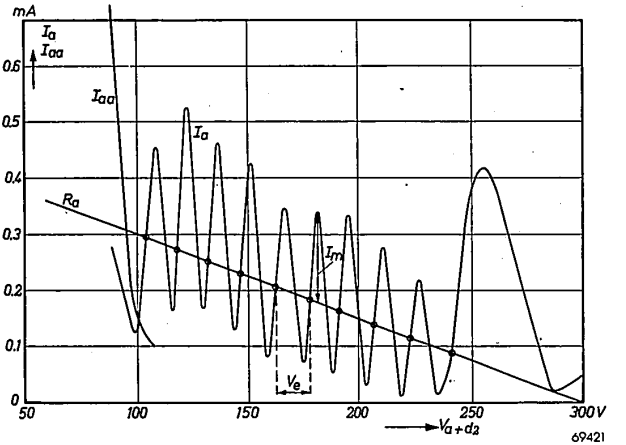



Fig. 6. The characteristics $I_a$ and $I_{aa}$ as functions of $V_{a+d_2}$ measured on a practical valve. $V_b = V_{a0} = 300$ V, $V_{aa} = 300$ V, $V_{d_1} = 150$ V, $V_s = 300$ V via $R_4 = 39$ kΩ (see fig. 7), $R_a = 680$ kΩ (fig. 7), $I_{aa\,max} = 1{\cdot}26$ mA for $V_{a+d_2} = 75$ V. The marked points indicate the stable positions of the beam.

plate, which will counteract the displacement. Thus a small displacement of the beam from the stable position will give $d_2$ a voltage variation which brings the beam back to the equilibrium position. By analogous reasoning it follows that a small displacement of the beam from one of the nine other

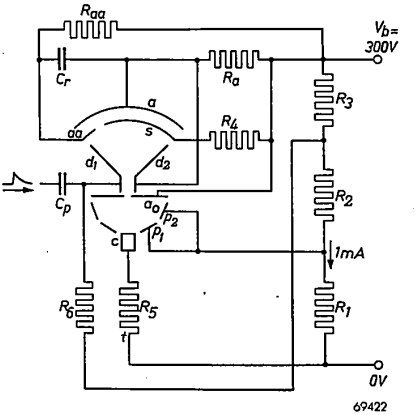



Fig. 7. Supply diagram of the counter valve for measuring the characteristic of fig. 6. Electrode indications as in fig. 4. $V_b = 300$ V, $R_a = 680$ kΩ, $R_{aa} = 470$ kΩ, $R_1 = 10$ kΩ, $R_2 = 150$ kΩ, $R_3 = 150$ kΩ, $R_4 = 39$ kΩ, $R_5 = 8$ kΩ, $R_6 = 33$ kΩ, $C_r = 100$ pF, $C_p = 560$ pF. Pulses of the shape as in fig. 8 are applied to the deflection plate $d_1$ via the capacitor $C_p$. Each pulse causes the beam to move up one step to the next stable position.

points will result in a change in the voltage of $d_2$ which will remove the beam farther from the unstable position.

The operation of counting can now be explained with the aid of figs 6 and 7. It should be remarked that in the circuit of fig. 7 only one supply voltage of 300 volts is used. The d.c. voltages for the deflection plate $d_2$ and the cathode plates $p_1$ and $p_2$ are obtained with the aid of the potentiometer circuit $R_1$, $R_2$, $R_3$. The slotted screens is connected to the voltage supply via the resistor $R_4$, while the gun anode $a_0$ is connected to it directly.

The counting anode $a$, together with the deflection plate $d_2$, is connected to the voltage supply via the resistor $R_a$ and coupled with the auxiliary anode $aa$ via a capacitor $C_r$.

To bring the valve into operation pulses of a shape as shown in fig. 8 are applied to the deflection plate $d_1$. The leading slope of these pulses must be much steeper than their tail slope. To explain this, let us assume that the beam is in a stable position in a circuit as in fig. 7. To $d_1$ we apply a pulse as in fig. 8. The crest voltage $V_i$ (e.g., 15 volts) of this pulse must be about equal to the voltage distance $V_c$ between two adjacent stable positions of fig. 6. The leading slope of the pulse is so great that owing to the parasitic capacitances the potential of $d_2$ cannot immediately readjust itself; it practically keeps its original value, so that as a result of the pulse voltage the beam moves up one place to the left. The following decrease of the pulse voltage, on the other hand, is so slow that the potential $V_{d_2}$ changes practically as fast as that of $d_1$ and in the same sense. This causes the beam to stay in this new stable position.

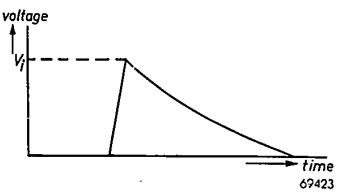



Fig. 8. Shape of the pulses applied to the deflection plate $d_1$ in fig. 7.

Practical demands for the pulse appear to be:

(1) leading slope $\mathrm{d}V/\mathrm{d}t \geqq 4(I_m/C_a)$;

(2) tail slope $-(\mathrm{d}V/\mathrm{d}t) \leqq \frac{1}{2}(I_m/C_a)$;

(3) $\frac{1}{2} \times$ voltage distance $V_e <$ crest voltage $V_i < \frac{3}{2} V_e$,

where $I_m$ is the amplitude of the wave of the characteristic of fig. 6 and $C_a$ the total capacitance of $a$ and $d_2$, parasitic capacitances included.

For counting in the decimal system up to numbers of more than one digit it is necessary to connect a series of these counter valves in cascade.

With the tenth pulse the beam of the first valve must then be brought back into the zero position and a pulse has to be applied to the next valve. This flyback is performed by the auxiliary anode *aa* (fig. 7), on which at the tenth pulse the full beam current falls (see fig. 6), in contradistinction to the anode *a*, on which through the slots of the screen *s* never more than a small part of the beam can fall.

In the last stable position the anode current is, e.g., only one-fifth of the full beam current. When the full beam impinges on the auxiliary anode the latter will be charged negatively, so that its voltage will drop considerably. In the coupling capacitor $C_r$ a displacement current will then arise which is greater than the anode current was in the last stable position, so that the anode voltage and also $V_{d_2}$ will further decrease and the beam will move towards $d_1$. The beam will then stop and move in the opposite direction, owing to the following effects:

(1) upon reaching the edge of $d_1$ the beam current will partly pass on to this deflection electrode;
(2) as $V_{aa}$ decreases $I_{Raa}$ increases and becomes an increasing part of the beam current; so less displacement current will be available;
(3) at a low value of $V_{aa}$ only a part of the beam current will reach the auxiliary anode (operation beyond the knee of the $I_{aa}$-$V_{aa}$ characteristic).

Apparently owing to these effects the current through $C_r$ and $R_a$ will decrease, $V_{d_2}$ will increase and the beam moves to the right across *aa*. When the beam leaves *aa* the current through the capacitor $C_r$ will be in the opposite sense, giving $d_2$ an extra positive voltage pulse and throwing the beam to the right. This will be discussed in more detail in section 5.

## 4. Computation of the counting speed *)

Fig. 9 is a simplified diagram of the counting circuit when the beam jumps from one position into the next during the counting. Fig. 10 gives the

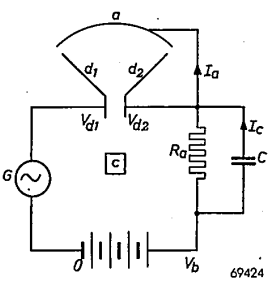



Fig. 9. Part of the diagram of fig. 7, needed for the computation of the counting speed.

*) For a list of symbols; see appendix.

corresponding characteristic of the anode current $I_a$ as a function of the difference $V_{d_2} - V_{d_1}$ of the voltages of the two deflection plates. This characteristic can be imagined as being the sum of a sinusoidal and a linear part. It can then be represented by the formula

$$I_a = [V_b - V_{d_2} + V_{d_1}]/R_a + I_m \sin [2\pi(V_{d_2} - V_{d_1})/V_e],$$

where $V_b$ is the difference of the supply d.c. voltages of the deflection plates and $V_e$ the voltage difference between two stable positions. Here $V_{d_1}$ and $V_{d_2}$ are determined with respect to the point $O$ (fig. 9). Moreover, for the outer circuit we have

$$I_a = \frac{V_b - V_{d_2}}{R_a} + I_C,$$

so that

$$I_C = \frac{V_{d_1}}{R_a} + I_m \sin [2\pi\,(V_{d_2} - V_{d_1})/V_e]\,.$$

If the generator $G$ (fig. 9) gives a voltage $V_{d_1} = V_G$, then

$$I_C \equiv -C\,\frac{\mathrm{d}V_{d_2}}{\mathrm{d}t} = \frac{V_G}{R_a} + I_m \sin [2\pi\,(V_{d_2} - V_G)/V_e]\,.$$

From this differential equation $V_{d_2} = \varphi(t)$ can in principle be computed, if $V_G = \Psi(t)$ is given.

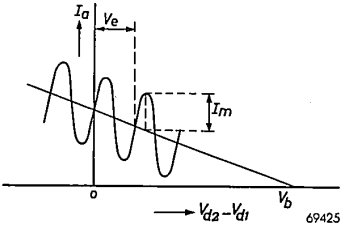



Fig. 10. Computation of the counting speed.

The first term of the right-hand member will in our case be small in comparison with the second term (fig. 10), and to simplify the equation we drop this first term. In our calculation we neither take into account the influence of the flyback circuit. To obtain the value of $C$ we add the parallel capacitance $C_{aa}$ (see fig. 14) of the flyback anode to that of the anode $a$. Moreover we ignore the influence of the resistance $R_{aa}$. This gives about the same error as when disregarding the influence of $R_a$ by dropping the first term of the right-hand member of the differential equation; so to obtain a simple solution giving an impression of the order of magnitude of the counting speed we put $R_{aa} = \infty$, $R_a = \infty$. Now the function $I_a = f(V_{d_2} - V_{d_1})$ changes as shown in fig. 11. The full-drawn curve represents the

new characteristic, the dashed curve the old one. The differential equation now becomes:

$$\frac{dV_{d2}}{dt} + \frac{I_m}{C} \sin\left[2\pi(V_{d2} - V_G)/V_e\right] = 0.$$

Let

$$2\pi(V_{d2} - V_G)/V_e = q, \qquad CV_e/2\pi I_m = r$$

and

$$\frac{2\pi}{V_e}\frac{dV_G}{dt} = -u = \text{constant},$$

so that $V_G$ is a linear function of the time $t$. Then

$$t - t_0 = r \int_{q_0}^{q_t} \frac{dq}{ur - \sin q},$$

where $q_0$ is the value of $q$ at the moment $t_0$ and $q_t$ that at the moment $t$.

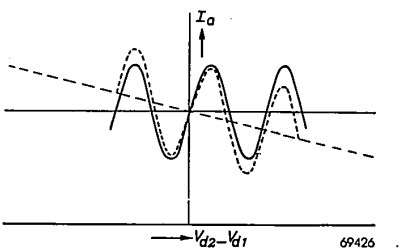



Fig. 11. Computation of the counting speed.

Fig. 12 shows how the solution of this integral is found. The quantity $(ur - \sin q)^{-1}$ has been plotted as a function of $q$. Thus the area bounded between two values of $q$, the $q$-axis and the curve gives the value $(t - t_0)/r$.

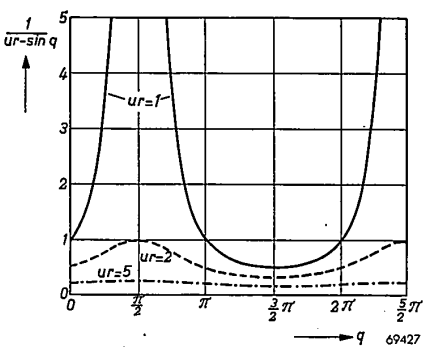



Fig. 12. The function $f(q) = (ur - \sin q)^{-1}$ occurring in the computation of the counting speed

For these values we take $q = 0$ and $q = 2\pi$. In this way we find the time needed by the beam to move from one stable position into the next when $ur$ is given, i.e. with a certain $r$ given by the circuit and valve characteristic and a certain $u$ given by the shape of the pulse.

It is to be noted that for values of $ur$ less than unity there is no displacement of the beam. The time required for this when $ur = 1$ would be infinitely long. This reasoning, however, is not quite correct, as in that case the term $V_G/R_a$ may not be ignored. This case, however, is of no importance, since the height of the pulse then becomes very great. For $ur$ the greater time required appears to become small.

Summarizing, from fig. 12 it may be concluded that for the leading slope of the pulse we must have $ur > 1$ and for the initial tail slope $ur < 1$.

Fig. 13 gives the result of a graphical determination of the function $u(t-t_0) = \int_0^{q_t} ur\, dq/(ur-\sin q)$ for different values of $ur = -(C/I_m)\,(dV_G/dt)$. The quantity $ur$ is determined by the leading or the tail slope of the pulse, while $u(t - t_0) = 2\pi/V_e\ (V_{C0} - V_{G_t})$ is a measure for the difference between the initial and the final value of $V_G$, i.e. for the crest voltage of the pulse. The value of $q_t$ determines the position of the beam.

When $q_t = 0, \pm 2\pi, \pm 4\pi, \ldots$, the adjustment is stable, and the purpose is to go from one stable position to the next by means of the pulse. The pulse is assumed to have a triangular shape with a leading slope greater than the tail slope. In the graph of fig. 12 this means that the leading slope must be such that, beginning at the stable position $q = 0$, within a reasonable

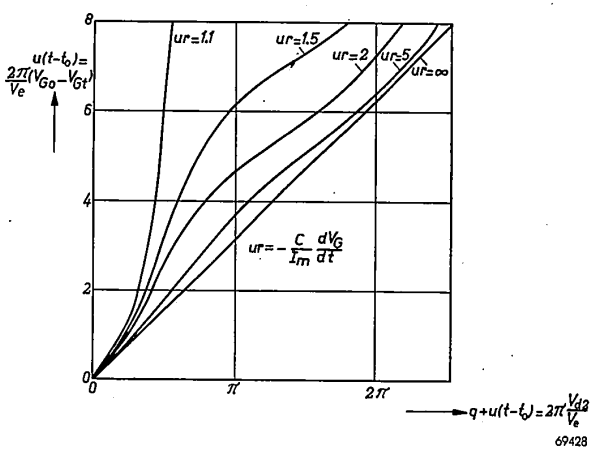



Fig. 13. Result of the graphical determination of the function $u(t-t_0) = \int_0^{q_t} ur(ur-\sin q)^{-1}\, dq$ occurring in the computation of the counting speed.

time the line $q_t = \pi$ must be passed. Moreover, $ut$ may not become too great. For $ur = 2$ these conditions are fulfilled. The beam is then beyond the unstable point of intersection of the load line and the characteristic of the valve.

In practice a further demand will be that, notwithstanding a certain divergency in $I_m$ and $V_c$, the pulse displaces the beam from one position to the next. Taking these variations into account, we arrive at the following conditions:

for the leading slope: $|ur| \equiv \left|\frac{C}{I_m}\frac{\mathrm{d}V_G}{\mathrm{d}t}\right| > 3$;

for the tail slope : $|ur| \equiv \left|\frac{C}{I_m}\frac{\mathrm{d}V_G}{\mathrm{d}t}\right| < 0{\cdot}6$.

When, e.g., the average $I_m$ is 0·1 mA, $V_c = 12$ V and $C = 25$ pF,

the leading slope $= 1{\cdot}2.10^7$ V/sec

the tail slope $= 2{\cdot}4.10^6$ V/sec.

With this shape of the pulse we get for the pulse $ut$:

14 V $<$ crest voltage $<$ 24 V.

In practice this demand is also influenced by the fact that the anode resistance $R_a$ and the resistance $R_{aa}$ of the auxiliary anode are parallel via the capacitance $C_r$ (fig. 7). Moreover the boundaries for the crest voltage of a pulse with steeper leading slope turn out to become lower and comparatively wider.

A superposition of similar pulses will result in a greater tail slope. With maximum tail slope the resolving power will equal approximately the descending time of the pulse, thus in the case given above about 6 μsec. If the tail slope is less than half or one-third the permissible maximum value, then a superposition of 2 or 3 pulses will not cause any "fall-back", but in that case the difference between the boundaries between which the pulse voltage has to be kept becomes smaller.

Since in practice the counting-speed, especially for equally spaced pulses, is mainly determined by the flyback time of the beam from the last position (9) to the first (0), it is necessary to consider more closely the auxiliary anode and the flyback of the beam. We then get an impression of the magnitude of the coupling capacitance between anode and auxiliary anode and of the time needed for the full flyback.

## 5. The flyback time

### (a) *Method of calculation*

Fig. 14 shows the anode, the auxiliary anode and the external circuit with the anode resistance $R_a$, the auxiliary-anode resistance $R_{aa}$, the

flyback capacitance $C_r$, the anode capacitance $C_a$ and the capacitance $C_{aa}$ of the auxiliary anode.

Kirchhoff's equations can be drawn up for this circuit. The currents to counting anode and auxiliary anode, however, are non-linear functions of the voltage $V_{d_2}$, giving rise to non-linear differential equations. In principle it is possible to solve these differential equations for all values of the capacitances and resistances and for all shapes of the characteristics mentioned. This, however, would not, or only with much trouble, give an insight into the influence of variations in the characteristics and in the impedance values upon the flyback time.

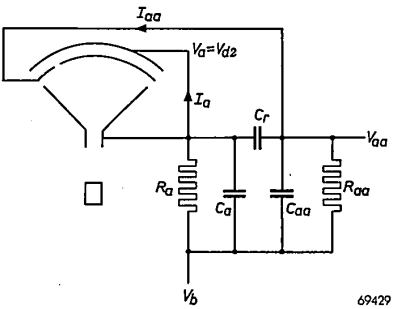



Fig. 14. Computation of the flyback time.

The flyback is therefore split up into a number of phases corresponding to parts of the valve characteristics. These characteristics are idealized in such a manner that for each phase of the flyback a system of linear differential equations can be solved.

Figs 15 and 16 give an example of $V_{d_2}$ (= $V_a$) and $V_{aa}$, respectively, as functions of time. At the left and right the flyback is visible. The separate vertical lines are voltage measures; their length represents a voltage of 100 volts. It can be seen that already in position 9 the voltage $V_{aa}$ of the auxiliary anode falls about 30 volts as a result of that anode being struck by a part of the beam. As soon as the flyback starts the voltages of anode and auxiliary anode both fall about 40 volts. Then $V_{d_2}$ remains almost constant during a short time, while $V_{aa}$ decreases slowly to more than 200 volts below the voltage reached without beam current. After that, $V_{aa}$ remains constant for some time and then quickly rises 140 volts. In the same time $V_{d_2}$ rises about 140 volts, first slowly then rapidly. $V_{d_2}$ has now become rather too high and drops a few volts, while $V_{aa}$ slowly returns asymptotically to the rest position, its voltage changing about 70 volts.

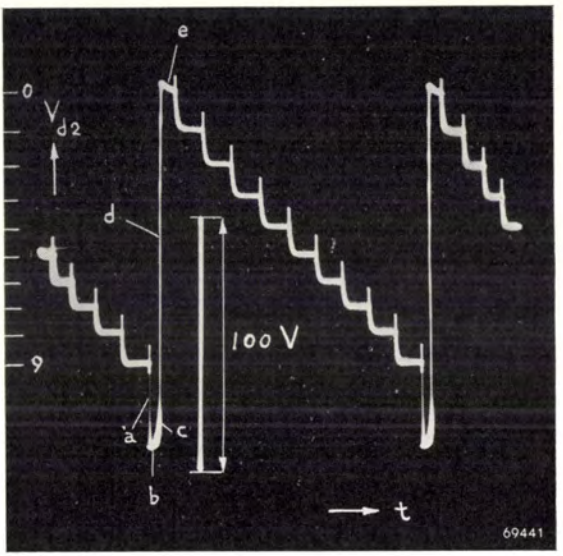



Fig. 15. Oscillogram of $V_{d2}$ ($= V_a$) as a function of time.

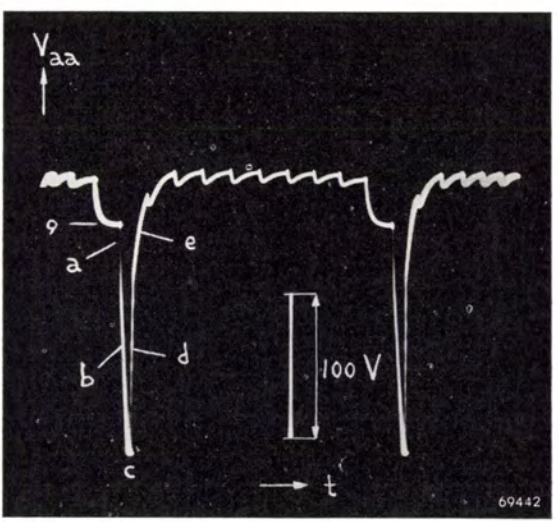



Fig. 16. Oscillogram of $V_{aa}$ as a function of time.

(b) *The valve characteristics and their idealization*

From the investigation into the influence of the valve characteristics upon the duration of the flyback it appeared that two ranges of the characteristics are very important: the time of flyback is mainly determined by the periods of time the beam stays in the two outermost positions. Special attention has therefore been paid to these parts, and by means of linear functions we have tried to find a fairly exact approximation of the real characteristics at least in these ranges. The behaviour of the valve when the beam is in the transitory range between impinging on the counting anode and on the auxiliary anode is less important for computing the flyback time. This also holds good for passing the positions 9, 8, ..., 0 during the flyback. For these ranges a less exact approximation of the characteristics suffices, and thus has been chosen in order to avoid mathematical difficulties.

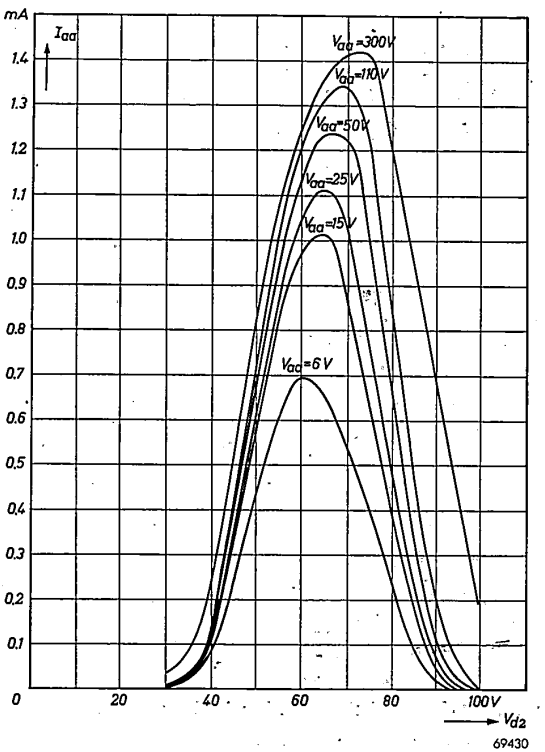



Fig. 17. The valve characteristic giving the auxiliary-anode current $I_{aa}$ as a function of the deflecting voltage $V_{d2}$ (= $V_a$) with $V_{aa}$ as a parameter and the following voltages on the other electrodes (for the indications see fig. 4): $V_s$ = 250 V via 47 kΩ, $V_{tc}$ = 300 V, $V_{d1}$ = 150 V, $V_{a0}$ = 300 V, $V_{p1+p2}$ = − 3 V.

Fig. 17 gives an example of the real characteristics $I_{aa} = f(V_{d_2})$. Idealizing, these can be represented by the lines of fig. 18. In both cases $V_{aa}$ has been taken as a parameter. From the computation it appeared that without much objection the characteristic of fig. 19 can also be applied. Here the slope $\mathrm{d}I_{aa}/\mathrm{d}V_{d_2}$ is infinite on the left side too. During the time the beam impinges on this side of the auxiliary anode $V_{d_2}$ actually remains constant, while the auxiliary-anode voltage decreases considerably. The characteristic on the left side of fig. 19 results in a first-order differential equation for $V_{aa}$, the solution of which at the initial values given in practice differs little from that of the second-order differential equation that follows from the characteristics of fig. 18. In the following computation, therefore, we give the more surveyable results of the characteristic of fig. 19.

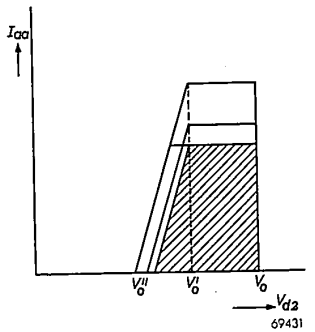



Fig. 18. Idealized characteristic of fig. 17, with infinite steepness on the right side.

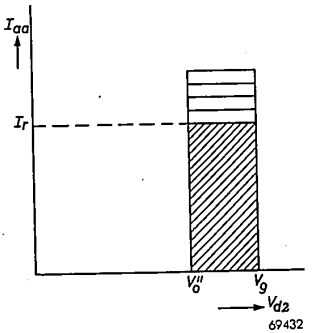



Fig. 19. Idealized characteristic of fig. 17 with infinite steepness on both sides.

Fig. 20 gives the real characteristic $I_{aa} = f(V_{aa})$ with $V_{d_2}(= V_a)$ as a parameter. The corresponding idealized characteristics are given in fig. 21. It is assumed that, for $V_{aa} = V_r$, $\mathrm{d}I_{aa}/\mathrm{d}V_{aa}$ suddenly becomes infinite.

So $V_r$ is the "knee" voltage of the pentode-like characteristic. The vertical part of the characteristic $I_{aa} = f(V_{aa})$ corresponds to the hatched part of the characteristics in figs 18 and 19. The sides of the latter characteristics correspond to the part $V_{aa} > V_r$ in the characteristics of fig. 21.

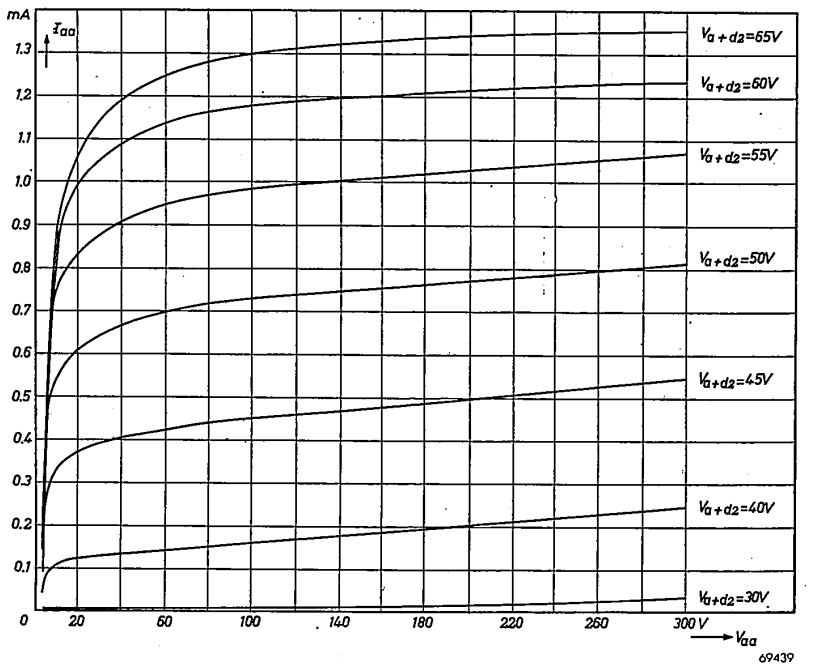



Fig. 20. The auxiliary-anode current $I_{aa}$ as a function of $V_{aa}$ with the deflecting voltage $V_{d2}$ (= $V_a$) as a parameter and the voltages of fig. 17 on the other electrodes.

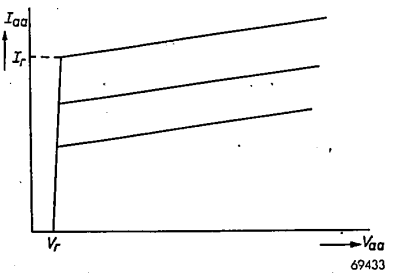



Fig. 21. Idealized characteristic of fig. 20. At the "knee" voltage $V_r$ the steepness is assumed to become infinite.

It is assumed that for the computation of the flyback time the idealized characteristic of fig. 22 can be used instead of the characteristic $I_a = f(V_{d_2})$ of fig. 6. The slowly descending line has replaced the undulatory characteristic. In section 5 it is shown how to compute the breakdown of the flyback occurring when the amplitudes of the waves in this characteristic are too great or the flyback capacitance is too small. In the following sections we will assume $I_a = 0$ for $V_{d_2} < V_9$, and $I_a = (V_b - V_a)/R_a$ for $V_0 \geqq V_{d_2} \geqq V_9$.

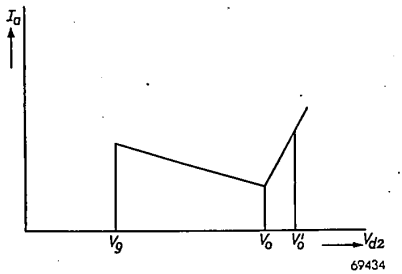



Fig. 22. Idealized characteristic of fig. 6.

In reality there is a continuous transition between the flowing of anode current and auxiliary-anode current when $V_{d_2}$ changes. The properties of this transition appear to have an influence upon whether the flyback will occur or not, but the influence upon the flyback time is very small. The same holds for the wave characteristic $I_a = f(V_{d_2})$ for $V_9 < V_{d_2} < V_0$. It is possible that owing to incorrect dimensions of the circuit elements the flyback is incomplete, and then the beam stops in the position 1, 2 or even higher. If, however, the beam does not stop, the flyback time appeared to be only slightly influenced by the separate oscillations in the characteristic $I_a = f(V_{d_2})$. Between $V_9$ and $V_0$, therefore, this characteristic is considered to be a straight line coinciding with the average of the wave line of the true characteristic. For $V_{d_2} > V_0$ the characteristic is also given by a straight line, with a different slope, and in practice the time the beam stays on the right of $V_0$ appears to contribute considerably to the flyback time. This contribution will also be calculated. Further, it will be investigated what demands the characteristic $I_a = f(V_{d_2})$ has to meet in order to prevent the beam from stopping before the point $V_{d_2} = V_0$ is reached.

(c) *Formation of the differential equations*

To derive the differential equations we start with the circuit of fig. 14 used in practice and the idealized characteristics just described:

$$I_a = f_1 (V_{d_2}) \tag{1}$$

and

$$I_{aa} = f_2 (V_{d_2}, V_{aa}), \tag{2}$$

with $V_{d_2} = V_a$.

The circuit gives the following system of simultaneous differential equations:

$$I_{aa} = \frac{V_b - V_{aa}}{R_{aa}} - C_{aa} \dot{V}_{aa} - C_r (\dot{V}_{aa} - \dot{V}_a), \tag{3}$$

$$I_a = \frac{V_b - V_a}{R_a} - C_a \dot{V}_a - C_r (\dot{V}_a - \dot{V}_{aa}); \tag{4}$$

$C_a$ and $C_{aa}$ are the capacitances to earth of the anode and the auxiliary anode, respectively, $C_r$ is the coupling capacitance of these anodes and the dots refer to derivatives with respect to time.

To compute the flyback time we split up the flyback into a number of phases. Suppose that, owing to a pulse, the beam is in a position such that $V_{d_2} = V_g$. From there $V_{d_2}$ passes through the following phases, indicated in fig. 23.

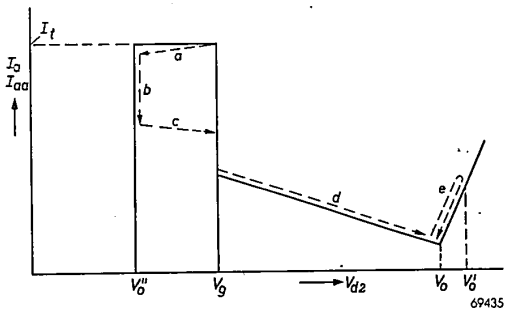



Fig. 23. Illustrating the different phases of the flyback.

*Phase a:* $V_{d_2}$ falls to $V_0''$, the voltage of the auxiliary anode decreases. As a result $I_{aa}$ becomes a little less than $I_t$; $I_a$ is zero.

*Phase b:* $V_{d_2}$ remains constant: $V_{d_2} = V_0''$; the flyback voltage decreases to $V_{aa} = V_r$. Likewise $I_{aa}$ decreases.

*Phase c:* $V_{d_2}$ rises to $V_g$, while $V_{aa}$ remains constant and equal to $V_r$. Meanwhile $I_a$ remains zero.

*Phase d:* $V_{d_2}$ rises to $V_0$, while $V_{aa}$ increases; $I_{aa}$ is now zero; $I_a$ travels down the wave characteristic schematically represented by a descending straight line.

*Phase e:* $V_{d_2}$ rises to $V_0'$ and then asymptotically approaches $V_0$.

(d) *Computation of phase a*

We take $V_b = 0$ volt, so that all potentials are considered with respect to $V_b$; eqs (1), (2), (3) and (4) then become

$$I_a = 0\,,$$

$$I_{aa} = I_t + \varepsilon V_{aa}\,,$$

$$I_{aa} = -\frac{V_{aa}}{R_{aa}} - C_\tau \dot{V}_{aa} + C_r \dot{V}_a\,,$$

$$I_a = -\frac{V_a}{R_a} - C_a \dot{V}_a + C_r \dot{V}_{aa}\,,$$

where $1/\varepsilon$ is the internal resistance, $C_\tau = C_r + C_{aa}$ and $C_a = C_r + C_a$. From this can be derived

$$\dot{V}_a = -\frac{V_a}{\Delta\sigma} - \frac{V_{aa}}{\varrho}\frac{1-\Delta}{\Delta} - \frac{I_m}{C_r}\frac{1-\Delta}{\Delta},$$

where

$$\Delta = \frac{C_a C_\tau - C_r^2}{C_a C_\tau} \ll 1\,,$$

$$\sigma = R_a C_a\,, \qquad 1/\varrho = \frac{1}{C_r}\left(\varepsilon + \frac{1}{R_{aa}}\right).$$

In practice the third term of the right-hand side appears to be the most important.

If, e.g. we take $C_r = 60$ pF, $C_a = 68$ pF, $C_\tau = 64$ pF, $R_a = 0{\cdot}8$ MΩ (thus $\sigma = 55$ μsec), $\varepsilon + 1/R_{aa} = 3.10^{-6}$ $A/V$ (thus $\varrho = 20$ μsec $\Delta = 0{\cdot}17$, $I_m = 1{\cdot}25$ mA, and assuming that the average values of $V_a$ and $V_{aa}$ are $-220$ V and $-40$ V respectively, then the average value of $\dot{V}_a$ becomes

$$\dot{V}_a = -69.10^6 \text{ V/sec.}$$

The time between $V_a = -200$ V and $V_a = -240$ V thus becomes

$$t_a = \frac{40}{69} \text{ μsec} = 0{\cdot}58 \text{ μsec}.$$

The course of $V_{aa}$ is given by

$$\dot{V}_{aa} = \dot{V}_a \frac{C_a}{C_r} + \frac{V_a}{R_a C_r} = 1{\cdot}13\,\dot{V}_a + \frac{V_a}{48.10^{-6}}.$$

Thus in 0·58 μsec $V_{aa}$ drops 48 V.

(e) *Computation of phase b*

Eq. (4) can also be written as

$$\dot{V}_{aa} = \frac{V_a}{R_a C_r} + \frac{C_a}{C_r}\dot{V}_a \approx \frac{V_a}{R_a C_r} + \dot{V}_a,$$

$I_a$ being zero and $C_a \approx C_r$. In this phase $V_a (= V_{d_2})$ is practically constant ($V_0''$): it changes not more than 5 V, as appears from fig. 15. In the same time, however, $V_{aa}$ falls some 125 V (fig. 16). From this it may be concluded that $\dot{V}_a \ll \dot{V}_{aa}$, so that

$$\dot{V}_{aa} \approx \frac{V_a}{R_a C_r} = \frac{V_0''}{R_a C_r}.$$

The duration of this phase becomes

$$t_b = \frac{V_r - V_{aa9}}{\dot{V}_{aa}} = R_a C_r \frac{V_r - V_{aa9}}{V_0''}.$$

Hence if $V_{aa}$ falls from $V_{aa9} = -75$ V to $V_r = -200$ V the time needed with $V_0'' = -250$ V, $R_a C_r = 48$ μsec becomes

$$t_b = 24 \text{ μsec}.$$

(f) *Computation of phase c*

During this phase $I_a = 0$ and $V_{aa} = V_r$. The latter appears from the fact that eqs (3) and (4) can be combined into

$$-\Delta C_\tau \dot{V}_{aa} = I_{aa} + \frac{V_r}{R_{aa}} + \frac{C_r V_a}{\sigma}.$$

With the values of $V_r$ and $V_a$ occurring in practice $\dot{V}_{aa}$ can be positive only if $I_{aa} < I_r$. This, however, is not possible, since $I_{aa} > I_r$ for $V_{aa} > V_r$ (see fig. 21). For $\dot{V}_{aa} = 0$, $I_{aa}$ remains smaller than $I_r$, and this is the only possible solution of this differential equation.

From eq. (4) we get

$$I_a = -\frac{V_a}{R_a} - C_a \dot{V}_a = 0, \qquad V_a = V_0'' e^{-t/\sigma}.$$

Therefore the duration of this phase is

$$t_c = \sigma \ln \frac{V_0''}{V_9}.$$

Taking $\sigma = 55$ μsec and $V_0''/V_9 = 1{\cdot}2$, we get

$$t_c = 10 \text{ μsec}.$$

(g) *Computation of phase d*

Here $I_{aa} = 0$ and $I_a = -V_a/R_a$. From (3) and (4) it then follows that

$$\dot{V}_{aa} = -\frac{V_{aa}}{\Delta R_{aa} C_\tau},$$

so that

$$V_{aa} = V_r \exp(-t/\Delta R_{aa} C_\tau)$$

and, since $\dot{V}_a = C_r/C_a \dot{V}_{aa}$,

$$V_a = V_9 - V_r \frac{C_r}{C_a} [1 - \exp(-t/\Delta R_{aa} C_\tau)].$$

The duration $t_d$ of this phase follows from

$$V_0 = V_9 - V_r \frac{C_r}{C_a} [1 - \exp(-t_d/\Delta R_{aa} C_\tau)];$$

$$t_d = \Delta R_{aa} C_\tau \ln \frac{C_r V_r}{C_a (V_9 - V_0)}.$$

Taking $\Delta = 0{\cdot}17$, $R_{aa} C_\tau = 40$ μsec, $C_r/C_a = 0{\cdot}88$, $V_r = -200$ V, $V_9 - V_0 = -125$ V, then

$$t_d = 2{\cdot}3 \text{ μsec}.$$

(h) *The value of the coupling capacitance*

It is useful, after this computation of the phase $d$, to calculate also the value of $C_r$, this value being especially determined by the fact that if $C_r$ is too small there is no complete flyback, the beam stopping in a position higher than 0. For this computation it is assumed that the oscillations in the characteristic have a negligible influence upon the operation, except for the crest in the $I_a$-$V_a$ characteristic where the beam stops. This means that where the preceding oscillations occur the positive differences between the load line and the wave line are neutralized by the negative ones. To compute $C_r$ it is assumed that with a certain $V_{d_2}$ equal to $V_{ds}$, at the time $t = t_s$, this difference sudenly reaches the value $I_{as}$, and that in this situation $\dot{V}_a$ becomes zero. This means that with $V_{d_2} < V_{ds}$ eqs (3) and (4) are applied, but that with $t = t_s$ we have

$$V_{aa} - V_r = -\frac{C_a}{C_r} (V_9 - V_{ds}).$$

Eq. (3) then becomes

$$\dot{V}_{aa} = -\frac{V_r - \frac{C_a}{C_r} (V_9 - V_{ds})}{R_{aa} C_\tau},$$

and eq. (4)

$$I_a = -\frac{V_{ds}}{R_a} - \frac{C_r V_r - C_a (V_9 - V_{ds})}{R_{aa} C_\tau}.$$

This is a linear function of $V_{ds}$, so that a straight line can be drawn at a distance

$$\frac{V_{ds}}{R_a} + I_{as} = -\frac{C_r}{C_\tau}\frac{V_r}{R_{aa}} + \frac{C_a}{C_\tau}\frac{V_9}{R_{aa}} - \frac{C_a}{C_\tau}\frac{V_{ds}}{R_{aa}}.$$

above the load line $I_a = -V_{d2}/R_a$. When the crests of the $I_a$-$V_a$ characteristic rise above this line the beam will stop. So the crest current $I_m$ has to be smaller than $(V_{ds}/R_a) + I_{as}$.

The last relation for $I_{as}$ shows what is required for the minimum of $C_r$ when $I_m$ is given, viz.

$$\frac{C_r}{C_a} > -\frac{I_m R_{aa}\frac{C_{aa}}{C_a} + V_{ds} - V_9}{I_m R_{aa} + V_r + V_{ds} - V_9}.$$

In fig. 24 the minimum value of $C_r/C_a$ has been plotted as a function of the quantity $I_m R_{aa}/DV_a$, where $C_{aa}/C_a$ and $V_r/DV_a$ have been taken as parameters ($DV_a = V_{ds} - V_9$). It is seen that it is most important to make $-V_r$ as large as possible, in order to get a small $C_r$ and thus a small flyback time when $I_m$ is given.

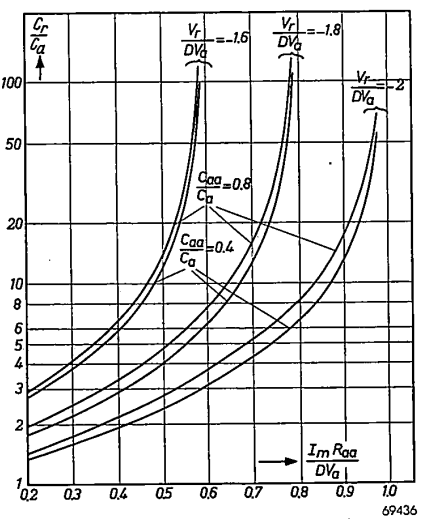



Fig. 24. Minimum value of $C_r/C_a$ needed for a complete flyback

There is another possibility permitting of a small value of $C_r$. The focusing plates $p_1$ are coupled, via a small capacitance, to the screen $s$ (fig. 4). This screen is connected to a positive voltage via a resistor which gives about 50 volts potential drop. During the phases $a$, $b$ and $c$ the screen current is zero and the screen potential will rise. During phase $d$ the screen current is about 1 mA; the screen potential will be lowered and, via the coupling capacitor and the focusing plates, will cut off the beam. $C_r$ may have a much lower value, since there are no crests in the $I_a$-$V_a$ characteristic to stop the beam.

(i) *Comptutation of phase e*

Here we choose another zero for $V_{d_2}$, namely $V_{d_2} = V_0$; let $\varphi_a = V_{d_2} - V_0$. During this phase we have

$$I_{aa} = 0, \qquad I_a = S\varphi_a - \frac{\varphi_a + V_0}{R_a}.$$

$S$ is the difference between $\mathrm{d}I_a/\mathrm{d}V_{d_2}$ and the slope of the load line, for $V_{d_2} > V_0$. Eqs (3) and (4) now become

$$0 = -\frac{V_{aa}}{R_{aa}} - \dot{V}_{aa}\, C_\tau + \dot{\varphi}_a\, C_r\,,$$

$$0 = S\varphi_a - \dot{\varphi}_a\, C_a + \dot{V}_{aa}\, C_r\,.$$

Eliminating $V_{aa}$ and $\dot{V}_{aa}$ we get

$$\ddot{\varphi}_a + \dot{\varphi}_a \frac{\alpha + \tau}{\Delta\alpha\tau} + \varphi_a \frac{1}{\Delta\alpha\tau} = 0\,,$$

where $\alpha = C_a/S$, $\tau = R_{aa}C_\tau$ and $\Delta = [C_aC_\tau - C_r^2]/C_aC_\tau$. Solving this differential equation, we find for $\Delta \ll 1$ and $\alpha \ll \tau$, and with the boundary conditions $V_{aa} = V_{aa0}$ and $\varphi_a = 0$, at $t = 0$,

$$\varphi_a = V_{aa0}\frac{C_r}{C_a}\frac{\alpha(\alpha+\tau)}{(\alpha+\tau)^2 - \Delta\alpha\tau}\left[\exp(-t/t_1) - \exp(-t/t_2)\right],$$

with $t_1 = \Delta\alpha\tau/(\alpha+\tau)$ and $t_2 = \alpha + \tau$.

The term $\exp(-t/t_1)$ will soon become zero, since $t_2 \gg t_1$. Assuming that at $t = t_e$, $\varphi_a$ has become equal to $-V_{aa0}/n$, we find:

$$t_e = (\alpha+\tau)\ln\frac{nC_r\alpha(\alpha+\tau)}{C_a\{(\alpha+\tau)^2 - \Delta\alpha\tau\}}.$$

For $\Delta \ll 1$, $\alpha \ll \tau$ and $C_r \approx C_a$ we may take for this the following approximation:

$$t_e = R_{aa}C_\tau \ln(n/R_{aa}S)\,.$$

If $n$ is not given but the difference $I_{ae}$ between the load line and the real $I_a$ at the time $t = t_e$ is given, then we find

$$t_e = R_{aa}C_\tau \ln \frac{I_{as}}{I_{ae}}.$$

$I_{as}$ is given by the choice of $C_r$; $C_r$ must have such a value that the $I_{as}$ defined in section 5(h) lies far enough above the $I_a$ actually present when the last crest in $I_a$ before $V_0$ is passed. At the time $t = t_e$, however, the anode current must be so small that the valve can count again, and this means that the part of the anode current above the load line must decrease to about one-third of its original value during this phase. From this it follows that

$$t_e \approx R_{aa}C_\tau \ln 3 \approx 1{\cdot}1\, R_{aa}C_\tau .$$

Taking $R_{aa}C_\tau = 50$ μsec, then

$$t_e = 55 \text{ μsec}.$$

The value of $t_e$ appears to be highly dependent upon the demands that are made as to the reliability of the circuit. For instance, $I_{as}$ can be chosen so low that, in one valve, the beam just does not stop, and $I_{ae}$ so high that with a large pulse the valve just starts counting again. There will then, however, be a great chance that an interchange of the counter valves, a small variation in the value of the pulse or a change in the supply voltage will result in the flyback stopping in position 1, thus two positions being counted with one pulse. It is necessary, therefore, to make $\ln (I_{as}/I_{ae})$ at least equal to 1.

(j) *The total flyback time*

From the foregoing it appears that the total flyback time $\tau_r$ is determined mainly by the phases $b$, $c$, $d$ and $e$. We found

$$t_b = R_a C_r \frac{V_r - V_{aa_0}}{V_0''},$$

$$t_c = R_a(C_r + C_a)\ln \frac{V_0''}{V_9},$$

$$t_d = \Delta R_{aa}C_\tau \ln \frac{C_r V_r}{C_a(V_9 - V_0)},$$

$$t_e = 1{\cdot}1\, R_{aa}(C_r + C_{aa}).$$

Assuming for the capacitances, the resistances and the voltages the values used above, we arrive at

$$\tau_r = t_b + t_c + t_d + t_e \approx 100 \text{ μsec}.$$

This is still a rather high amount, which may be understood from the fact that virtually the flyback is a univibrator operating with a very small current whereas it should give off a very large voltage. With such a small beam current time is needed to collect enough charge to produce the pulse for returning the beam to the zero position, and then a similar interval of time is needed again to carry off the surplus of charge till the rest position is reached.

## 6. The counting rate

In evaluating the maximum counting rate we distinguish three possibilities:

(a) the pulses appear at random but are separated at least by a certain dead time $\tau_d$ (gas counters, etc.);
(b) the pulses appear with equal invariable intervals (frequency dividing);
(c) the pulses appear with equal intervals the repetition rate of which is not known (frequency determination).

(a) If the pulses appear at random the dead time $\tau_d$ may exceed the flyback dead time $\tau_r$ and the normal counting dead time $\tau_t$ of the decade counter valve. $\tau_d$ may have a value somewhere between $\tau_r$ and $\tau_t$, and it may be smaller than $\tau_r$ and $\tau_t$.

If the average number of pulse-producing particles per second is $N_0 (N_0 \gg 1)$ and $N_c$ is the number of pulses detected per second by the first decade counter valve, then these $N_c$ pulses are produced in the sensitive time only, this time being the total time (one second) less the sum of the dead times. Thus, for

$$\tau_d > \tau_r \quad \text{and} \quad \tau_d > \tau_t ,$$

$$N_0 = \frac{N_c}{1 - N_c \tau_d} .$$

It is obvious that in this case the dead time of the gas counter is the only limiting factor. Consequently for dead times exceeding $10^{-4}$ sec the decade counter valve with the simple flyback circuit described in section 5 does not influence the counting rate. But with

$$\tau_r > \tau_d > \tau_t ,$$

$$N_0 = \frac{N_c}{1 - 0{\cdot}9\, N_c \tau_d - 0{\cdot}1\, N_c \tau_r} .$$

In this case the counting rate is limited by an apparent dead time

$$\tau_d' = 0{\cdot}9 \tau_d + 0{\cdot}1\, \tau_r .$$

Similarly, for $\tau_r < \tau_d < \tau_t$ the apparent dead time is

$$\tau_d' = 0{\cdot}9\ \tau_t + 0{\cdot}1\ \tau_d$$

and for $\tau_d < \tau_r$ and $\tau_d < \tau_t$

$$\tau_d' = 0{\cdot}9\ \tau_t + 0{\cdot}1\ \tau_r\ .$$

Assuming $\tau_t = 4$ μsec and $\tau_r = 100$ μsec, the apparent dead time as a function of the dead time of the gas-discharge counter valve is given in table I:

TABLE I

| $\tau_d$ | $\tau_d'$ |
|---|---|
| ≧100 | $\tau_d$ |
| 90 | 91 |
| 50 | 55 |
| 10 | 19 |
| ≦ 4 | 13·6 |

(b) If the valve is used for frequency dividing with a dividend lower than ten, a circuit may be used which counts to a lower scale than the decimal one (see section 7). Generally the flyback time will then be shorter than 100 μsec. Moreover it is possible to make the interval between two pulses a small part of the flyback time, then counting to a smaller number than the dividend. One or more pulses then appear during the flyback time. If, e.g., the dividend is 7 and the flyback time for counting to 5 is 60 μsec, two pulses will appear during the flyback time. The minimum time between two pulses is then 20 μsec and the counting rate should be less than 50 000 per sec.

(c) If the pulses have equal intervals but the number of pulses per second is not known, then the interval between two pulses should always be greater than the flyback time. The same applies for the counting of pulses with unequal intervals where no single pulse may be omitted. For these applications a very short flyback time would be highly desirable. In section 8 some circuits will be given which shorten the flyback time considerably.

## 7. Counting in a lower-than-ten system

For some applications a scale lower than ten is desirable. For this purpose the circuit of fig. 25 may be used. It incorporates a diode, the cathode of which is kept at a voltage between 100 and 230 V corresponding to a stable position of the beam in the counter valve. It is clear that the anode of the diode which is connected to the anode and deflection plate of the decade counter valve cannot have a voltage higher than that of the cathode

of the diode without immediately falling to the cathode voltage. Thus at the flyback the anode voltage will not rise to position 0 but to 1,2, etc., corresponding to a scale of 9, 8, etc.

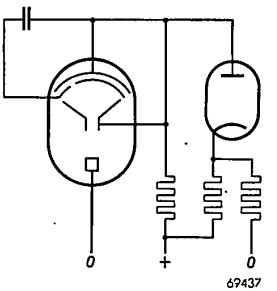



Fig. 25. Counting in a lower-than-ten system.

## 8. The use of a trigger for the flyback

The flyback time can be considcrably shortened by employing a separate univibrator instead of the built-in one. It is possible to build a univibrator with a resolving power less than 0·1 μsec [5]), but such a univibrator cannot furnish the voltage drop of about 150 V necessary for the counter-valve beam to jump from position 9 to position 0. If there is a stray capacitance of 20 pF then a valve yielding 30 mA needs 0·1 μsec to produce a charge up to 150 V. Moreover, unless a secondary-emission valve is used, a valve gives off a negative charge, and when a common valve is used the positive charge needed for the flyback of the counter valve has to be supplied via a resistance.

Fig. 26 gives an example of a circuit with the aid of which a rather high resolving power was obtained. The anode and the deflection plate of the

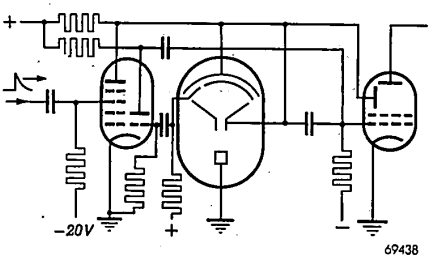



Fig. 26. Circuit for a large resolving power. A separate integrating valve counts the pulses. The flyback consists of a univibrator incorporating a secondary-emission valve.

counter valve are connected to the hexode anode of a triode-hexode and to the dynode of a secondary-emission valve. Every time a positive voltage pulse is given to the third grid of the hexode the charging of the deflection plate by the current pulse is sufficiently negative to cause the beam to jump from one position to the next. It is easy to calculate that with a normal radio valve this happens within $10^{-7}$ sec and obviously only the duration of the voltage pulse limits the dead time $\tau_t$. The flyback is obtained in a similar manner, but the deflection plate should lose a 10-fold charge within a very short time. At the tenth pulse the electron beam hits the auxiliary anode and charges the control grid of the triode negatively. The resulting positive voltage of the triode anode opens the control grid of the secondary-emission valve, and as the negative internal resistance of the dynode may reach [6]) about 100 Ω the deflecting voltage will increase very rapidly till the anode voltage of this valve is reached. With the circuit of fig. 26 a flyback time of 0·3 μsec has been attained. Fig. 27 shows an oscillogram of the anode voltage in such a circuit.

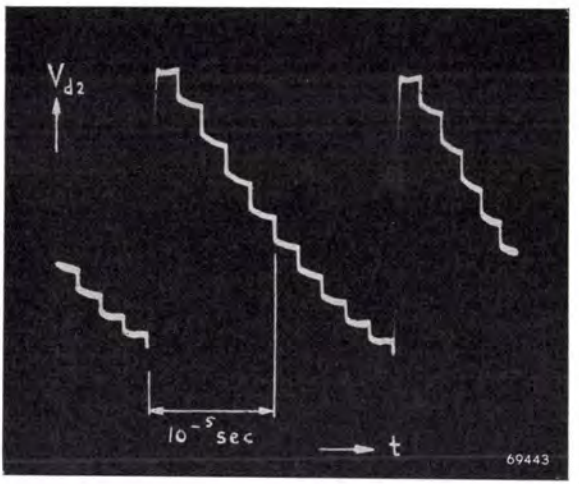



Fig. 27. Oscillogram of the deflection-plate voltage with the circuit of fig. 26. The pulses were 2 μsec apart.

**Acknowledgements**

In the course of this work valuable contributions were made by E. van Barneveld, J. Domburg, Z. van Gelder, H. Groendijk, J. de Hartog and G. Hepp of the Philips Research Laboratories and by K. Rodenhuis,

C. van der Velden and L. Wolters of the Philips Radio-Tube Development Laboratory.

*Eindhoven, July 1951*

## Appendix

### LIST OF SYMBOLS

| | |
|---|---|
| $C$ | see fig. 9 |
| $C_a$ | anode capacitance (parasitic capacitances included) |
| $C_{aa}$ | auxiliary-anode capacitance |
| $C_r$ | coupling capacitance between anode and auxiliary anode (flyback capacitance) |
| $C_a$ | $C_r + C_a$ |
| $C_\tau$ | $C_r + C_{aa}$ |
| $DV_a$ | $V_{ds} - V_9$ |
| $e$ | base of natural logarithms |
| $I_0$ | beam current |
| $I_a$ | anode current |
| $I_{aa}$ | auxiliary-anode current |
| $I_{ae}$ | value of anode current at the end of phase $e$ |
| $I_{as}$ | value of anode current when $\dot{V}_{d_2}$ becomes zero |
| $I_C$ | current through $C$ (fig. 9) |
| $I_m$ | wave amplitude of the characteristic $I_a = f(V_{d_2})$ |
| $I_r$ | value of $I_{aa}$ when $V_{aa} = V_r$ |
| $I_{Raa}$ | current through auxiliary-anode resistance (fig. 7) |
| $I_t$ | maximum value of $I_{aa}$ |
| $n$ | see section 5 (i) |
| $N_0$, $N_c$ | see section 6 |
| $q$ | $2\pi\ (V_{d_2} - V_C)/V_e$ |
| $q_0$, $q_t$ | values of $q$ at the moments $t_0$ and $t$, respectively |
| $r$ | $CV_e/2\pi I_m$ |
| $R_a$ | anode resistance |
| $R_{aa}$ | auxiliary-anode resistance |
| $S$ | $(\mathrm{d}I_a/\mathrm{d}V_{d_2}) - 1/R_a$ for $V_{d_2} > V_0$ |
| $t$ | time |
| $t_0$ | fixed value of $t$ |
| $t_1$ | $\Delta a\tau/(a + \tau)$ |
| $t_2$ | $a + \tau$ |
| $t_a \ldots t_e$ | time intervals of the flyback phases |
| $t_s$ | value of $t$ where $\dot{V}_{d_2}$ becomes zero |
| $u$ | $-(2\pi/V_e)\ (\mathrm{d}V_C/\mathrm{d}t)$ |

IPR2025-00217 EXHIBIT 2010 Page 114 of 485

| | |
|---|---|
| $V_0$, $V_0'$, $V_0''$, | |
| $V_0$ | see figs 18 to 23 |
| $V_a$ | anode voltage |
| $V_{a0}$ | gun-anode voltage |
| $V_{aa}$ | auxiliary-anode voltage |
| $V_{aa0}$ | value of $V_{aa}$ in position 0 |
| $V_{aa9}$ | value of $V_{aa}$ in position 9 |
| $V_b$ | battery voltage |
| $V_{d1}$ | voltage of deflection plate not connected to anode |
| $V_{d2}$ | voltage of deflection plate connected to anode ($= V_a$) |
| $V_{ds}$ | value of $V_{d2}$ where $\dot{V}_{d2}$ becomes zero |
| $V_e$ | wavelength of characteristic $I_a = f(V_{d2})$ |
| $V_G$ | pulse voltage ($= V_{d1}$) |
| $V_{G0}$, $V_{Gt}$ | values of $V_G$, at the moments $t_0$ and $t$, respectively |
| $V_i$ | crest voltage of pulse |
| $V_r$ | see fig. 21 |
| $V_s$ | screen voltage |
| $\alpha$ | $C_a/S$ |
| $\Delta$ | $(C_aC_\tau - C_r^2)/C_aC_\tau$ |
| $\varepsilon$ | reciprocal of the internal resistance of the counter valve |
| $\pi$ | 3·14 ... |
| $p$ | $C_r/(\varepsilon + 1/R_{aa})$ |
| $\sigma$ | $R_aC_a$ |
| $\tau$ | $R_{aa}C_\tau$ |
| $\tau_d$,$\tau_d'$,$\tau_t$ | see section 6 |
| $\tau_r$ | total flyback time |
| $\varphi_a$ | $V_{d2} - V_0$ |

## REFERENCES


1) J. L. H. Jonker, Valves with a ribbon-shaped electron beam, Philips Res. Rep. **5**, 6-22, 1950.
2) H. Alfvén and H. Romanus, Valve with trochoidal electronic motion, Nature **160**, 614-615, 1947.
3) D. L. Hollway, An electron-beam decimal-counter tube, Nature **165**, 856-857, 1950.
4) J. R. Pierce, Rectilinear electron flow in beams, J. appl. Phys. **11**, 548-554, 1940.
5) J. R. Tillman, Transition of an Eccles-Jordan circuit, Wireless Engr **28**, 101-110, 1951.
6) A. J. W. M. van Overbeek, Voltage-controlled secondary-emission multipliers, their construction and application, Wireless Engr **28**, 114-125, 1951.

# INSTABILITY OF THIN-WALLED CYLINDERS SUBJECTED TO INTERNAL PRESSURE *)

by J. A. HARINGX

539.371

**Summary**

It is shown that under certain conditions a thin-walled cylinder may buckle when subjected to internal pressure. The critical value of this pressure can easily be calculated by means of Euler's well-known formula. Most of the formulae for pressure-loaded cylinders given in current text-books, however, fail to predict this behaviour and should therefore be applied with caution in the case of great lengths and/or high pressures.

**Résumé**

Dans certaines conditions, un cylindre à minces parois soumis à des pressions intérieures peut être sujet à „flambage". La valeur critique de cette pression peut se calculer facilement à l'aide de la formule bien connue de Euler. Toutefois, la plupart des formules, indiquées dans les manuels usuels pour le calcul des cylindres soumis à pression, ne soufflent mot de ce flambage; aussi, dans le cas de cylindres très longs et de pressions éleveés ne les utilisera-t-on qu'avec certaines précautions.

**Zusammenfassung**

Es wird gezeigt, daß unter bestimmten Bedingungen ein dünnwandiger Zylinder knicken kann, wenn er innerem Druck ausgesetzt wird. Der kritische Wert dieses Druckes läßt sich auf Grund der bekannten Eulerschen Formel leicht errechnen. Die meisten der in den einschlägigen Lehrbüchern enthaltenen Formeln für druckunterworfene Zylinder berücksichtigen nicht die Möglichkeit dieses Verhaltens und sind deshalb in Fällen großer Länge und bzw. oder hohen Druckes mit Vorsicht anzuwenden.

## 1. Introduction

The problem of elastic stability of thin-walled cylinders subjected to radial and axial forces has occupied many investigators since the middle of the last century [1]). A number of well-known scientists have in turn made such contributions as would lead one to consider this problem — at least in the idealized case of simple boundary conditions and perfectly straight cylinders — as being solved. Yet in current text-books, such as those of Biezeno-Grammel [2]), Flügge [3]), Love [4]), Pflüger [5]), Prescott [6]) and Timoshenko [7]), formulae are given which are far from identical. It is true that always a similar relation is found from which the critical combinations of the radial and axial pressures as a function of the dimensions of the cylinder can be derived, but considered in detail the coefficients occurring in this relation show mutual differences.

These differences are of minor importance in the case where the radial

*) Originally published in Dutch in De Ingenieur **63**, O39-O41, 1951.

pressure is exerted on the cylinder externally. When, however, the formulae are applied to the special case of a cylinder subjected to internal pressure, the mutual differences appear to be really of very essential importance. In this respect the formulae may be divided into two groups, viz.

(a) the group according to which the thin-walled cylinder unloaded in the axial direction would have to buckle under a certain internal pressure (Flügge, Pflüger, Prescott);

(b) the group according to which an internal pressure would never lead to instability (Biezeno-Grammel, Love, Timoshenko).

As far as the author knows, the possibility of a thin-walled cylinder buckling under an internal pressure has never before been considered in literature [8]). Yet this case of instability is very obvious, as will be shown in section 2 in an elementary way. This means that the formulae of the last-mentioned group are inadequate to give a useful answer under all circumstances. With cylinders of great length and/or high pressure these formulae have to be applied with caution.

The more general formulae for thin-walled cylinders subjected simultaneously to radial and to axial pressures, which formulae are to be preferred on account of the foregoing remarks, will be submitted to a closer examination in section 3. Further, for one set of deviating formulae the cause of the essential differences in the result will be given in the appendix.

Though the discussion of the problem of instability of bellows subjected to internal pressure has been deferred for another paper [9]), it is to be mentioned that it was the investigation into the cause of the — more or less unexpected — buckling of this elastic element which drew the author's attention to the similar behaviour of thin-walled cylinders.

## 2. The thin-walled cylinder subjected to internal radial pressure

Imagine a cylinder closed at the ends and filled with a liquid under pressure. Unless special precautions are taken the wall will then have to bear not only the radial pressure but also an axial load. Since in the case of loading to be investigated here it is desired to confine ourselves to the radial pressure, at least at one end of the cylinder a frictionless piston will have to be applied. The liquid pressure acting in the axial direction is then taken up by the piston rod and not by the cylinder wall. If a cylinder of diameter $2a$, length $l$ and wall thickness $h (h \ll a)$ is subjected to an internal pressure $q'$, the axial force amounts to $P = \pi a^2 q'$ and thus we get the situation as represented diagrammatically in fig. 1.

From the calculation given below it appears that as the pressure is gradually increased the cylinder remains straight until, upon a certain pressure being reached, it suddenly deflects. This happens in such a way that the shape shows a great similarity to that of the straight bar buckling under

an axial load. If, for instance, both ends are hinged the cylinder assumes the shape given in fig. 2, the cross-section remaining circular and the centre line bending according to the curve $y = f(z)$.

If the cylinder filled with liquid is cut at $z$ according to a plane normal to the centre line, at this cross-section the normal force $N \approx P$, the transverse force $D \approx -P\mathrm{d}y/\mathrm{d}z$ and the bending moment $M = Py$ have to be applied in order to maintain the equilibrium. The liquid, which owing to the axial discharge of the cylinder wall transmits the whole of the normal force, naturally cannot contribute towards the transverse force and the bending moment, so that these two load components have to be taken up completely by the cylinder wall.

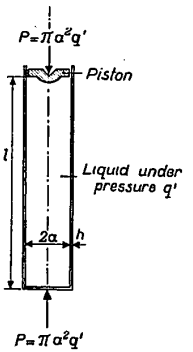



Fig. 1. The thin-walled cylinder subjected to internal pressure, its wall being unloaded in the axial direction.

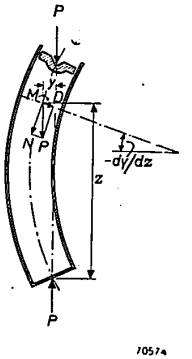



Fig. 2. The cylinder of fig. 1 in the deflected state.

The moment of inertia of the ring-shaped section with respect to one of its diameters amounts to $I = \pi a^3 h$. With Young's modulus of elasticity $E$, the rigidity of the cylinder with respect to bending is equal to $EI$. Thus, if the cylinder buckles the curvature $\mathrm{d}^2y/\mathrm{d}z^2$ of its centre line must, when ignoring the shear effect, satisfy the relation

$$\mathrm{d}^2y/\mathrm{d}z^2 = -M/EI.$$

With $M = Py$ this leads to the differential equation

$$\mathrm{d}^2y/\mathrm{d}z^2 + Py/EI = 0.$$

For the boundary conditions given ($y = 0$ at $z = 0$ and $z = l$) a deflection appears to be possible only if $P$ satisfies the relation

$$P = \pi^2 EI/l^2, \qquad (1)$$

or, on account of the relations $P = \pi a^2 q'$ and $I = \pi a^3 h$, if the internal pressure $q'$ is equal to

$$q' = \pi^2 Eah/l^2 . \tag{2}$$

Since the method of calculation given here agrees entirely with that followed in deriving the condition for instability of the thrust-loaded straight rod, it is not surprising that eq. (1) is identical with Euler's well-known formula. This means, therefore, that instability likewise occurs when the critical force $P$ given in (1) is not transferred into a radial pressure by means of the piston and the liquid, but is directly applied to the wall as an axial load. This will be reverted to in the next section.

### 3. The thin-walled cylinder subjected to radial and axial pressures

From the theory of shells it is known that the thin-walled cylinder subjected to radial and axial pressures may not under all circumstances be regarded as a profiled rod. The wall may be divided both along the circumference and in the longitudinal direction into a number of fields where the wall is bent alternately inward and outward.

If, in the middle of the wall given by the coordinates $a$, $\varphi$ and $z$, we denote the additional components of the displacement in the radial, tangential and axial directions respectively by $u_0$, $v_0$ and $w_0$, the deformation mentioned can be described by the relations (cf. Biezeno and Grammel [2]), p. 603)

$$\left.\begin{aligned} u_0 &= U \cos p\varphi \sin(\lambda z/a) , \\ v_0 &= V \sin p\varphi \sin(\lambda z/a) , \\ w_0 &= W \cos p\varphi \cos(\lambda z/a) , \end{aligned}\right\} \quad (\lambda = \pi an/l) \tag{3}$$

where $U$, $V$ and $W$ are constants, and $p$ and $n$ each represent a positive integer. In the special case where $p = 1$ and $n = 1$ the cross-section of the cylinder remains circular and the centre line is bent according to half a sine. Thus this form of deflection is the same as that discussed in the previous section.

In the more general case where the cylinder is loaded not only with a radial pressure $q = -q'$ but also with the axial compressive force $Q$ per unit circumference, Flügge's calculation [3]), for instance, leads, with Poisson's ratio $\nu$ and the notations

$$q_1 = \frac{(1-\nu^2)qa}{Eh}, \qquad q_2 = \frac{(1-\nu^2)Q}{Eh}, \qquad k = \frac{h^2}{12\,a^2}, \tag{4}$$

to the condition of instability

$$c_1 + c_2 k = c_3 q_1 + c_4 q_2 . \tag{5}$$

The coefficients $c_i (i = 1, 2, 3, 4)$ occurring in this equation are determined,

in a somewhat more complete form than that given originally by Flügge, by

$$\left.\begin{aligned} c_1 &= (1-\nu^2)\lambda^4, \\ c_2 &= (\lambda^2+p^2)^4 - 2\{\nu\lambda^6+3\lambda^4p^2+(4-\nu)\lambda^2p^4+p^6\} + \\ &\qquad + (4-3\nu^2)\lambda^4 + 2(2-\nu)\lambda^2p^2+p^4, \\ c_3 &= (\lambda^2+p^2)^2(p^2-1) - \lambda^2p^2+2\nu\lambda^4, \\ c_4 &= (\lambda^2+p^2)^2\lambda^2 + 2(1+\nu)\lambda^4+\lambda^2p^2. \end{aligned}\right\} \quad (6)$$

To apply this result as a check upon the elementary case given in the previous section we need only introduce the substitutions $p = 1$, $n = 1$, $\lambda \ll 1$ and $k \ll 1$. It is then found that to a first approximation instability occurs if

$$-q_1 + 2q_2 \approx (1-\nu^2)\lambda^2, \quad (7)$$

or — owing to $q = -q'$, $\lambda = \pi a/l$ and eq. (4) — if

$$q' + 2Q/a \approx \pi^2 Eah/l^2. \quad (8)$$

For the cylinder wall unloaded in the axial direction $Q = 0$ applies, so that the result in this case is identical with the result (2) deduced in the elementary way. On the other hand when the cylinder wall is axially loaded $\pi a^2(q' + 2Q/a)$ represents the total external force applied axially on the ends. Thus, it appears to be this total force which, substituted in Euler's well-known formula, gives the condition of instability (8), so that with respect to this instability it is immaterial whether the force is exerted directly on the wall or transmitted entirely or partly by means of a piston to the liquid contained in the cylinder.

If the cylinder is closed at its ends, a case which often occurs, the wall has to bear per unit of circumference an axial tensile force $-Q = \frac{1}{2} q'a$. From eq. (8) it follows that this tensile force is just sufficient to cancel the buckling effect of the radial pressure. Probably this will be one of the reasons why the possibility of a cylinder buckling when subjected to internal pressure has not become common experience.

Flügge explains the results of his calculation by means of diagrams which take into account the presence of internal pressures ($q' < 0$). But then he is considering only the effect of these pressures upon the axial compressive force, without mentioning that an internal radial pressure may in itself cause the buckling of a thin-walled cylinder unloaded in the axial direction. It is therefore to be concluded that this possibility also escaped his notice.

A comparison of Flügge's results with, for instance, those of Biezeno and Grammel [2]) shows that the condition of buckling given by the last-mentioned authors can also be written in the form of eq. (5) with the

IPR2025-00217 EXHIBIT 2010 Page 120 of 485

coefficients $c_1$, $c_2$ and $c_4$ according to (6) but with the coefficient $c_3$ equal to

$$c_3 = (\lambda^2 + p^2)^2 (p^2 - 1) + \nu\lambda^4.$$

The term $\lambda^2 p^2$ required for the result (7) has disappeared from the equation for $c_3$ and so the formula of Biezeno and Grammel does not lead to the desired result. The reason for this discrepancy is briefly explained in the appendix.

Whilst Plüger [5]) and Prescott [6]) confirm the result of Flügge's calculation, that the thin-walled cylinder unloaded in the axial direction will buckle under a certain internal pressure, Love [4]) and Timoshenko [7]) give as a result, in accordance with the Biezeno-Grammel calculation, that in no case will an internal pressure lead to instability. From the elementary calculation given in section 2 it is clear that only the group of authors first mentioned have the right view of the problem.

Here it is to be emphasized that the differences described find expression only in the special case of internal pressure, because only then is the solution $p = 1$ a possible one. In the cases where the radial pressure is exerted on the cylinder externally we always have $p \neq 1$ and the mutual differences are of minor importance.

## 4. Summary

The problem of the thin-walled cylinder subjected to radial and axial pressures is dealt with in many text-books for applied mechanics.

For the idealized case of simple boundary conditions and perfectly straight cylinders it is possible to give a solution which in general is fairly reasonable. There appear to be, however, mutual differences in the formulae given, and, although these differences are in many cases of minor importance, in the case of a cylinder loaded by internal pressure, as dealt with here, they lead to essentially different results.

By means of an elementary calculation where the thin-walled cylinder is regarded as a profiled straight rod, it can be shown that the cylinder unloaded in the axial direction will buckle under a certain internal pressure. This phenomenon is predicted by only a limited number of formulae [e.g. Flügge [3]), Pflüger [5]) and Prescott [6])], which are therefore to be preferred to the other group of formulae [e.g. Biezeno and Grammel [2]), Love [4]) and Timoshenko [7])] according to which a cylinder subjected to internal pressure would never be able to become unstable.

It appears that the internal pressure necessary for the buckling of a cylinder unloaded in the axial direction by means of a piston (fig. 1) can be calculated by putting the force exerted on the piston equal to the axial buckling force calculated from Euler's formula. When investigating the danger of buckling we have to deal with the total axial force, no matter

whether this force is transmitted directly to the cylinder wall as an axial load or, entirely or partly, by means of a piston and liquid as a radial load.

In consequence of the great rigidity of the cylinder with respect to bending, the kind of instability pointed out here will occur only in the case of high pressures and/or great lengths of the cylinder. Furthermore special precautions have to be taken to unload the cylinder wall in the axial direction entirely or partly, for if the cylinder is closed at its ends the axial tensile force in the wall is just sufficient to cancel the buckling effect resulting from the radial pressure.

## Appendix

The difference between the results of Flügge [3]) and those of Biezeno and Grammel [2]) occurs because the last-mentioned authors have taken no account of the axial component $-q' \, \partial u/\partial z$ of the hydrostatic pressure when calculating the additional load in the deflected state of the wall. When their second eq. (VII, *30*, 6) with $q = -q'$ is replaced by

$$Z = -\frac{q\partial^2 w_0}{a\partial\varphi^2} + q\frac{\partial u_0}{\partial z},$$

the coefficient of $U$ in their third eq. (VII, *30*, 10) increases with the term $\lambda q_1$. Supplemented in this way their eqs (VII, *30*, 10) conform to the corresponding equations of Flügge.

*Eindhoven, January 1952*

## REFERENCES


1) For a brief historical survey see, e.g., Sanden and Folke, Über Stabilitätsprobleme dünner kreiszylindrischer Schalen, Ing. Archiv **3**, 24, 1932.
2) C. B. Biezeno and R. Grammel, Technische Dynamik, Springer, Berlin (1939), p. 604.
3) W. Flügge, Statik und Dynamik der Schalen, Springer, Berlin (1934), p. 193.
4) A. E. H. Love, Mathematical theory of elasticity, University Press, Cambridge (1927), p. 574.
5) A. Pflüger, Stabilitätsprobleme der Elastostatik, Springer, Berlin (1950), p. 133.
6) J. Prescott, Applied elasticity, Dover Publ., New York, (1946), p. 551, eq. (17.127).
7) S. Timoshenko, Theory of elastic stability, McGraw Hill, New York (1936), p. 450.
8) After the completion of this paper I learned from Dr A. Klinkenberg of the Bataafse Petroleum Maatschappij (The Hague) that the buckling of drill pipes had been a point of discussion at the Spring meeting of the Southwestern District, Division of Production, American Petroleum Institute (March 1951). In commenting upon Klinkenberg's paper Mr H. B. Woods gave a formula for the buckling of drill pipes. According to this formula buckling of a thin-walled cylinder may occur when the axial tensile force per unit of circumference is smaller than half the difference between the internal and external hydrostatic pressures multiplied by the radius of the cylinder. This limit agrees with relation (8) for infinitely long cylinders.


IPR2025-00217 EXHIBIT 2010 Page 122 of 485

R 187 *Philips Res. Rep.* 7, 119-130, 1952

# A SIMPLE ELECTROSTATIC ELECTRON-OPTICAL SYSTEM WITH ONLY ONE VOLTAGE

by P. SCHAGEN, H. BRUINING and J. C. FRANCKEN

537.533.331

**Summary**

In this article 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 appertaining to this system are calculated and the influence of the hole in the anode is estimated. These 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 $n$ equals 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.

**Résumé**

Description d'un système électronoptique simple, utilisable dans les tubes de prise de vue et dans les tubes convertisseurs d'images. Ce système est constitué par deux électrodes sphériques concentriques; l'électrode intérieure, l'anode, est percée d'un trou. On calcule les trajectoires des électrons dans ce système et on estime l'influence du trou percé dans la sphère anodique. A une distance $6(n-1)R_a/(n-4)$ derrière le trou dans l'anode on obtient une image renversée de la cathode grossie $3/(n-4)$ fois, $n$ étant le rapport des diamètres des sphères et $R_a$ le rayon de la sphère anodique. Des systèmes pratiques, se rapprochant du modèle décrit, ont permis d'obtenir des images de bonne définition et de bonne forme géométrique.

**Zusammenfassung**

Es wird ein einfaches elektronenoptisches System beschrieben, das in Bildzerlegerröhren und Bildwandlern angewendet werden kann. Es besteht aus zwei konzentrischen kugelförmigen Elektroden, von denen diejenige mit dem kleineren Radius, die Anode, mit einem Loch versehen ist. Die Elektronenbahnen in diesem System werden berechnet und der Einfluss des Anodenloches wird abgeschätzt. Man erhält ein umgekehrtes Bild der Kathode in einem Abstand gleich $6(n-1)R_a/(n-4)$ hinter dem Loch in der Anodenkugel, mit einer Vergrösserung gleich $3/(n-4)$, wobei $n$ das Verhältnis des Kathodenzum Anodenradius und $R_a$ der Anodenradius ist. Mit praktisch ausgeführten Systemen, die dem beschriebenen Modell sehr ähnlich sind, wurden sehr scharfe und verzerrungsfreie Bilder erhalten.

## 1. Introduction

The simplest electron-optical system with two electrodes consists of a flat anode opposite a flat cathode. This construction was first used for image converting in 1934 by G. Holst, J. H. de Boer, M. C. Teves and C. F. Veenemans [1]) of this laboratory (see fig. 1).

The uniform field between the electrodes does not focus the electron pencils emitted from the cathode. Electrons emitted from a point $P$ of the cathode with only a tangential velocity component $v_r$ — perpendicular to the electrostatic field — describe parabolas with their apex at $P$. They strike the anode surface at a distance [2])

$$\varrho = 2d \sqrt{\frac{\varepsilon_r}{e\Phi_a}} \tag{1}$$

from the intersection $Q$ of the axis of these parabolas with the anode surface, where

$d$ = distance between anode and cathode,
$\Phi_a$ = potential difference between anode and cathode (in volts),
$\varepsilon_r$ = $\frac{1}{2}mv_r^2$ = initial (tangential) energy in W. sec.

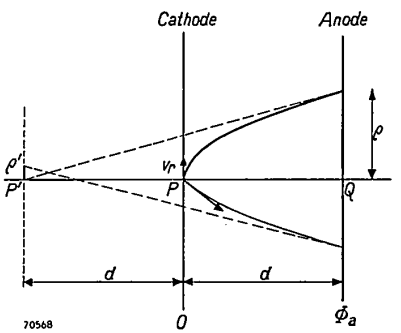



Fig. 1. Diagram of electron paths in an electron-optical system consisting of a flat anode opposite a flat cathode.

The tangents of the parabolas at the anode surface intersect in a point $P'$ behind the cathode at a distance from $P$ equal to $d$. This point $P'$ may be considered to be the virtual image point of $P$, seen from $Q$. Its position is independent of the magnitude of $v_r$ (or $\varepsilon_r$).

If, however, the electrons emitted from $P$ also possess a velocity component $v_z$ in the direction of the field lines, the virtual image point is no longer $P'$. It can be deduced that the tangents of all electron trajectories, taken at the anode surface, intersect the plane through $P'$ parallel to the electrodes within a circle of radius

$$\varrho' = \frac{\varepsilon_0}{e\Phi_c'}, \tag{2}$$

where $\varepsilon_0 = \varepsilon_r + \varepsilon_z$ = maximum energy of emission, and $\Phi'_c$ = potential gradient of the electrostatic field.

This effect is usually called "chromatic aberration", in analogy with light-optical terminology.

The disadvantages of this system for use in image converters are, therefore:

(a) a poor definition, given by equation (1);

(b) the field gradient at the cathode surface, $\Phi'_c = \Phi_c/d$, may easily cause field emission in cases where high voltages and small values of $d$ are employed in order to get a small value of $\varrho'$;

(c) the magnification is always equal to unity, which is a serious limitation of the system.

An effective method of reducing the field gradient at the cathode is provided by shaping the cathode and anode surfaces as concentric spheres, the anode sphere having the smaller radius (see figure 2). In this way the field gradient is reduced to

$$\Phi'_c = \frac{\Phi_a}{nd}, \tag{3}$$

where $n$ is the ratio of radii of cathode and anode spheres, $d$ the distance between cathode and anode (in this case the difference of the radii) and $\Phi_a$ the potential of the anode if the cathode potential is taken to be zero.

The unsharpness of the demagnified "image" on the anode sphere is

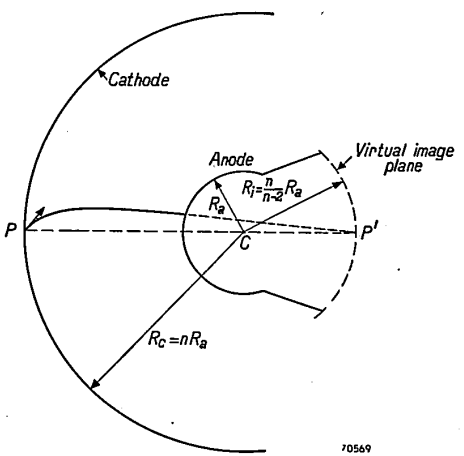



Fig. 2. Diagram of an electron path in an electron-optical system consisting of concentric spheres, the anode sphere having the smaller radius.

given by a formula similar to eq. (1) (see eq. (14)), but it is relatively more serious because of the demagnification of the image. In this case a virtual image lies beyond the anode sphere. In section 2 of this paper it will be shown that it is situated on a sphere concentric with the others and of radius

$$R_i = \frac{n}{n-2} R_a, \tag{4}$$

where $R_a$ is the radius of the anode sphere.

This virtual image cannot be made real, even if it were possible to realize an anode in the shape of an ideal grid penetrable for electrons without loss of energy. As the space beyond the anode is field free, there would always be a bend in the potential resulting in a negative-lens action.

By providing a small hole in the anode sphere a real image can be obtained, as is shown in section 3 of this paper. Due to the negative-lens action of the hole this image is not situated on the sphere given by equation (4) but at a greater distance beyond the anode (see figure 3).

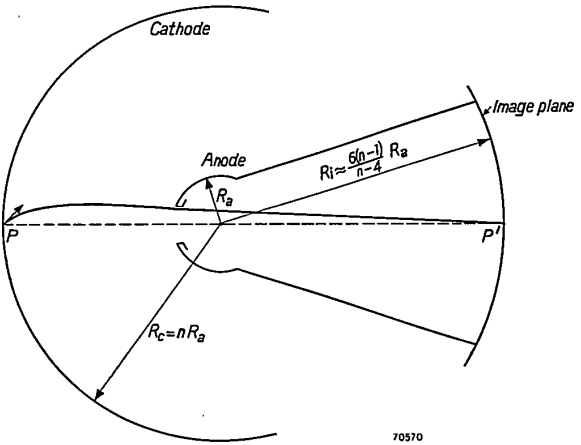



Fig. 3. Schematic illustration of the influence on the electron paths of a hole in the anode sphere.

## 2. Theory of virtual image formation in a system consisting of two concentric spheres

### (a) *The field equation*

The potential field is considered between two concentric spheres with centre $C$ and radii of cathode and anode sphere equal to $R_c$ and $R_a$ respectively.

Let the potential on the cathode sphere be $\Phi_{R_c} = 0$, and on the anode sphere $\Phi_{R_a} = \Phi_a$.

At a distance $r$ from $C$, with $R_a < r < R_c$, the potential is in this case

$$\Phi_r = \frac{\Phi_a}{n-1}\left(\frac{R_c}{r} - 1\right) = \Phi_0\left(\frac{R_c}{r} - 1\right), \tag{5}$$

where

$$n = \frac{R_c}{R_a} \quad \text{and} \quad \Phi_0 = \frac{\Phi_a}{n-1}. \tag{6}$$

(b) *The ray equation*

The centre $C$ is taken as the origin of a system of spherical coordinates $r$, $\varphi$ and $\vartheta$. Without loss of generality, the calculations may be restricted to the case $\vartheta$ = constant, and this plane is represented in figure 4.

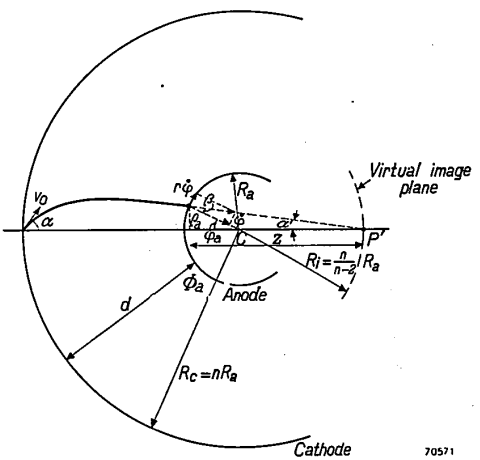



Fig. 4. Diagram for calculating the ray equation.

An electron is now considered, leaving the surface of the cathode sphere with a velocity $v_0$ at an angle $\alpha$ with the normal, and with the direction of $v_0$ in the plane of drawing. This electron moves in the central field of force between the spheres, and the path is accordingly determined by Kepler's law of equal areas

$$r^2\,\dot{\varphi} = C\,, \tag{7}$$

and the energy equation

$$\tfrac{1}{2}\,mv^2(r) = \tfrac{1}{2}\,mv_0^2 + e\Phi_r\,. \tag{8}$$

IPR2025-00217 EXHIBIT 2010 Page 127 of 485

The constant $C$ in equation (7) is found from the boundary condition $R_c\dot{\varphi}_0 = v_0 \sin \alpha$, whilst the velocity $v$ in eq. (8) is composed of its two components according to

$$v^2 = \dot{r}^2 + (r\dot{\varphi})^2 .$$

Eliminating $t$ from equation (7) and (8) yields the ray equation

$$\mathrm{d}\varphi = \frac{R_c \sin \alpha \, \frac{\mathrm{d}r}{r^2}}{\left\{\left(1 - \frac{e\Phi_0}{\frac{1}{2}mv_0^2}\right) + \frac{e\Phi_0}{\frac{1}{2}mv_0^2}\,\frac{R_c}{r} - \left(\frac{R_c}{r}\sin\alpha\right)^2\right\}^{1/2}}. \tag{9}$$

By introducing the new variable $u = R_c/r$ and substituting $\varepsilon_0 = \frac{1}{2}mv_0^2$, where $\varepsilon_0$ represents the initial energy of the electron, the ray equation (9) becomes

$$\mathrm{d}\varphi = \frac{-\sin\alpha \, \mathrm{d}u}{\{(1-p) + pu - u^2\sin^2\alpha\}^{1/2}}, \quad \text{where } p = \frac{e\Phi_0}{\varepsilon_0} = \frac{e\Phi_a}{(n-1)\frac{1}{2}mv_0^2}. \tag{10}$$

The solution of equation (10) by integration is found to be

$$\cos(\varphi + C_1) = -p + 2u\sin^2\alpha\{p^2 + 4(1-p)\sin^2\alpha\}^{-1/2},$$

where $C_1$ will be found from the condition that for $t = 0$, $\varphi = \varphi_0$ and $r = R_c$, or $u = 1$. Without loss of generality $\varphi_0$ may be taken as zero, leading to

$$\tan\tfrac{1}{2}\varphi = \frac{-\cos\alpha + \{(1-p) + pu - u^2\sin^2\alpha\}^{1/2}}{(p/\sin\alpha) - (u+1)\sin\alpha}. \tag{11}$$

(c) *The Gaussian image*

First the point will be calculated where the ray, according to eq. (11), intersects the anode sphere. For this point $r = R_a$, or $u = n$ may be substituted, leading to

$$\tan\tfrac{1}{2}\varphi_a = \frac{-\cos\alpha + \{(1-p) + pn - n^2\sin^2\alpha\}^{1/2}}{(p/\sin\alpha) - (n+1)\sin\alpha} \text{ (see figure 4).} \tag{12}$$

The practical value of $p$ may be very large, for instance of the order of $10^4$ or $10^5$. For the Gaussian image all terms not containing $p$ may be ignored with respect to $p$, yielding, for not exceptionally large values of $n$,

$$\tan\tfrac{1}{2}\varphi_a \approx \left\{\frac{p(n-1)\sin^2\alpha}{p^2}\right\}^{1/2}.$$

Since in this case $\varphi_a$ is also very small, this may be written as

$$\varphi_a \approx 2\left\{\frac{(n-1)\,\varepsilon_0 \sin^2\alpha}{e\Phi_0}\right\}^{1/2} = 2(n-1)\left(\frac{\varepsilon_\varphi}{e\Phi_a}\right)^{1/2}, \tag{13}$$

where $\varepsilon_\varphi = \varepsilon_0 \sin^2\alpha$, the component of the initial energy in the $\varphi$-direction (parallel to the cathode surface).

To this approximation, all electrons leaving the cathode sphere with this velocity in the $\varphi$-direction strike the anode sphere on a circle of radius

$$\varrho_a \approx R_a\varphi_a \approx 2(n-1)\,R_a\left(\frac{\varepsilon_\varphi}{e\Phi_a}\right)^{1/2} = 2d\left(\frac{\varepsilon_\varphi}{e\Phi_a}\right)^{1/2}, \tag{14}$$

where $d$ equals the distance between cathode and anode spheres. We thus find the same formula as in the case of two flat parallel electrodes (see eq. (1)).

It is now possible to determine the position of the virtual image point $P'$ of $P$. This is done by computing the tangent to the ray (11) in the point $r = R_a (u = n)$, and determining where this tangent intersects with the axis at an angle $\alpha'$.

For this purpose, in the vicinity of $u = n$, the ray equation may be written as

$$\tfrac{1}{2}\varphi \approx \left\{\frac{(u-1)\sin^2\alpha}{p}\right\}^{1/2} = (u-1)^{1/2}\left(\frac{\varepsilon_\varphi}{e\Phi_0}\right)^{1/2}. \tag{15}$$

The tangent to the ray in $r = R_a$ forms an angle $\beta$ with the normal in $r = R_a$, $\varphi = \varphi_a$, determined by

$$\tan\beta = \left(\frac{r\dot{\varphi}}{\dot{r}}\right)_{u=n} = \left(r\frac{\mathrm{d}\varphi}{\mathrm{d}r}\right)_{u=n} = \left(r\frac{\mathrm{d}\varphi}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}r}\right)_{u=n} = \left(-\frac{\mathrm{d}\varphi}{\mathrm{d}u}r\frac{R_c}{r^2}\right)_{u=n} = -\left(u\frac{\mathrm{d}\varphi}{\mathrm{d}u}\right)_{u=n}. \tag{16}$$

With eq. (10)

$$\tan\beta = \frac{n\sin\alpha}{\{(1-p) + pn - n^2\sin^2\alpha\}^{1/2}} \approx \frac{n}{(n-1)^{1/2}}\left\{\frac{\varepsilon_0\sin^2\alpha}{e\Phi_0}\right\}^{1/2}. \tag{17}$$

Since $\beta$ is very small, this may be written in virtue of eqs (6) and (14) as

$$\beta \approx n\left(\frac{\varepsilon_\varphi}{e\Phi_a}\right)^{1/2}.$$

The angle $\alpha'$ is equal to $\alpha' = \varphi_a - \beta \approx (n-2)\,(\varepsilon_\varphi/e\Phi_a)^{1/2}$. (18)

The distance $z$ of $P'$ beyond the anode sphere is given by

$$z \approx \frac{\varrho_a}{\tan\alpha'} \approx \frac{\varrho_a}{\alpha'}.$$

With equations (15) and (18)

$$z \approx \frac{2d}{n-2} = \frac{2(n-1)}{n-2} R_a. \tag{19}$$

This distance is independent of $\varepsilon$ and $\alpha$, which means that $P'$ is a sharp virtual image point of $P$. All the points on the sphere $r = R_c$ have virtual image points on a concentric sphere of radius

$$R_i = z - R_a = \frac{n}{n-2} R_a. \tag{20}$$

The virtual image, for which in equation (12) all terms not containing $p$ have been ignored with respect to $p$, may be termed the virtual *Gaussian image*. The magnification of this virtual image is given by

$$M = \frac{R_i}{R_c} = \frac{1}{n-2}. \tag{21}$$

(d) *Chromatic aberration*

We shall now calculate the distance from the intersection $P''$ of the ray with the axis to the centre $C$ of the system without ignoring the small terms in eq. (12). This will lead to an expression for the distance $\Delta$ of the virtual image point $P''$ to the Gaussian image point $P'$ as a function of $v_0$ and $\alpha$. From the maximum value of $\Delta$ the first-order chromatic aberration will be calculated.

From figure 5 it follows that

$$\overline{P''C} = \frac{R_a \sin \varphi_a}{\tan(\varphi_a - \beta)} - R_a \cos \varphi_a,$$

where $\varphi_a$ and $\beta$ are determined by eqs (12) and (17), viz.

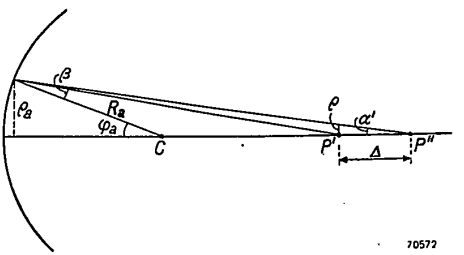



Fig. 5. Diagram for calculating the chromatic aberration.

$$\tan\frac{\varphi_a}{2} = \frac{-\cos\alpha + \{(1-p) + pn - n^2\sin^2\alpha\}^{1/2}}{(p/\sin\alpha) - (n+1)\sin\alpha},$$

and

$$\tan\beta = \frac{n\sin\alpha}{\{(1-p) + pn - n^2\sin^2\alpha\}^{1/2}}.$$

Evaluating $\Delta = \overline{P''C} - \overline{P'C}$, where $\overline{P'C} = R_i$ of eq. (20), leads to the expression

$$\Delta = R_a \frac{n(n-1)}{(n-2)}\left[\frac{\frac{n-2}{2}\frac{e\Phi_a}{\varepsilon_0}}{1 - n\sin^2\alpha + \cos\alpha\{1 - n^2\sin^2\alpha + e\Phi_a/\varepsilon_0\}^{1/2}} - 1\right]^{-1}. \quad (22)$$

If terms not containing $e\Phi_a/\varepsilon_0$ are again ignored with respect to $e\Phi_a/\varepsilon_0$ expression (22) for $\Delta$ becomes

$$\Delta = 2R_a \frac{n(n-1)}{(n-2)^2}\left(\frac{\varepsilon_0}{e\Phi_a}\right)^{1/2}\cos\alpha. \quad (23)$$

Equation (23) shows that for electrons leaving the cathode surface with only a tangential initial velocity component ($\cos\alpha = 0$), $\Delta = 0$. This means that those electrons form the virtual Gaussian image, in analogy with the case of flat parallel electrodes. The chromatic aberration is, in consequence, due to the radial components of the initial velocities of the electrons, $\varepsilon_r = \varepsilon_0\cos^2\alpha$.

The radius of the circle of confusion in the Gaussian image point is

$$\varrho = \Delta\tan\alpha' \approx \Delta\,\alpha'.$$

Together with eqs (23) and (18) this determines $\varrho$ as

$$\varrho \approx 2R_a\frac{n(n-1)}{(n-2)^2}\left(\frac{\varepsilon_0}{e\Phi_a}\right)^{1/2}\cos\alpha\,(n-2)\left(\frac{\varepsilon_0\sin^2\alpha}{e\Phi_a}\right)^{1/2} =$$

$$= R_a\frac{n(n-1)}{n-2}\frac{\varepsilon_0}{e\Phi_a}\sin 2\alpha. \quad (24)$$

The resolving power is given by $\delta = 2\varrho/M$ for $(\sin 2\alpha)_{\max} = 1$, where $M = 1/(n-2)$ (from eq. (21)), and thus becomes

$$\delta = 2n(n-1)R_a\frac{\varepsilon_0}{e\Phi_a} = 2(n-1)R_c\frac{\varepsilon_0}{e\Phi_a}. \quad (25)$$

By substituting the field gradient at the cathode $\Phi_c' = \Phi_a/(n-1)R_c$, derived by differentiating eq. (5), eq. (25) may also be written as

$$\delta = 2\frac{\varepsilon_0}{e\Phi_c'}. \tag{26}$$

This is the same expression as that found for a system consisting of a flat cathode with a uniform electric field, in connection with an electric lens (cf. eq. (2)).

Equation (26) does not represent the resolving power corresponding to the smallest possible circle of confusion. Slightly beyond the Gaussian image a smaller diameter of the virtual electron pencil can be found.

### 3. Concentric spheres with a pierced anode

Since a real image of the cathode was required, the anode sphere was provided with a hole to let the electrons pass (see figure 3). This introduces an additional negative lens, the focal length of which may be approximated by the following expression [2]):

$$f = \frac{4\Phi_h}{\Phi_2' - \Phi_1'}, \tag{27}$$

where $f$ = focal length of the lens, $\Phi_h$ = potential in the centre of the hole, $\Phi_2' - \Phi_1'$ = difference in electric field gradient of the two sides of the lens. Equation (27) can also be written as

$$f = \frac{4\{D\,(\Phi_2' - \Phi_1')/2\pi + \Phi_a\}}{\Phi_2' - \Phi_1'}, \tag{28}$$

thus introducing the diameter of the hole $D$ and changing over to the anode potential $\Phi_a$. Equation (28) becomes

$$f = \frac{2}{\pi}D + \frac{4\Phi_a}{\Phi_2' - \Phi_1'},$$

which, in the case of concentric spheres with $\Phi_2' = 0$ and $\Phi_1' = n\Phi_a/(n-1)R_a$, gives

$$f = -4R_a\left(\frac{n-1}{n} - \frac{k}{2\pi}\right), \tag{29}$$

if $D/R_a = k$ is introduced.

With the formula $1/f = 1/s + 1/s'$, where $s$ = virtual object distance, equal to $2(n-1)R_a/(n-2)$ from eq. (19) and $s'$ = distance of real image behind the hole in the anode, the value of $s'$ is

$$s' = R_a\frac{4\dfrac{n-1}{n-2}\left(\dfrac{n-1}{n} - \dfrac{k}{2\pi}\right)}{\dfrac{2(n-1)}{n} - \dfrac{n-1}{n-2} - \dfrac{k}{\pi}}. \tag{30}$$

For a very small hole, where $k/2\pi \ll 1$, eq. (30) becomes

$$s' = R_a \frac{4(n-1)}{n-4}.$$

For practical applications it is useful to derive a simpler formula by approximating $(n-1)/n - (k/2\pi)$ in eq. (29) by $\frac{3}{4}$, since the maximum value of $(n-1)/n-(k/2\pi)$ is about equal to 1 for very large $n$ and very small $k$, whilst the minimum value is about $\frac{1}{2}$ for $k_{max} = 2$ and $n \approx 4$. This leads to $f = -3\ R_a$ and to

$$s' = \frac{6(n-1)}{n-4} R_a. \tag{31}$$

The magnification is in this case

$$M = \frac{\frac{6(n-1)}{n-4} R_a}{\frac{2(n-1)}{n-2} R_a} \frac{1}{n-2} = \frac{3}{n-4}. \tag{32}$$

## 4. Significance of the theory for practical applications

For practical applications an electrode system as described in section 3 is in most cases not very suitable because the diameter of the tube becomes too large. Instead of using a cathode with a spherical surface, a design of tube may be employed where only part of the surface is spherical, the rest of the envelope being part of a cylinder (see figure 6). Testing of the theory in its most exact form will therefore be hardly possible in practical cases, since the ratio $n$ of the radii $R_c$ and $R_a$ does not entirely determine the shape of the field. Because the field does not correspond to the field in the pure-spherical model, the position of the anode has to

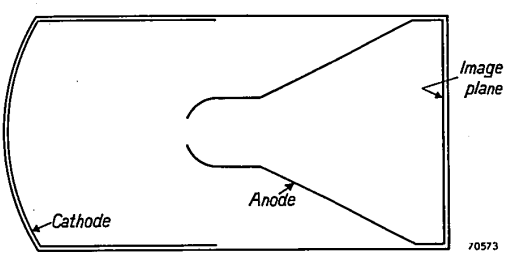



Fig. 6. Diagram of a tube for practical applications employing the electron optical system consisting of two concentric spheres.

be determined experimentally in order to obtain an image from a maximum area of the cathode.

Notwithstanding the departure from the theoretical model in practical constructions, it has been found that the formulas provide a very useful approximation. If a certain magnification or demagnification is required it appears to be useful to start experiments with the ratio of radii of the spheres as given by eq. (32). Since under experimental conditions $\varepsilon_0 \ll e\Phi_a (\Phi_a \geqq 10$ kV and $\varepsilon_0 \approx 1$ eV), the image has a considerable depth of focus. Therefore good pictures on a flat surface have been obtained in various experimental tubes, where a small pin-cushion distortion was the most prominent aberration. A description of a tube in which such a system is used will be published elsewhere [3]).

*Eindhoven, September 1951*

REFERENCES:

1) G. Holst, J. H. de Boer, M. C. Teves and C. F. Veenemans, Physica **1**, 297, 1934.
2) A. A. Rusterholz, Elektronenoptik I, Grundzüge der theoretisehen Elektronenoptik, Birkhäuser, Basel, 1950, pp. 51 and 126.
3) M. C. Teves and T. Tol, Philips tech. Rev. **14**, 65-75, 1952.

R 188 Philips Res. Rep. 7, 131-157, 1952

# ON THE SCATTERING MATRIX OF SYMMETRICAL WAVEGUIDE JUNCTIONS *)

by A. E. PANNENBORG 621.392.26:621.317.3

**Summary**

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. General considerations about the consequences of structural symmetry of microwave circuits lead up to an extensive discussion of junctions consisting of two parallel sections of rectangular waveguide having one of their sides in common. The theory for directional couplers obeying these symmetry requirements is derived. Two instruments, viz. an attenuator and a standard matching transformer, both having a directional coupler as basic unit, are described in detail. The last chapter, finally, deals with resonant directional couplers.

**Résumé**

Description systématique du comportement électromagnétique des circuits micro-ondes à l'aide de la „scattering matrix". L'auteur étudie en particulier les résonateurs. Il examine, du point de vue utilisation pratique, les résultats fondamentaux obtenus par Tomonaga sur cette question. Des considérations générales sur les conséquences de la symétrie structurale des circuits micro-ondes conduisent à une discussion détaillée de joints consistant en deux sections en parallèle de guides d'ondes rectangulaires ayant un coté commun. L'auteur en dérive la théorie des coupleurs directionnels répondant à ces conditions de symétrie. Il décrit ensuite en détail deux appareils, à savoir un atténuateur et un transformateur „normal", comportant tous deux un coupleur directionnel comme unité de base. Enfin, le dernier chapitre est consacré aux coupleurs de résonance directionnels.

**Zusammenfassung**

Es wird das elektromagnetische Verhalten kurzwelliger Kreise mit Hilfe der „scattering matrix" systematisch entwickelt. Besondere Aufmerksamkeit wird den Resonatoren gewidmet. Die von Tomonaga über diesen Gegenstand gewonnenen grundlegenden Resultate werden bis zu ihrer praktischen Anwendbarkeit entwickelt. Allgemeine Betrachtungen über die Folgerungen, die sich aus struktureller Symmetrie von Kurzwellenkreisen ergeben, führen zu einer ausführlichen Diskussion von Verbindungsstücken, welche aus zwei parallelen Gliedern bestehen, rechteckigen Hohlleitern, die eine Seite gemeinsam haben. Es wird die Theorie für „directional couplers", welche diesen Symmetrieforderungen genügen, abgeleitet. Es folgt eine ausführliche Beschreibung zweier Instrumente, nämlich eines Abschwächers und eines normalen Anpassungstransformators, die beide einen Richtungskoppler als Grundeinheit besitzen. Das letzte Kapitel schließlich behandelt Resonanz-Richtungskoppler.

*) Thesis, Technical University of Delft, February, 1952.

**CONTENTS**

INTRODUCTION

Chap. I. FUNDAMENTAL THEOREMS
- I. 1. General remarks
- I. 2. Fields in waveguides
- I. 3. Scattering matrix
- I. 4. Shift in position of terminal reference planes
- I. 5. Symmetry of the scattering matrix
- I. 6. Lossless junctions
- I. 7. Frequency dependence of a unitary scattering matrix

Chap. II. RESONATORS
- II. 1. Quality factor
- II. 2. Fundamental considerations
- II. 3. Resonators with one output lead
- II. 4. Resonators with two output leads
- II. 5. Losses in the output circuit

Chap. III. SYMMETRY ANALYSIS OF WAVEGUIDE JUNCTIONS
- III. 1. Representation of a symmetry group
- III. 2. Relation between the symmetry of a junction and its scattering matrix
- III. 3. Field distribution in a symmetry plane

Chap. IV. JUNCTIONS OF FOUR WAVEGUIDES WITH TWO-FOLD PLANAR SYMMETRY
- IV. 1. The general case
- IV. 2. The nature of the solution
- IV. 3. The degenerate case
- IV. 4. The wave matrix
- IV. 5. Shift of the reference planes

Chap. V. DIRECTIONAL COUPLERS
- V. 1. Condition for perfect directivity
- V. 2. Bethe-hole coupler
- V. 3. Systems of identical junctions connected in cascade
- V. 4. Frequency dependence of directional couplers
- V. 5. Variable attenuator
- V. 6. Standard matching transformer

Chap. VI. RESONANT SLOTS
- VI. 1. Single slots
- VI. 2. Measurements on single slots
- VI. 3. Resonant directional coupler

## INTRODUCTION

The physical configuration of microwave circuits is in all but the simplest cases too intricate to enable one to solve Maxwell's equations subject to the boundary conditions imposed. For practical purposes, however, it is not necessary to know explicitly the components of the electromagnetic field at every point of the interior of the structure. It often suffices to know how the structure affects the field in the leads connecting it to the outside.

This approach to the behaviour of microwave circuits, sometimes called transducer theory, has been made up till now in the majority of cases with the aid of equivalent circuits. It has proved itself to be a powerful tool for the analysis of waveguide junctions. The physical insight so obtained is very illuminating, especially for the radio engineer accustomed to low-frequency lumped-circuit concepts.

On the other hand for structures with more than two or three output leads the equivalent circuit merely serves to provide a well-defined framework in which numerical values obtained from experiments can be inserted. In these more intricate cases the value of the representation as a guide to better understanding seems at least doubtful; anyway much of the charm of the method is lost. The way out of this difficulty is found in an alternative method, where the description of the field in the output leads of a junction is made by specifying not the ratios of electric and magnetic components, but rather the ratios of amplitudes of travelling waves incident on and emergent from the junction. In analogy with other branches of physics the matrix expressing the relation between these wave amplitudes is called the *scattering matrix* of the junction. Up till now the scattering matrix has been used to a limited extent only and — for the reasons stated above — especially in the case of junctions with more than two output leads.

Apart from the general theorems supplied to transducer theory by Maxwell's equations, a source of information is often available in the form of the symmetries pertaining to the structure under consideration. In very simple cases the consequences of symmetry can be found by inspection. For more intricate situations, a formalism has been developed by Dicke [1]); it introduces symmetry operators, subject to which the solution of Maxwell's equations must remain invariant. In this way it sometimes becomes possible to construct a number of special solutions, the most general solution then being a linear combination of these. The number of parameters needed to specify the behaviour of a junction can thus be considerably reduced. For the symmetrical structures with many output leads, where the formalism is most useful, the scattering matrix is preferable

to impedance or admittance representations. Therefore the symmetry formalism usually is applied to the former method.

As stated above, the scattering matrix has found little application for simple circuits. Especially in the case of resonators the impedance or admittance description has been used almost exclusively. This is quite logical from a historical point of view as work in the microwave range originates from a gradual extension of techniques well known for longer wavelengths. In this way it was natural to transfer the familiar concept of an LCR circuit to microwave resonators. That this procedure is quite correct was proved subsequently by several authors, e.g. by Slater [2]).

On the other hand Tomonaga [3]) has shown how the scattering matrix can be used equally well for the description of the behaviour of resonators. The results are, of course, identical with those from calculations on an impedance basis and whilst to the radio engineer the simple LCR representation for a resonator is very attractive, the scattering description should appeal for analogous reasons to the physicist. Furthermore the latter method seems to involve slightly less specific assumptions about the nature of the resonant structure.

In Chap. I a short recapitulation is given of the fundamental properties of the scattering matrix. This seems justified since elsewhere in literature[1]),[4]) the theorems are derived, using almost invariably the impedance matrix as an intermediate station. Chapter II contains a slight extension of Tomonaga's theory together with some examples demonstrating the practical usefulness of this theory. Chapter III gives an outline of the symmetry formalism. In Chap. IV the general properties of structures with four output leads having a high degree of symmetry are derived. In Chap. V these results are applied to waveguide systems with special reference to directional couplers, while finally Chap. VI deals with analogous systems incorporating resonant elements.

## CHAPTER I. FUNDAMENTAL THEOREMS

### I. 1. General remarks

As in the following pages we shall be concerned with certain applications of transducer theory to microwave structures and as these investigations will be carried out with the aid of the scattering matrix, it seems proper to start with the derivation of some fundamental properties of this matrix. Because extensive demonstrations constituting the rigorous foundation of transducer theory can be found elsewhere[1]),[4]), we will confine ourselves to a recapitulation of those proofs that are of direct importance for the contents of the following chapters. In this connection it is supposed that the reader is familiar with waveguide theory.

Before embarking on the subject we will state some conventions and restrictions that will be adhered to throughout.

(i) The time factor is understood to be $e^{-i\omega t}$. Unless specifically stated otherwise $\omega$ is real and the fields vary harmonically with time.
(ii) All structures considered are linear and passive. Anisotropic media are allowed with the limitation that the tensors describing their permittivity, permeability, and conductivity must be symmetrical.
(iii) It is possible to surround a structure by a closed surface of finite dimensions on which all field quantities are zero except where this surface cuts the output leads.
(iv) The leads, in which the terminal fields relating to a structure are defined, consist of ideal (i.e. lossless and cylindrical) waveguides, not containing anisotropic materials.
(v) In each output lead power is transported by one mode only.
(vi) Terminal surfaces are drawn at sufficient distance from any discontinuity so that the effect of non-propagating modes can be ignored.
(vii) Rationalized MKS(Giorgi) units are used.

Some points need some comment. Condition (iii) merely requires that the structure and its output leads are effectively shielded. Point (v) constitutes no important restriction. In fact, because of the orthogonality properties of the characteristic solutions (called modes) of Maxwell's equations for an ideal waveguide, a system of two modes in one waveguide is within the formalism of transducer theory fully equivalent to a system of two modes each in a different waveguide. Point (vi), finally, is only of significance when the interconnection between two transducers is considered. It states that the connecting leads should be sufficiently long to prevent non-propagating modes excited in the interior of one transducer from reaching any discontinuity in the interior of the other transducer. If this condition is not fulfilled, the behaviour of the combination of two structures cannot be accurately predicted from the transducer data of the separate structures. Indeed, it should always be remembered that transducer theory combined with the necessary experiments gives for most applications sufficient, but never complete information on the repartition of the electromagnetic field.

As a consequence of points (i), (ii), and (vii) Maxwell's equations have the form

$$\left.\begin{aligned} \operatorname{curl} \mathbf{H} &= (-i\omega\varepsilon + \sigma)\,\mathbf{E}\,, \\ \operatorname{curl} \mathbf{E} &= i\omega\mu\mathbf{H}\,, \end{aligned}\right\} \tag{1}$$

the permittivity $\varepsilon$, the permeability $\mu$ and the conductivity $\sigma$ being independent of time and of the field vectors.

## I. 2. Fields in waveguides

We next recall without proof some results from waveguide theory as far as they are needed in later sections.

Let the $z$-axis be parallel to the axis of the waveguide. The general expressions for the electric and magnetic fields of a waveguide mode can be written as follows [2]):

$$\left.\begin{aligned}\mathbf{E} &= \mathbf{E}_t(a\,e^{-i\beta z} + b\,e^{i\beta z}) + \mathbf{k}E_z(a\,e^{-i\beta z} - b\,e^{i\beta z})\,,\\ \mathbf{H} &= \mathbf{H}_t(a\,e^{-i\beta z} - b\,e^{i\beta z}) + \mathbf{k}H_z(a\,e^{-i\beta z} + b\,e^{i\beta z})\,.\end{aligned}\right\} \tag{2}$$

Here $E_z$ and $H_z$ are the longitudinal components; $\mathbf{k}$ represents the unit vector in the positive $z$-direction. The transverse components $\mathbf{E}_t$ and $\mathbf{H}_t$ are vector functions (derivable from $E_z$ and $H_z$) lying in a plane perpendicular to the $z$-axis; $a$ and $b$ are amplitude coefficients for the waves in the negative and the positive $z$-direction respectively. The propagation constant $\beta$ is real for propagating modes. As the $z$-dependence has been written explicitly in eqs (2), the quantities $E_z$, $H_z$, $\mathbf{E}_t$ and $\mathbf{H}_t$ are functions of the transverse coordinates only.

Equations (2) are valid for all three classes of modes: transverse electromagnetic or TEM($E_z = 0$, $H_z = 0$), transverse electric or TE($E_z = 0$) and transverse magnetic or TM($H_z = 0$).

The transverse electric and magnetic fields are proportional to each other; the relation is

$$\mathbf{E}_t = Z_0\,(\mathbf{k} \times \mathbf{H}_t)\,, \tag{3}$$

where $Z_0$ is a constant which is real for propagating modes.

If a section of waveguide for which the field components are given by eqs (2) is fed by a generator sending power in the negative $z$-direction, where the guide is terminated by a passive load, it is natural to define a reflection coefficient as the ratio of the amplitudes of the waves incident on and reflected by the load. As the transverse electric field is most easily accessible to measurement, we base the definition on this component; the reflection coefficient $\Gamma$ thus derived from eqs (2) is

$$\Gamma = \frac{b\,e^{i\beta z}}{a\,e^{-i\beta z}} = \frac{b}{a}\,e^{2i\beta z}. \tag{4}$$

To determine the expression for power flow, first the $z$-component of the time-average Poynting vector $\mathbf{S}$ is formed, viz.

$$S_z = \tfrac{1}{4}\,\mathbf{k}(\mathbf{E} \times \mathbf{H}^* + \mathbf{E}^* \times \mathbf{H}). \tag{5}$$

The asterisk denotes the complex conjugate. By substituting eqs (2) and

(3) in eq. (5) it becomes clear that only the transversc field components ntribute to $S_z$:

$$S_z = \frac{|E_t|^2}{2Z_0} \{|b|^2 - |a|^2\} = \tfrac{1}{2} Z_0 |H_t|^2 \{|b|^2 - |a|^2\}. \tag{6}$$

The power flow $P$ in the positivc $z$-direction is obtained by integrating $S_z$ over the cross-section $\Omega$ of the guide. But first we introduce the normalization condition. The expression for power flow is brought into its most simple form by putting

$$\frac{1}{Z_0} \int_\Omega |E_t|^2 \, \mathrm{d}\Omega = Z_0 \int_\Omega |H_t|^2 \, \mathrm{d}\Omega = 1. \tag{7}$$

With this condition $P$ becomes

$$P = \tfrac{1}{2} \{|b|^2 - |a|^2\}. \tag{8}$$

When applying eq. (8) to the setup with generator and load as used for deriving eq. (4) it is to be noted that it represents the power flow from the load towards the generator.

### I. 3. Scattering matrix

Consider a waveguide junction subject to the conditions stated in Sec. 1. In each of the $N$ output leads a coordinate system is introduced, the $z$-axis being parallel to the axis of the waveguide and pointing outward from the junction. The electromagnetic field in each lead can be written in the form of eqs (2), indicating the presence in output lead $(n)$ of an incident wave of amplitude $a_n$ and an outgoing wave of amplitude $b_n$. The linearity of Maxwell's equations together with condition (ii), Sec. 1, requires that the amplitudes $b_n$ of the outgoing waves are linear functions of the amplitudes $a_n$ of the incident waves. Grouping the quantities $a_n$ and $b_n$ together in column matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ respectively, this linear relation can be formally represented by

$$\boldsymbol{B} = \boldsymbol{S}\boldsymbol{A}, \tag{9}$$

the elements of the *scattering matrix* $\boldsymbol{S}$ being independent of the amplitudes, $a_n$ and $b_n$, and of time.

The physical meaning of the elements of $\boldsymbol{S}$ is closely related to the reflection coefficient $\Gamma$ introduced in eq. (4). In fact, applying eq. (9) to the case of a structure with only one lead, i.e. a waveguide terminated by a load, the single remaining element of $\boldsymbol{S}$ is seen to be identical with $\Gamma(z = 0)$. In the general case with $N$ output leads the diagonal elements of the square matrix $\boldsymbol{S}$ of order $N$ relate outgoing waves to waves incident in the same lead; they are of the same nature as a reflection coefficient.

The off-diagonal elements, relating the outgoing wave in one lead to waves incident in other leads, may be called transmission coefficients.

If we put all $a_n$ equal to zero except $a_m$, which situation is in practice achieved by connecting a generator to lead $(m)$, all other leads being terminated by a matched load, and choose $a_m$ equal to 1, the elements of the $m$th column of $\boldsymbol{S}$ give the amplitudes of the outgoing waves.

### I. 4. Shift in position of terminal reference planes

Just as the phase angle of the reflection coefficient $\Gamma$ defined by eq. (4) depends on the specific position along the lead, the phase angles of the elements of the scattering matrix $\boldsymbol{S}$ also vary with the choice of the origins of the respective $z$-axes. It is in this connection useful to study the transformation of $\boldsymbol{S}$ caused by a change of the origin of any $z_n$-axis. Referring to fig. 1, let the $z'_n$-axis be shifted over a distance $l_n$ with respect to the $z_n$-axis away from the junction.

The appropriate phase shift is

$$\varphi_n = \beta_n l_n . \tag{10}$$

Denoting with a prime the quantities relating to the new origin, we have

$$a'_n = a_n e^{-i\varphi_n}, \qquad b'_n = b_n e^{i\varphi_n}, \tag{11}$$

or in matrix notation

$$\boldsymbol{A}' = \boldsymbol{\Phi}^{-1}\boldsymbol{A}, \qquad \boldsymbol{B}' = \boldsymbol{\Phi}\boldsymbol{B}. \tag{12}$$

$\boldsymbol{\Phi}$ is a diagonal matrix with elements

$$\Phi_{nn} = e^{i\varphi_n}, \qquad \Phi_{nm} = 0 \ (n \neq m). \tag{13}$$

Solving eqs (12) for $\boldsymbol{A}$ and $\boldsymbol{B}$ and inserting the result in eq. (9) we obtain

$$\boldsymbol{\Phi}^{-1}\boldsymbol{B}' = \boldsymbol{S}\boldsymbol{\Phi}\boldsymbol{A}', \tag{14}$$

or

$$\boldsymbol{B}' = \boldsymbol{\Phi}\boldsymbol{S}\boldsymbol{\Phi}\boldsymbol{A}'. \tag{15}$$

Hence it is seen that the scattering matrix $\boldsymbol{S}'$, related to the new origins, is connected to $\boldsymbol{S}$ by the simple relation

$$\boldsymbol{S}' = \boldsymbol{\Phi}\boldsymbol{S}\boldsymbol{\Phi}. \tag{16}$$

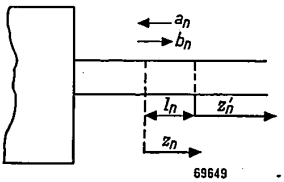



Fig. 1. An output lead of a waveguide junction showing the old and new $z$-axes.

## I. 5. Symmetry of the scattering matrix

The well-known principle of reciprocity in electrodynamics has a direct bearing on the scattering matrix. In fact it can be demonstrated that if reciprocity holds, as is guaranteed in our case by the conditions set out in Sec. 1, the scattering matrix is symmetrical.

The proof runs as follows. Let us consider the vector identity

$$\operatorname{div}(\mathbf{E}_1\times\mathbf{H}_2-\mathbf{E}_2\times\mathbf{H}_1)=\mathbf{H}_2\operatorname{curl}\mathbf{E}_1-\mathbf{E}_1\operatorname{curl}\mathbf{H}_2-\mathbf{H}_1\operatorname{curl}\mathbf{E}_2+\mathbf{E}_2\operatorname{curl}\mathbf{H}_1. \quad (17)$$

Here the subscripts 1 and 2 refer to two solutions of Maxwell's equations for the same junction and the same frequency but for different terminal conditions. Eliminating the curls with the aid of eqs (1) we find the right-hand side of eq. (17) to be identically zero. Hence

$$\operatorname{div}(\mathbf{E}_1\times\mathbf{H}_2-\mathbf{E}_2\times\mathbf{H}_1)=0. \quad (18)$$

Equation (18) is now integrated over the volume $V$ of the junction and this integral is transformed into an integral over the surface $F$ enclosing the junction and cutting the output leads perpendicularly to their respective $z$-axes. This gives

$$\int_V \operatorname{div}(\mathbf{E}_1\times\mathbf{H}_2-\mathbf{E}_2\times\mathbf{H}_1)\,\mathrm{d}V=\int_F(\mathbf{E}_1\times\mathbf{H}_2-\mathbf{E}_2\times\mathbf{H}_1)\,\mathrm{d}\mathbf{F}=0. \quad (19)$$

$\mathrm{d}\mathbf{F}$ is normal to the surface $F$. As the junction has been assumed to be perfectly shielded, the only contributions different from zero in eq. (19) come from those portions $\Omega_n$ of the surface $F$ where it cuts the output leads. Therefore

$$\int_F(\mathbf{E}_1\times\mathbf{H}_2-\mathbf{E}_2\times\mathbf{H}_1)\mathrm{d}\mathbf{F}=\sum_n\int_{\Omega_n}(\mathbf{E}_1\times\mathbf{H}_2-\mathbf{E}_2\times\mathbf{H}_1)\,\mathbf{k}_n\,\mathrm{d}\Omega_n=0. \quad (20)$$

We now identify solution 1 with the case that there is an incident wave in lead (1) only.

Solution 1:
$$\left.\begin{aligned} a_n &= 0 \quad n\neq 1,\\ b_n &= S_{n1}\,a_1 \quad .\end{aligned}\right\} \quad (21)$$

In the same way

Solution 2:
$$\left.\begin{aligned} a_n &= 0 \quad n\neq 2,\\ b_n &= S_{n2}\,a_2 \quad .\end{aligned}\right\} \quad (22)$$

Substituting eqs (21) and (22) together with eqs (2), (3) and (7) into eq. (20) we find that the only contributions come from leads (1) and (2). Evaluating these we obtain from lead (1)

$$(a_1 e^{-i\beta_1 z_1}+S_{11}\,a_1 e^{i\beta_1 z_1})\,S_{12}\,a_2 e^{i\beta_1 z_1}+ S_{12}\,a_2 e^{i\beta_1 z_1}(a_1 e^{-i\beta_1 z_1}-S_{11}\,a_1 e^{i\beta_1 z_1})=2\,S_{12}\,a_1 a_2, \quad (23)$$

and from lead (2)

$$-S_{21}\,a_1 e^{i\beta_1 z_1}(a_2 e^{-i\beta_1 z_1} - S_{22}\,a_2 e^{i\beta_1 z_1}) -$$
$$-(a_2 e^{-i\beta_1 z_1} + S_{22}\,a_2 e^{i\beta_1 z_1})\,S_{21}\,a_1 e^{i\beta_1 z_1} = -2\,S_{21}\,a_1 a_2\,. \quad (24)$$

According to eq. (20) the quantities (23) and (24) add up to zero. Hence

$$S_{12} = S_{21}\,. \quad (25)$$

Now it is clear that only for the sake of clarity have we conducted the above argument specifically for leads (1) and (2). Nothing special has been assumed about them. Therefore eq. (25) can be generalized to

$$S_{nm} = S_{mn}\,, \quad (26)$$

or, in matrix notation,

$$S = \tilde{S}, \quad (27)$$

where $\tilde{S}$ represents the transpose of $S$. This completes the proof that the scattering matrix is symmetrical. It is to be noted that the validity of this result is closely linked with the particular choice of the normalization condition, eq. (7).

### I. 6. Lossless junctions

If a junction is lossless, it follows from physical considerations that, once a stationary state has been attained ($\omega$ real), the net power flow into the junction must be zero. It will be proved that as a consequence the scattering matrix of the junction must be unitary. We recall that a matrix is unitary if the product of its transpose with its complex conjugate yields the unit matrix.

Preceding this proof a general theorem will be derived that will be needed later on. Let us consider the identity

$$\operatorname{div}(\mathbf{E}\times\mathbf{H}^* + \mathbf{E}^*\times\mathbf{H}) = \mathbf{H}^*\operatorname{curl}\mathbf{E} - \mathbf{E}\operatorname{curl}\mathbf{H}^* + \mathbf{H}\operatorname{curl}\mathbf{E}^* - \mathbf{E}^*\operatorname{curl}\mathbf{H}. \quad (28)$$

The curls in eq. (28) can be eliminated with the aid of Maxwell's equations (1). In this connection complex values of $\omega$ are allowed; on the other hand $\varepsilon$ and $\mu$ are restricted to real values *). The result is

$$\operatorname{div}(\mathbf{E}\times\mathbf{H}^* + \mathbf{E}^*\times\mathbf{H}) = i(\omega-\omega^*)\{\mu|H|^2 + \varepsilon|E|^2\} - 2\sigma|E|^2\,. \quad (29)$$

This equation is integrated over the volume $V$; the left-hand side can be converted into a surface integral. Thus

$$\int_F (\mathbf{E}\times\mathbf{H}^* + \mathbf{E}^*\times\mathbf{H})\,\mathrm{d}\mathbf{F} = i(\omega-\omega^*)\int_V \{\mu|H|^2 + \varepsilon|E|^2\}\,\mathrm{d}V - 2\int_V \sigma|E|^2\,\mathrm{d}V. \quad (30$$

It is to be noted that

$$\tfrac{1}{4}\int_V \{\mu|H|^2 + \varepsilon|E|^2\}\,\mathrm{d}V = W\,. \quad (31)$$

*) This is only a formal restriction. Dielectric losses are already accounted for by $\sigma$. Losses of a magnetic nature could be easily included by adding a real term proportional to $|H|^2$ on the right-hand side of eq. (29).

IPR2025-00217 EXHIBIT 2010 Page 144 of 485

is the energy stored in the volume $V$. Further

$$\tfrac{1}{2}\int_V \sigma|E|^2\,\mathrm{d}V = D \tag{32}$$

is the power dissipated in the volume $V$, and finally

$$\tfrac{1}{4}\int_F (\mathbf{E}\times\mathbf{H}^* + \mathbf{E}^*\times\mathbf{H})\,\mathrm{d}\mathbf{F} = P \tag{33}$$

represents the net power flowing through $F$ out of $V$. Introducing the symbols $W$, $D$ and $P$ in eq. (30) we find

$$P = i(\omega-\omega^*)\,W - D. \tag{34}$$

If eq. (30) or eq. (34) is applied to a lossless junction for stationary states, the right-hand side vanishes so that

$$P = \tfrac{1}{4}\int_F (\mathbf{E}\times\mathbf{H}^* + \mathbf{E}^*\times\mathbf{H})\,\mathrm{d}\mathbf{F} = 0\,. \tag{35}$$

The contributions of the various output leads to eq. (35) can be written explicitly. Thus

$$P = \tfrac{1}{2}\sum_n (b_n b_n^* - a_n a_n^*) = 0\,. \tag{36}$$

Introducing matrix notation we find that

$$\tilde{\boldsymbol{B}}\,\boldsymbol{B}^* - \tilde{\boldsymbol{A}}\,\boldsymbol{A}^* = 0\,. \tag{37}$$

Now $\boldsymbol{B}$ can be eliminated by applying eq. (9), whence

$$\widetilde{\boldsymbol{S}\boldsymbol{A}}\,\boldsymbol{S}^*\boldsymbol{A}^* - \tilde{\boldsymbol{A}}\boldsymbol{A}^* = 0, \tag{38}$$

or

$$\tilde{\boldsymbol{A}}(\tilde{\boldsymbol{S}}\,\boldsymbol{S}^* - \boldsymbol{I})\boldsymbol{A}^* = 0\,. \tag{39}$$

Since eq. (30) is valid for all values of $\boldsymbol{A}$, it can be concluded that

$$\tilde{\boldsymbol{S}}\,\boldsymbol{S}^* = \boldsymbol{I}. \tag{40}$$

Thus the scattering matrix of a lossless junction is unitary.

### I. 7. Frequency dependence of a unitary scattering matrix

From a reasoning similar to that of the two previous sections a general expression for the frequency dependence of the scattering matrix of a lossless junction can be derived. The result is equivalent to Foster's reactance theorem in network theory.

Consider the identity

$$\begin{aligned}\operatorname{div}(\mathbf{E}_1\times\mathbf{H}_2^* + \mathbf{E}_2^*\times\mathbf{H}_1) &= \\ &= \mathbf{H}_2^*\operatorname{curl}\mathbf{E}_1 - \mathbf{E}_1\operatorname{curl}\mathbf{H}_2^* + \mathbf{H}_1\operatorname{curl}\mathbf{E}_2^* - \mathbf{E}_2^*\operatorname{curl}\mathbf{H}_1.\end{aligned} \tag{41}$$

Here the subscripts 1 and 2 refer to two solutions of Maxwell's equations for the same lossless structure but for different (real) frequencies. With the aid of eqs (1) the curls in the right-hand side of eq. (41) can be eliminated:

$$\operatorname{div}(\mathbf{E}_1 \times \mathbf{H}_2^* + \mathbf{E}_2^* \times \mathbf{H}_1) = (\omega_1 - \omega_2)\{\mu \mathbf{H}_1 \mathbf{H}_2^* + \varepsilon \mathbf{E}_1 \mathbf{E}_2^*\}. \quad (42)$$

As before eq. (42) is integrated over the volume $V$, yielding

$$\int_F (\mathbf{E}_1 \times \mathbf{H}_2^* + \mathbf{E}_2^* \times \mathbf{H}_1)\,\mathrm{d}\mathbf{F} = i(\omega_1 - \omega_2) \int_V \{\mu \mathbf{H}_1 . \mathbf{H}_2^* + \varepsilon \mathbf{E}_1 . \mathbf{E}_2^*\}\,\mathrm{d}V. \quad (43)$$

For $\omega_1 = \omega_2$ eq. (43) is seen to become identical with eq. (35) as it should be.

Next we assume $\omega_2$ to differ from $\omega_1$ by an infinitesimal amount $\delta\omega$. The field quantities with index 2 then differ from $\mathbf{E}_1$ and $\mathbf{H}_1$ by small variations only, viz.

$$\left.\begin{aligned} \delta\mathbf{E} &= \mathbf{E}_2 - \mathbf{E}_1\,, \\ \delta\mathbf{H} &= \mathbf{H}_2 - \mathbf{H}_1\,. \end{aligned}\right\} \quad (44)$$

Substituting eqs (44) in eq. (43) we obtain to the first order

$$\int_F (\mathbf{E} \times \delta\mathbf{H}^* + \delta\mathbf{E}^* \times \mathbf{H})\mathrm{d}\mathbf{F} = -i\delta\omega \int_V \{\mu|H|^2 + \varepsilon|E|^2\}\,\mathrm{d}V = -4\,iW\delta\omega\,, \quad (45)$$

where the subscript 1 has been dropped as it is no longer needed.

On the left-hand side of eq. (45) the contributions from the various terminal surfaces can be introduced explicitly. With the aid of eqs (2), (3) and (7) we find

$$\int_F (\mathbf{E} \times \delta\mathbf{H}^* + \delta\mathbf{E}^* \times \mathbf{H})\,\mathrm{d}\mathbf{F} = 2 \sum_n (b_n \delta b_n^* - a_n \delta a_n^*)\,. \quad (46)$$

Equation (46) can be written in matrix notation and combined with eq. (45) resulting in

$$\tilde{\boldsymbol{B}}\delta\boldsymbol{B}^* - \tilde{\boldsymbol{A}}\delta\boldsymbol{A}^* = -2\,i\,W\delta\omega\,. \quad (47)$$

Now $\boldsymbol{B}$ can be eliminated according to

$$\tilde{\boldsymbol{B}}\delta\boldsymbol{B}^* = \widetilde{\boldsymbol{SA}}\,(\delta\boldsymbol{S}^*\,\boldsymbol{A}^* + \boldsymbol{S}^*\,\delta\boldsymbol{A}^*) = \tilde{\boldsymbol{A}}\,\tilde{\boldsymbol{S}}\,\delta\boldsymbol{S}^*\,\boldsymbol{A}^* + \tilde{\boldsymbol{A}}\delta\boldsymbol{A}^*. \quad (48)$$

Substitution of eq. (48) in eq. (47) gives

$$\tilde{\boldsymbol{A}}\,\tilde{\boldsymbol{S}}\,\delta\,\boldsymbol{S}^*\boldsymbol{A}^* = -2\,i\,W\,\delta\omega\,. \quad (49)$$

Taking the complex conjugate of eq. (49) and putting

$$\frac{\delta \boldsymbol{S}}{\delta\omega} = \frac{\mathrm{d}\boldsymbol{S}}{\mathrm{d}\omega}, \quad (50)$$

we finally obtain the equation for the derivative of the scattering matrix of a lossless junction with respect to frequency, viz.

$$\tilde{A}^* \tilde{S}^* \frac{\mathrm{d}S}{\mathrm{d}\omega} A = 2\,i\,W. \tag{51}$$

We shall now apply this general result to the special case of a single waveguide terminated by a lossless structure. The matrices in eq. (51) then reduce to scalars; $S$ degenerates to the reflection coefficient $\Gamma(z = 0)$. Hence

$$a^* \Gamma^* \frac{\mathrm{d}\Gamma}{\mathrm{d}\omega} a = 2\,i\,W. \tag{52}$$

As the termination is lossless, the "matrix" $\Gamma$ is unitary or

$$\Gamma\,\Gamma^* = 1. \tag{53}$$

This equation is satisfied by putting

$$\Gamma = e^{i\varphi}, \tag{54}$$

where $\varphi$ is restricted to real values. We further notice that the power incident on the junction $P_i$ is given by

$$P_i = \tfrac{1}{2}\, a^* a. \tag{55}$$

If eqs (54) and (55) are substituted into eq. (52), the final result

$$\frac{\mathrm{d}\varphi}{\mathrm{d}\omega} = \frac{W}{P_i} \tag{56}$$

is obtained. This equation is important not so much for the quantitative information which it gives as for the fact that the right-hand side is essentially positive. Thus

$$\frac{\mathrm{d}\varphi}{\mathrm{d}\omega} > 0, \tag{57}$$

or in words: for a lossless termination the phase angle of the reflection coefficient always increases with frequency.

## CHAPTER II. RESONATORS

In this chapter we shall be concerned with transducer theory for structures exhibiting resonance phenomena. As stated in the introduction, in writing the next pages we have been inspired by a little-known paper of Tomonaga [3]). This paper contains much valuable material, but it has one

serious shortcoming in that it deals with lossless structures only. Although this idealization is justified for many practical configurations, it forms especially for resonators a prohibitive restriction. We shall, accordingly, extend Tomonaga's theory so as to include the effect of losses.

## II. 1. Quality factor

We start by reviewing the essential property in which a resonator differs from other junctions. Let us assume that a certain amount of energy has been fed into a junction. If now the junction is left to itself and if it contains a resonating element, outgoing exponentially damped waves will appear in the output leads, which are for the moment assumed to be terminated by matched loads. In other words the amount of electromagnetic energy, stored inside the junction, is gradually decreasing because of the power loss through the various output leads and the dissipation inside the junction proper. It is this gradual exponential decrease of the stored energy (and at the same time of the amplitudes of the corresponding outgoing waves) which forms the characteristic feature of a resonator.

It is well known that microwave resonators have an infinite number of resonance frequencies. In the following we will develop the theory under the following restrictions. The angular frequency $\omega$ shall lie close to one resonance frequency and be far removed from all others. This particular resonance frequency shall moreover be simple, i.e. there shall be only one mode of oscillation associated with it. Both conditions are fulfilled in many practically important cases. The theory, however, may readily be extended so as to include more general possibilities.

The theorem derived in Sec. I.6, eq. (34), furnishes quantitative information on the decrement of the oscillation in a resonator. We apply this equation,

$$P = i(\omega_c - \omega_c^*)\, W - D, \tag{1}$$

to the case when no power is incident on the junction. Then $P$ stands for the power leaving the junction through the various output leads, $D$ is the power dissipated within the junction and $W$ denotes the stored energy. The damped oscillation is described as an oscillation with a complex angular frequency $\omega_c$.

It will be remembered that the quality factor $Q$ of a resonator can be defined as

$$Q = \frac{\omega_0 W}{D + P}, \tag{2}$$

where $\omega_0$ is the real part of $\omega_c$.

Equation (2) can be rewritten as

$$\frac{1}{Q} = \frac{1}{\omega_0 W}\left\{D + \sum_n P_n\right\}, \tag{3}$$

or

$$\frac{1}{Q_L} = \frac{1}{Q_U} + \sum_n \frac{1}{Q_{E_n}}. \tag{4}$$

Apparently $Q_L^{-1}$ is the sum of the contribution from internal losses $D/\omega_0 W = Q_U^{-1}$ plus the contributions from each output lead individually $P_n/\omega_0 W = Q_{E_n}^{-1}$. Here a notation is employed which is customary in microwave work. $Q_L$, to be identified with $Q$ in eq. (2), is called the loaded $Q$ and represents the quality factor of the resonator loaded by all output leads. The partial $Q$'s introduced in eq. (4) are divided into the unloaded $Q$, denoted by $Q_U$, and various external $Q$'s, denoted by $Q_{E_n}$, each representing the loading effect of one output lead.

Attention should be drawn to the fact that in eq. (3) the terms $D$ and $P_n$ appear in a similar way. Whilst in eq. (I.30) their different origins are clearly indicated, in eq. (3) there is left only a formal difference in notation. The situation is, indeed, not altered if the term involving $D$ is supposed to be due to an extra output lead, the junction itself being lossless. The advantage gained by this point of view is that we can start calculations with a unitary *) scattering matrix for the idealized junction, introducing the internal losses only at a later stage. It should never be forgotten, however, that this is a purely formal way of approach and that, whilst $N$ output leads of the junction are accessible for direct measurement, the fictitious lead numbered $(N + 1)$ is not.

This procedure, viz. the concentration of the distributed losses within the junction into a fictitious extra output lead, is identical with the representation of the internal losses by a lumped resistance in equivalent-circuit theory. Compared with the latter our point of view seems to be slightly superior in that no further assumption about the nature of the fictitious output lead is necessary whereas we must always assign a definite location to a lumped resistance.

The relation between the decrement of damped oscillations in a resonator and its $Q$-factor is obtained by combining eqs (1) and (2). Thus

$$i(\omega_c - \omega_c^*) = \frac{\omega_0}{Q}. \tag{5}$$

## II. 2. Fundamental considerations

Reverting now our attention to the scattering matrix we recall that its elements represent the ratio between amplitudes of outgoing and incident

*) The scattering matrix of a lossless junction is unitary. This property is a direct result of the conservation of energy within the junction. The lemma therefore is true only for real frequencies as is also evident from the proof in Sec. I. 6.

waves. In the preceding section we have discussed the state of a resonator with only outgoing waves in the output leads. A complex frequency $\omega_c$ was used to describe these damped waves. Considering the elements of the scattering matrix as functions in the complex $\omega$-plane, it is clear that some, if not all, elements become infinite if the frequency approaches the value $\omega_c$. The most obvious possibility is that this singularity is a simple pole. Consequently we write tentatively

$$S_{nm} = \frac{p_{nm}}{\omega - \omega_c} + R_{nm}, \tag{6}$$

where $p_{nm}$ and $R_{nm}$ are independent of $\omega$. By the latter assumption we restrict the validity of eq. (6) a priori to that frequency range where the non-resonant frequency-dependence of the junction is completely obscured by the properties of the resonating element. Equation (6) therefore is at best valid in the neighbourhood of resonance.

With the aid of the following argument *) it can be shown that eq. (6) is a plausible approximation. Suppose that along lead $(m)$ a pulse of infinitesimal duration is fed into the resonator. This incoming signal can be represented by [5])

$$a_m = \lim_{\tau \to 0} \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{\sin(\omega\tau/2)}{\omega\tau/2} e^{-i\omega t} \, d\omega . \tag{7}$$

The outgoing waves $b_n$, caused by $a_m$, are obtained by multiplying the integrand in eq. (7) by the appropriate element $S_{nm}$. Inserting eq. (6) we have

$$b_n = p_{nm} \lim_{\tau \to 0} \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{\sin(\omega\tau/2)}{\omega\tau/2} \frac{e^{-i\omega t}}{\omega - \omega_c} d\omega + \\ + R_{nm} \lim_{\tau \to 0} \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{\sin(\omega\tau/2)}{\omega\tau/2} e^{-i\omega t} \, d\omega . \tag{8}$$

The first term on the right-hand side of eq. (8) can be evaluated by closing the path of integration by a semicircle of infinite radius in the negative imaginary half plane. Taking the limit $\tau \to 0$ and applying the theorem of residues we find

$$b_n = -i \, p_{nm} e^{-i\omega_c t} + R_{nm} \lim_{\tau \to 0} \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{\sin(\omega\tau/2)}{\omega\tau/2} e^{-i\omega t} \, d\omega . \tag{9}$$

*) This proof was suggested to the author by Prof. Dr H. B. G. Casimir, Director, Philips Research Laboratories.

The last term in eq. (9) represents an outgoing pulse of which the amplitude is determined by the non-resonant properties of the junction. The first term on the right in eq. (9) indicates that part of the energy contained in the incident pulse is temporarily stored within the junction and leaks away only gradually through the output leads. This behaviour agrees exactly with the views developed in Sec. 1 on resonators. The assumption that $S_{nm}$ has a simple pole for $\omega = \omega_c$ is thus seen to be well-founded.

From the preceding argument another important conclusion can be drawn. It follows from eq. (9) that the amplitudes of the outgoing damped waves, caused by the pulse incident in lead $(m)$, are proportional to the coefficients $p_{nm}$. Nothing special has been assumed about lead $(m)$; it is clearly immaterial through which lead the energy, giving rise to the outgoing damped waves, is brought into the resonator. Therefore the ratios of the coefficients $p_{nm}$ ($m$ fixed) should be independent of $m$. This fact together with the symmetry condition for the scattering matrix, which property is not restricted to real values of $\omega$ (cf. Sec. I, 5), requires, that all $p_{nm}$ have the form

$$p_{nm} = \pi_n \pi_m . \tag{10}$$

### II. 3. Resonators with one output lead

Instead of deriving the general theory for resonant junctions with any number of output leads we shall conduct the discussion along simpler lines by working out some practically important examples. Occasionally the obvious extensions to the more general case will be pointed out.

As a first example we shall analyse the behaviour of a resonator with one output lead; the coupling between resonator and output lead is supposed to be lossless. According to the argument developed at the end of Sec. 1 the situation can be schematically represented by fig. 2.1. The actual internal losses are accounted for by the matched termination in lead (2), while the resonant junction between the leads is lossless.

The scattering matrix for this structure is, according to eqs (6) and (10),

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{12} & S_{22} \end{pmatrix} = \begin{pmatrix} \dfrac{\pi_1^2}{\omega - \omega_c} + R_{11} & \dfrac{\pi_1 \pi_2}{\omega - \omega_c} + R_{12} \\ \dfrac{\pi_1 \pi_2}{\omega - \omega_c} + R_{12} & \dfrac{\pi_2^2}{\omega - \omega_c} + R_{22} \end{pmatrix}. \tag{11}$$

The symmetry condition for $S$ has been used in eq. (11). The unitary condition requires

$$\left.\begin{aligned} |S_{11}|^2 + |S_{12}|^2 = |S_{12}|^2 + |S_{22}|^2 = 1 , \\ S_{11} S_{12}^* + S_{12} S_{22}^* = 0 . \end{aligned}\right\} \tag{12}$$

These conditions must be fulfilled for all real values of $\omega$. If eq. (11) is substituted in eqs (12), terms appear in $(\omega-\omega_c)^{-1}$, $(\omega-\omega_c^*)^{-1}$ and $(\omega-\omega_c)^{-1}$ $(\omega-\omega_c^*)^{-1}$. The latter can be expanded in partial fractions, so that the resulting equations take the form

$$\frac{A}{\omega-\omega_c}+\frac{B}{\omega-\omega_c^*}+C=0, \tag{13}$$

where $A$, $B$ and $C$ are not dependent on $\omega$. If an equation like eq. (13) is to hold for all $\omega$, then $A=0$, $B=0$ and $C=0$ (note that $\omega_c$ is not real). The knowledge that the terms $C$ are zero immediately leads to the result that the matrix

$$\boldsymbol{R}=\begin{pmatrix} R_{11} & R_{12} \\ R_{12} & R_{22} \end{pmatrix} \tag{14}$$

is unitary. Actually this fact could have been deduced directly from eq. (11), as $\boldsymbol{R}$ represents the behaviour of the lossless junction far off resonance.

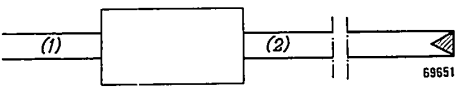



Fig. 2.1. Schematic representation of a resonator with one output lead. The matched termination in the (fictitious) lead (2) accounts for the internal losses.

The remaining equations are obtained from the coefficients of $(\omega-\omega_c)^{-1}$ and $(\omega-\omega_c^*)^{-1}$ in eqs (12) and are

$$\pi_1^*\{|\pi_1|^2+|\pi_2|^2\}+(R_{11}^*\pi_1+R_{12}^*\pi_2)\ (\omega_c-\omega_c^*)=0, \tag{15}$$

$$\pi_2^*\{|\pi_1|^2+|\pi_2|^2\}+(R_{12}^*\pi_1+R_{22}^*\pi_2)\ (\omega_c-\omega_c^*)=0. \tag{16}$$

Equations (15) and (16) are identical with the matrix equation

$$(\tilde{\Pi}\Pi^*)\Pi^*+(\omega_c-\omega_c^*)\boldsymbol{R}^*\Pi=0, \tag{17}$$

where the column matrix $\Pi$ is given by

$$\Pi=\begin{pmatrix}\pi_1\\ \pi_2\end{pmatrix}. \tag{18}$$

From eq. (17) $\boldsymbol{R}$ can be eliminated. As $\boldsymbol{R}$ is unitary, multiplication of eq. (17) by $\boldsymbol{R}$ and a subsequent change to the complex conjugate yields

$$(\tilde{\Pi}\Pi^*)\boldsymbol{R}^*\Pi-(\omega_c-\omega_c^*)\Pi^*=0. \tag{19}$$

By combination of eqs (17) and (19) we obtain

$$(\tilde{\Pi}\Pi^*)^2=-(\omega_c-\omega_c^*)^2. \tag{20}$$

As $\omega_c$ is the frequency of a damped oscillation, it has a negative imaginary part; hence the positive root of eq. (20) is the correct one, so that

$$\tilde{\Pi}\Pi^* = |\pi_1|^2 + |\pi_2|^2 = i(\omega_c - \omega_c^*) . \tag{21}$$

Equation (5) enables us to introduce the experimentally important parameters $\omega_0$ and $Q_L$:

$$|\pi_1|^2 + |\pi_2|^2 = \frac{\omega_0}{Q_L} . \tag{22}$$

In Sec. 1 we have seen that the internal losses and the loading due to the output lead contribute separately to $Q_L^{-1}$. Equation (4) adapted to the present problem can be written as

$$\frac{1}{Q_L} = \frac{1}{Q_E} + \frac{1}{Q_U} = \frac{1}{Q_1} + \frac{1}{Q_2} . \tag{23}$$

Comparison with eq. (22), where $Q_L^{-1}$ also appears as a sum of contributions from each output lead individually, compels us to conclude that

$$|\pi_1|^2 = \frac{\omega_0}{Q_1} = \frac{\omega_0}{Q_E} , \quad |\pi_2|^2 = \frac{\omega_0}{Q_2} = \frac{\omega_0}{Q_U} . \tag{24}$$

The relation between $\Pi$ and $\boldsymbol{R}$ can be derived by insertion of eq. (21) into eq. (17). The complex conjugate of the result is

$$\boldsymbol{R}\Pi^* = i\Pi . \tag{25}$$

As was pointed out before $\boldsymbol{R}$ describes the off-resonance behaviour of the resonator. The condition that the coupling between resonator and output lead is lossless implies total reflection far off resonance. Hence the off-diagonal elements of $\boldsymbol{R}$ must vanish and eq. (25) results in separate equations for $\pi_1$ and $\pi_2$. Anticipating the needs for the next section, however, we shall derive the consequences of eq. (25) without imposing any restriction on $\boldsymbol{R}$.

But for the asterisk eq. (25) would have been an eigenvalue equation for $\boldsymbol{R}$. It can be brought into this form by splitting up $\Pi$ into modulus and argument, viz.

$$\Pi = \begin{pmatrix} \pi_1 \\ \pi_2 \end{pmatrix} = \begin{pmatrix} e^{i\varphi_1} & 0 \\ 0 & e^{i\varphi_2} \end{pmatrix} \begin{pmatrix} |\pi_1| \\ |\pi_2| \end{pmatrix} = \Phi|\Pi| . \tag{26}$$

Substitution of eq. (26) in eq. (25) yields

$$\boldsymbol{R}\Phi^*|\Pi| = i\Phi|\Pi| . \tag{27}$$

Premultiplication of both sides of eq. (27) by $\Phi^{-1} = \Phi^*$ gives

$$\Phi^{-1}\boldsymbol{R}\Phi^{-1}|\Pi| \equiv \boldsymbol{R}'|\Pi| = i|\Pi| . \tag{28}$$

The matrix $\boldsymbol{R}'$, introduced in eq. (28), is derived from $\boldsymbol{R}$ by a simple shift of the reference planes in the output leads, as follows from Sec. I.4. The shift is chosen so as to make the quantities $\pi_n' = |\pi_n|$ real.

Equation (28) *is* an eigenvalue equation. The limitation it imposes on $\boldsymbol{R}'$ consists in the prescription of an eigenvalue $i$ with a corresponding eigenvector proportional to $|\Pi|$. The second eigenvalue of $\boldsymbol{R}'$ is a priori unknown but its magnitude should equal 1, because $\boldsymbol{R}'$, like $\boldsymbol{R}$, is unitary. The second eigenvector is fixed by the condition that it should be real and orthogonal to $|\Pi|$. We are now able to form $\boldsymbol{R}'$ from its eigenvalues and eigenvectors. The latter are normalized to unity. Let the second eigenvalue be $i\, e^{2i\eta}$, $\eta$ being real but otherwise unknown. In this way we obtain

$$\boldsymbol{R}' = \boldsymbol{M} \begin{pmatrix} i & 0 \\ 0 & i\, e^{2i\eta} \end{pmatrix} \boldsymbol{M}, \tag{29}$$

with

$$\boldsymbol{M} = \boldsymbol{M}^{-1} = \begin{pmatrix} \sqrt{\frac{Q_L}{Q_1}} & \sqrt{\frac{Q_L}{Q_2}} \\ \sqrt{\frac{Q_L}{Q_2}} & -\sqrt{\frac{Q_L}{Q_1}} \end{pmatrix}. \tag{30}$$

Reference to eq. (24) shows that the first column of the orthogonal matrix $\boldsymbol{M}$ is, indeed, proportional to $|\Pi|$.

Evaluating eq. (29) we find for $\boldsymbol{R}'$ the general expression

$$\boldsymbol{R}' = \begin{pmatrix} i - 2\frac{Q_L}{Q_2} \sin\eta\; e^{i\eta} & 2\frac{Q_L}{\sqrt{Q_1 Q_2}} \sin\eta\; e^{i\eta} \\ 2\frac{Q_L}{\sqrt{Q_1 Q_2}} \sin\eta\; e^{i\eta} & i - 2\frac{Q_L}{Q_1} \sin\eta\; e^{i\eta} \end{pmatrix}. \tag{31}$$

In the special case that the off-diagonal elements of $\boldsymbol{R}'$ vanish, eq. (31) reduces to the almost trivial form

$$\boldsymbol{R}' = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}. \tag{32}$$

The reflection coefficient $\Gamma$ seen in output lead (1), if output lead (2) is terminated by a matched load, becomes identical with $S_{11}$. Insertion of eqs (24) and (32) into eq. (11) yields

$$\Gamma' = S_{11}' = \frac{\omega_0}{Q_E(\omega - \omega_c)} + i. \tag{33}$$

Equation (33) represents the reflection coefficient in the output lead of the resonator at the particular reference plane fixed by eq. (28).

It is useful to eliminate $\omega_c$ from eq. (33) and to replace it by quantities better suited to measurement. To this purpose put, as is customary,

$$\omega - \omega_0 = \Delta\omega . \tag{34}$$

Then

$$\omega - \omega_c = \Delta\omega - \tfrac{1}{2}(\omega_c - \omega_c^*) . \tag{35}$$

With the aid of eqs (5) and (35) eq. (33) can be brought into its final form

$$\Gamma' = -i \frac{2Q_L}{Q_E(1 - 2i\, Q_L\, \Delta\omega/\omega_0)} + i . \tag{36}$$

The locus of $\Gamma'$ in the complex plane as a function of frequency is a circle of radius $Q_L/Q_E$ tangent to the unit circle at the point $i$, as is illustrated in fig. 2.2. It should be remembered that eq. (36) is valid for one particular reference plane. A shift of the reference plane simply corresponds to a rotation of the locus around the origin.

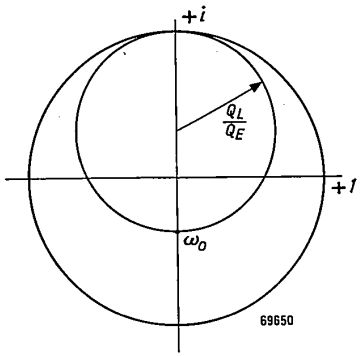



Fig. 2.2. Locus of the reflection coefficient in a waveguide terminated by a resonator.

## II. 4. Resonators with two output leads

As a second example of the use of a scattering matrix, the elements of which can be represented by eq. (6), we shall discuss the case of a resonator with two output leads attached to it. In the diagram of fig. 2.3 the output leads are designated (1) and (2). Output lead (3) accounts for the internal losses of the resonator, so that the scattering matrix pertinent to the rectangle containing the resonator is unitary. Both the coupling between the resonator and its output leads and the direct (i.e. off-resonance)

coupling between leads (1) and (2) are assumed to be lossless. As the matched load in lead (3) represents formally the losses within the resonator, no direct coupling will occur between lead (3) and the other two output leads. The off-resonance matrix $\boldsymbol{R}$ will therefore be

$$\boldsymbol{R} = \begin{pmatrix} R_{11} & R_{12} & 0 \\ R_{12} & R_{22} & 0 \\ 0 & 0 & R_{33} \end{pmatrix}. \tag{37}$$

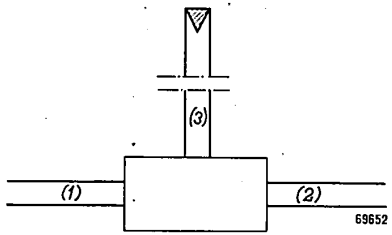



Fig. 2.3. Schematical representation of a resonator with two output leads. The matched termination in the (fictitious) lead (3) accounts for the internal losses.

For the present problem we require an obvious extension of some results of the preceding section. The proof, if needed, can be given in exact analogy to the treatment in Sec. 3. In fact, all matrix equations derived there are generally valid, irrespective of the order of the matrices, i.e. the number of output leads.

The matrix $\boldsymbol{R}$ is unitary. As a consequence the submatrix

$$\begin{pmatrix} R_{11} & R_{12} \\ R_{12} & R_{22} \end{pmatrix} \tag{38}$$

is also unitary.

The equations equivalent to eqs (23) and (24) are

$$\frac{1}{Q_L} = \frac{1}{Q_1} + \frac{1}{Q_2} + \frac{1}{Q_3}, \tag{39}$$

and

$$|\pi_1|^2 = \frac{\omega_0}{Q_1}, \quad |\pi_2|^2 = \frac{\omega_0}{Q_2}, \quad |\pi_3|^2 = \frac{\omega_0}{Q_3}. \tag{40}$$

To analyse the relation between $\boldsymbol{R}$ and $\boldsymbol{\Pi}$ it is useful to define the reference planes again in such a way that the elements of $\boldsymbol{\Pi}$ become real and eq. (28) applies. Because of eq. (37), eq. (28) falls apart in separate equations for $R'_{33}$ and the submatrix (38); the latter equation is

$$\begin{pmatrix} R'_{11} & R'_{12} \\ R'_{12} & R'_{22} \end{pmatrix} \begin{pmatrix} |\pi_1| \\ |\pi_2| \end{pmatrix} = i \begin{pmatrix} |\pi_1| \\ |\pi_2| \end{pmatrix}. \tag{41}$$

The conclusions to be drawn from this equation are almost literally the same as eqs (29) and (30) of the previous section; the only difference lies in the normalization of the modal matrix $\boldsymbol{M}$. Thus

$$\begin{pmatrix} R'_{11} & R'_{12} \\ R'_{12} & R'_{22} \end{pmatrix} = \boldsymbol{M} \begin{pmatrix} i & 0 \\ 0 & ie^{2i\eta} \end{pmatrix} \boldsymbol{M}, \tag{42}$$

where

$$\boldsymbol{M} = \boldsymbol{M}^{-1} = \begin{pmatrix} \sqrt{\frac{Q_E}{Q_1}} & \sqrt{\frac{Q_E}{Q_2}} \\ \sqrt{\frac{Q_E}{Q_2}} & -\sqrt{\frac{Q_E}{Q_1}} \end{pmatrix}. \tag{43}$$

Here we have introduced the symbol $Q_E$ to represent the total external loading of the resonator, which is given by

$$\frac{1}{Q_E} = \frac{1}{Q_1} + \frac{1}{Q_2}. \tag{44}$$

Evaluating eq. (42) we find

$$\begin{pmatrix} R'_{11} & R'_{12} \\ R'_{12} & R'_{22} \end{pmatrix} = \begin{pmatrix} i - 2\frac{Q_E}{Q_2} \sin\eta\, e^{i\eta} & 2\frac{Q_E}{\sqrt{Q_1 Q_2}} \sin\eta\, e^{i\eta} \\ 2\frac{Q_E}{\sqrt{Q_1 Q_2}} \sin\eta\, e^{i\eta} & i - 2\frac{Q_E}{Q_1} \sin\eta\, e^{i\eta} \end{pmatrix}. \tag{45}$$

Output lead (3) is now assumed to be terminated by a matched load. The scattering matrix for the resonant junction between leads (1) and (2) is obtained by omitting the third row and column from the matrix pertaining to the lossless junction discussed up till now. By eqs (6), (10) and (40) we have

$$\boldsymbol{S}' = \begin{pmatrix} \frac{\omega_0}{Q_1(\omega - \omega_c)} + R'_{11} & \frac{\omega_0}{\sqrt{Q_1 Q_2}(\omega - \omega_c)} + R'_{12} \\ \frac{\omega_0}{\sqrt{Q_1 Q_2}(\omega - \omega_0)} + R'_{12} & \frac{\omega_0}{Q_2(\omega - \omega_c)} + R'_{22} \end{pmatrix}. \tag{46}$$

It is possible to eliminate $\omega_c$ again by the introduction of $\Delta\omega = \omega - \omega_0$. If further eq. (45) is substituted in eq. (46), the final expression becomes

$$\boldsymbol{S}' = \begin{pmatrix} -i\,\dfrac{2Q_L}{Q_1(1-2iQ_L\,\Delta\omega/\omega_0)} + i - 2\,\dfrac{Q_E}{Q_2}\sin\eta\; e^{i\eta} & \\ -i\,\dfrac{2Q_L}{\sqrt{Q_1Q_2}(1-2iQ_L\,\Delta\omega/\omega_0)} + 2\,\dfrac{Q_E}{\sqrt{Q_1Q_2}}\sin\eta\; e^{i\eta} & \\ -i\,\dfrac{2Q_L}{\sqrt{Q_1Q_2}(1-2iQ_L\,\Delta\omega/\omega_0)} + 2\,\dfrac{Q_E}{\sqrt{Q_1Q_2}}\sin\eta\; e^{i\eta} & \\ -i\,\dfrac{2Q_L}{Q_2(1-2iQ_L\,\Delta\omega/\omega_0)} + i - 2\,\dfrac{Q_E}{Q_1}\sin\eta\; e^{i\eta} & \end{pmatrix}. \quad (47)$$

An instrument to which eq. (47) is applicable is the iris-coupled reaction wavemeter shown in fig. 2.4. It is not an example of eq. (47) in its most general form, as the symmetrical location of the two output leads obviously requires

$$Q_1 = Q_2\,. \quad (48)$$

After insertion of this condition into eq. (44), eq. (47) can be simplified to

$$\boldsymbol{S}' = \begin{pmatrix} -i\,\dfrac{Q_L}{Q_E(1-2iQ_L\,\Delta\omega/\omega_0)} + i - \sin\eta\, e^{i\eta}, & -i\,\dfrac{Q_L}{Q_E(1-2iQ_L\,\Delta\omega/\omega_0)} + \sin\eta\, e^{i\eta} \\ -i\,\dfrac{Q_L}{Q_E(1-2iQ_L\,\Delta\omega/\omega_0)} + \sin\eta\, e^{i\eta}, & -i\,\dfrac{Q_L}{Q_E(1-2iQ_L\,\Delta\omega/\omega_0)} + i - \sin\eta\, e^{i\eta} \end{pmatrix}.$$

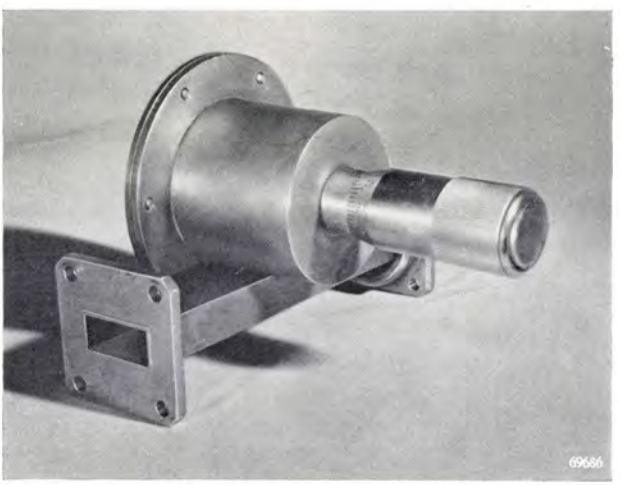



Fig. 2.4. Reaction wavemeter.

The transmission coefficient from lead (1) to lead (2) or vice versa is given by the off-diagonal elements in eq. (49). As a function of frequency it is represented in the complex plane by a circle of radius $Q_L/Q_1 = Q_L/2Q_E$. The position of this circle is determined by the value of $\eta$; its centre in particular lies on a circle, which we shall call $\eta$-circle, of radius $\frac{1}{2}$ centred on the imaginary axis at the point $(Q_L/2Q_3)i = (Q_L/2Q_U)i$, as is illustrated in fig. 2.5.

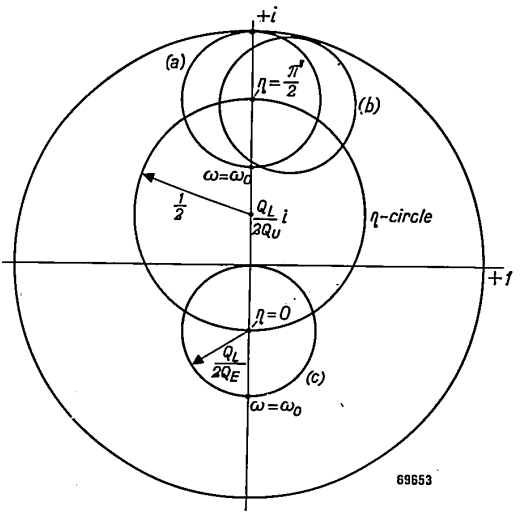



Fig. 2.5. Transmission-coefficient diagram for a resonator with two symmetrical output leads.

For the ideal reaction wavemeter, in which the off-resonance reflection caused by the coupling window is suitably compensated, the appropriate value of $\eta$ is $\frac{1}{2}\pi$. Its transmission coefficient is represented in fig. 2.5 by the circle (a).

The behaviour of a non-compensated reaction wavemeter is illustrated by the circle (b) the centre of which lies on the $\eta$-circle slightly off the imaginary axis.

The circle (c), for which $\eta = 0$, shows that eq. (49) also applies to wavemeters of the transmission type. It represents the transmission coefficient through a resonator with two symmetrically located output leads between which no direct coupling exists.

Dropping the restriction imposed by eq. (48) for symmetrical output leads we can deduce from eq. (47), for the case $\eta = 0$, that the condition for a match in output lead (1) at the resonance frequency is

$$\frac{2Q_L}{Q_1} = 1\,, \tag{50}$$

or

$$\frac{Q_U}{Q_1} = \frac{Q_U}{Q_2} + 1\,. \tag{51}$$

As a final remark it should be pointed out that the rather artificial difference between series and parallel resonant circuits, usual in equivalent-circuit theory, is completely avoided here.

### II. 5. Losses in the output circuit

The theory developed in the preceding section can be made, by a slight modification, to describe the behaviour of a resonator coupled to an ideal waveguide through a circuit of which the losses cannot be neglected. To this purpose the schematical representation of fig. 2.3. is redrawn in fig. 2.6. The output circuit is idealized to a frequency-independent T-junction. Its losses are lumped in the matched load of lead (2), while lead (3) is again included to account for the internal losses of the resonator. The configuration depicted in fig. 2.6 is a useful approximation for the analysis of a "cold" oscillator tube [6], [7]. Formally figs 2.3 and 2.6 are identical, so that eq. (47) is equally valid in the two cases.

If both leads (2) and (3) are terminated by a matched load, the reflection coefficient $\Gamma$ in lead (1) is given by $S_{11}$. Hence, from eq. (47)

$$\Gamma' = -i\,\frac{2Q_L}{Q_1(1 - 2iQ_L\,\Delta\omega/\omega_0)} + i - 2\,\frac{Q_E}{Q_2}\sin\eta\, e^{i\eta}\,. \tag{52}$$

The locus of $\Gamma'$ in the complex plane is commonly called the $Q$-circle. Once this circle has been obtained by experiment, all parameters of the resonator can be derived from it as follows:

From the relation between $\omega$ and $\Gamma'$ on the circle $Q_L$ can be determined. From the radius of the circle, which equals $Q_L/Q_1$, $Q_1$ is then calculated. It is seen from eq. (52) that the position of the centre of the $Q$-circle is

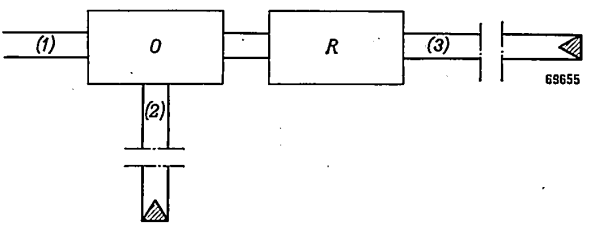



Fig. 2.6. Schematical representation of a resonator $R$ with output circuit $O$.

determined by $\eta$ and $Q_2$. For $\eta = 0$ the $Q$-circle is tangent with its off-resonance point to the unit circle at the point $i$; its centre lies for other values of $\eta$ on the $\eta$-circle, centred on the imaginary axis, with radius $Q_E/Q_2 = Q_1/(Q_1 + Q_2)$. In fig. 2.7 it is shown how the $\eta$-circle can be derived from the $Q$-circle by a geometrical construction. As a first step the off-resonance point $P$ must be located [7]), which can be done by extrapolation only, as our fundamental assumptions in Sec. 2 restrict the validity of eq. (52) to the neighbourhood of resonance. The diameter of the $Q$-circle through $P$ is drawn and parallel to it the diameter of the unit circle. The latter is to be identified with the imaginary axis; it should be realized that the reference plane for which eq. (52) applies is not known beforehand. The construction of the $\eta$-circle is now a simple matter, because it is known that (i) its centre lies on the imaginary axis, (ii) the centre of the $Q$-circle falls on it and (iii) it cuts the imaginary axis at a distance $(Q_1 - Q_L)/Q_1$ from the origin. The geometrical consequences are illustrated in fig. 2.7. After the $\eta$-circle has been constructed, $Q_2$ can be calculated from its radius. Finally $Q_3 = Q_U$ can be found with the aid of eq. (39) from the known values of $Q_L$, $Q_1$ and $Q_2$.

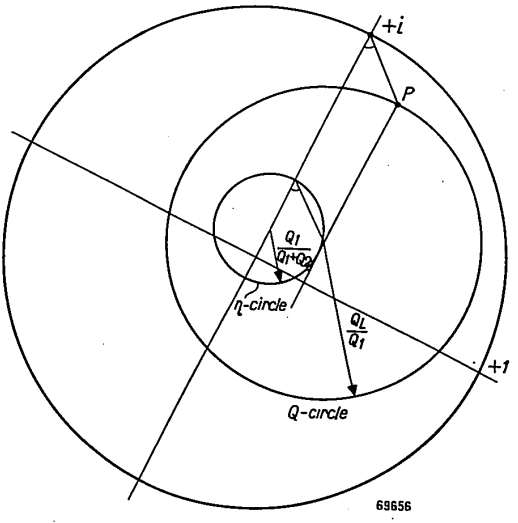



Fig. 2.7. The $Q$-circle and the $\eta$-circle.

*(To be continued)*

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

Reprints of the majority of these papers can be obtained on application to the Administration of the Research Laboratory, Kastanjelaan, Eindhoven, Netherlands. Those papers of which no reprints are available are marked with an asterisk (*).

**1962:** K. F. Niessen: On the deviations between theoretical and experimental values of the specific heat of superconductors (Physica **16**, 709-718, 1950).

The deviations mentioned in the title are explained qualitatively by the combination of two effects, the more important of which is based on the assumption that the thickness of Heisenberg's superconducting layer, covering part of the Fermi surface, does not drop abruptly to zero at the boundary of the layer. The other effect, which was more obvious but appeared to be of less importance, comes into play when the thickness of the superconducting layer decreases with increasing temperature.

**1963:** J. de Jonge, R. J. H. Alink and R. Dijkstra: Absorption spectrum and photo-decomposition of o-hydroxybenzene diazonium sulphate (Rec. Trav. chim. Pays-Bas **69**, 1448-1454, 1950).

The long-wave absorption of o-hydroxybenzene-diazonium salts in water is found at 3950 Å, but in 50% sulphuric acid as solvent it appears at 3500 Å. While the photo-decomposition of o-hydroxybenzene diazonium sulphate in water leads quantitatively to a derivative of cyclopentadiene, decompostion of the same compound in 50% sulphuric acid results for the greater part in catechol. The effect of sulphuric acid, both on the spectrum and the photo-decomposition, may be due to an increase in energy of the quinoid resonance structure of the hydroxybenzene-diazonium ion by the acid.

**1964:** D. Polder: Ferrite materials (Proc. Inst. elect. Engrs **97** II, 246-256, 1950).

Survey of electrical and magnetical behaviour of ferrite materials in connection with physico-chemical properties. The manufacture is dealt with and important possibilities for application of various ferrite materials are briefly reviewed. Some of the experimental methods are described for measuring the chemical and magnetical properties in various frequency ranges.

**1965*:** J. M. Stevels: Expérience et théories sur la facteur de puissance des verres en fonction de leur composition (Verres et Réfractaires **4**, 83-89, 1950). (Experiments and theory on power factor of glasses as a function of their composition; in French)

Translation of **R 127.**

**1966:** J. I. de Jong: Note on the reaction of urea and formaldehyde (Rec. Trav. chim. Pays-Bas **69**, 1568, 1950).

In contradistinction with other investigators a purely bimolecular reaction without an initial fall in $p_H$ is found when neutral formol is added to a freshly prepared solution of repeatedly crystallized urea. The decrease in $p_H$ and high initial reaction rate found by others may be due to small amounts of ammonium salts.

**1967:** J. D. Fast: Le vieillissement du fer et de l'acier. (Revue metallurgie **47**, 779-876, 1950) (Ageing of iron and steel; in French).

The normal commercial steels, such as are used on a large scale for constructing-purposes, show ageing phenomena both after rapid cooling from high temperatures (quench ageing) and after mechanical deformation at room temperature (strain ageing). There is a gradual, but unfavourable changing of the mechanical properties of the steel as, for instance, its hardness. The paper contains a survey of experiments carried out by the writer to investigate the above-mentioned quenching phenomena, starting from pure iron to which known small quantities of carbon, nitrogen or oxygen have been added. Also the influence of manganese on the ageing of $\alpha$-iron has been studied. It appears that oxygen being practically insoluble in $\alpha$-iron causes virtually no quench ageing; carbon and nitrogen both cause considerable quench ageing. This is ascribed to precipitation of these elements in the form of carbide and nitride respectively. Manganese appears to be capable of suppressing the quench ageing of iron containing nitrogen but not of iron containing carbon. It is assumed that in the crystal lattice the nitrogen atoms will preferably be in the vicinity of manganese atoms. The formation of iron nitride will take place at a perceptible rate only at high temperatures because the binding energy of nitrogen is fairly great there. As regards strain ageing the author's experiments prove that nitrogen is the main agent which causes this kind of ageing. A nitrogen content of $10^{-3}$ % gives already the maximum strain ageing even in the presence of manganese. Carbon does not cause strain ageing at room temperature, but at higher temperatures (100 °C) this elements also causes an increase of hardness in iron after deformation. The difference between C and N is ascribed to a difference in solubility.

**1968*:** Balth. van der Pol and H. Bremmer: Operational calculus based on the two-sided Laplace integral (Cambridge Univ. Press 1950, XIII + 424 pp., 97 fig.).

After a short introduction concerning the person of Oliver Heaviside and his "operational calculus" the Laplace transformation is dealt with at length. Apart from some historical remarks on the one-sided transformation (integration between 0 and $\infty$) the two-sided Laplace integral ($-\infty$ to $+\infty$) is introduced ab initio. The Fourier integral is dealt with as an introduction to the Laplace transformation. A number of simple examples and elementary rules are given after which the very essential conceptions of unit function and delta function are introduced. Two chapters are devoted to convergence questions and asymptotic expressions. Then follows among others the treatment of linear differential equations with constant coefficients, idem with variable coefficients, step functions and saw-tooth functions, differential equations and integral equations. Partial differential equations are dealt with both from the point of the Laplace transformation with one variable and from that of the "simultaneous" transformation in many variables. The special advantages of the operational calculus come to the fore in deducing unexpected and complex relations between functions and in dealing with transient phenomena in linear networks (filters).

**1969:** G. W. Rathenau and J. L. Meijering: Rapid oxydation of metals and alloys in the presence of ($MoO_3$) (Metallurgia **42**, 162-172, 1950).

The influence of $MoO_3$ on the oxidation in air of copper, aluminium bronze, silver, aluminium-silver, nickel, nichrome and two heat-resisting steels is measured as a function of temperature. Discontinuities in the curves are correlated with eutectic temperatures measured in mixtures of $MoO_3$ and oxides of the metals concerned. Metallographic examination of the oxidized specimens reveals some peculiar effects. Besides by the formation of liquid oxidic phases, the oxidation can be enhanced by $MoO_3$ in two other ways: incorporation of Mo in the crystal lattice of the oxide skin, and thermodynamic stabilization of $Ag_2O$ by formation of molybdates.

VOL. 7 NR 3 JUNE 1952

# Philips Research Reports

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

*R 189* *Res. Rep. 7, 161-168, 1952*

## SOME EXPERIMENTS AND THEORIES ON THE POWER FACTOR OF GLASSES AS A FUNCTION OF THEIR COMPOSITION, III

by J. M. STEVELS 621.317.374:621.315.612.6

**Summary**

It is shown that the power factor of borosilicate glasses (measured for greatly differing frequencies) is minimal on the straight line separating the accumulation region and the destruction region, 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.

**Résumé**

Il s'avère que le facteur de puissance de verres au borosilicate (mesuré à des fréquences très différentes) est minimum sur la droite de démarcation de la région d'accumulation et de la région de destruction, ce qui prouve que sur cette droite la rigidité du réseau est maximum. Ce maximum est moins prononcé pour les fortes concentrations de $SiO_2$ et disparaît même dans un verre de composition suivante: $SiO_2$ 75%-$B_2O_3$ 20,5%-$M_2O$ 4,5% (M = Na ou Li). Les résultats expérimentaux sont discutés du point de vue théorique.

**Zusammenfassung**

Es wird gezeigt, daß der (für sehr verschiedene Frequenzen gemessene) Verlustsfaktor von Borsilikatgläsern seinen Mindestwert auf der Geraden hat, die das Akkumulationsgebiet und das Destruktionsgebiet voneinander trennt, womit bewiesen wird, daß auf dieser Linie die Steifigkeit des Netzwerkes maximal ist. Dieses Maximum ist weniger ausgeprägt für die höheren $SiO_2$-Konzentrationen und verschwindet bei einer Zusammensetzung von 75% $SiO_2$-20,5% $B_2O_3$-4,5% $M_2O$ (M = Na oder Li). Die experimentellen Ergebnisse werden theoretisch ausgewertet.

### 1. Introduction

A few years ago the author studied the vitreous systems $M_2O$-$B_2O_3$-$SiO_2$ (M stands for Li, Na, K, etc.). Combining the results of density measurements with those obtained from X-ray analysis by Warren and co-workers [1]), it has been possible to show that the binary systems $M_2O$-$B_2O_3$ may be

divided into two regions, viz. the accumulation region *A.R.* (for $R = 1{\cdot}50\text{-}1{\cdot}61$, in which $R$ is the total amount of oxygen ions divided by the amount of boron ions) and the destruction region *D.R.* for $R > 1{\cdot}61$ [2]). In fig. 1, which represents all compositions of the system $M_2O\text{-}B_2O_3\text{-}SiO_2$ in the usual way, the base representing all $M_2O\text{-}B_2O_3$ systems is divided into two parts, with the transition point lying at a composition of 0·18 $M_2O$-0·82 $B_2O_3$ *).

In the *A.R.* the excess of oxygen ions is taken up by a change of the co-ordination number of the boron ions, so that no non-bridging oxygen ions are formed, whereas in the *D.R.* the excess of oxygen ions is taken up by the formation of non-bridging oxygen ions [1]).

This means that on adding increasing amounts of $M_2O$ to $B_2O_3$ the network becomes more and more rigid and after passing the transition from *A.R.* to *D.R.* it becomes weaker again. This is reflected in the trend of various physical properties, such as the coefficient of expansion, which passes through a minimum [3]).

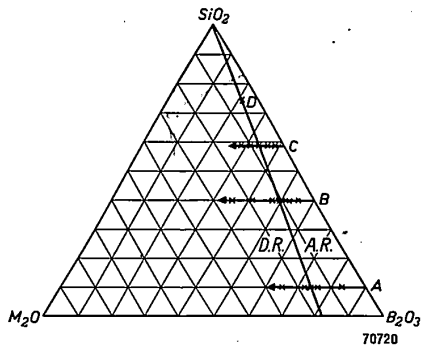



Fig. 1. The system $M_2O\text{-}B_2O_3\text{-}SiO_2$ showing the position of the accumulation region and the destruction region and the way in which the composition of the glasses investigated is varied.

When extending the idea of the two regions to the system $M_2O\text{-}B_2O_3\text{-}SiO_2$ it has already been made plausible [2]) that the transition from *A.R.* to *D.R.* must take place along a given straight line (the transition line), which is also given in fig. 1.

One might think that this transition line has the same significance as that of the point mentioned above and that it represents a line of maximum strength of the network. This has been tested by power-factor measurements. In this paper it is shown that this is true for a great part of the transition line, but when one studies the composition of about 75% $SiO_2$ this no longer holds.

*) All compositions in this paper are in mole percentages.

## 2. Experimental

Since the measurement of the dielectric power factor tan $\delta$ gives valuable information about the structure of glasses [4]), the author studied the trend of this physical property at two different frequencies when varying the composition, as given by the arrows *A*-*C* in fig. 1.

Series *A*

| | mole percentages | | | tan$\delta$ × $10^4$ for | |
|---|---|---|---|---|---|
| | $SiO_2$ | $B_2O_3$ | $M_2O$ | M = Li | M = Na |
| 1 | 10 | 83 | 7 | 97·5 | 73·5 |
| 2 | 10 | 76 | 14 | 70·9 | 43·0 |
| 3* | 10 | 74 | 16 | 28·8 | 20·2 |
| 4 | 10 | 72 | 18 | 52·0 | 34·8 |
| 5 | 10 | 65 | 25 | 112 | 66·2 |

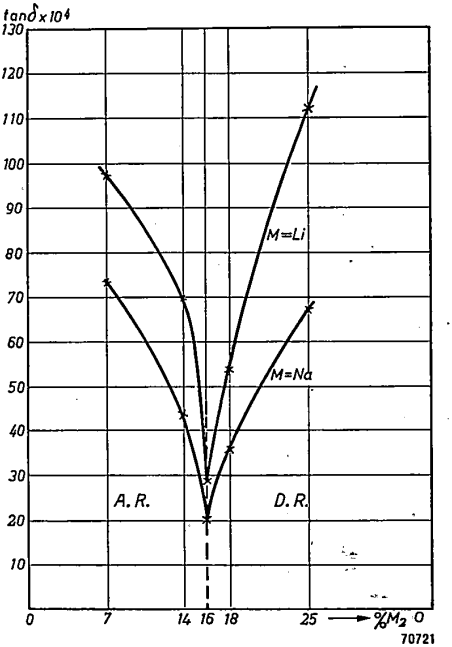



Fig. 2. The relation between the power factor tan$\delta$ (for $f = 1{\cdot}5 \times 10^6$ c/s and 20 °C) and the composition of glasses of series *A* containing 10 mole % $SiO_2$, 90 mole % ($B_2O_3 + M_2O$) for the cases M = Na and M = Li.

Series *B*

| | mole percentages | | | $\tan\delta \times 10^4$ for | |
|---|---|---|---|---|---|
| | $SiO_2$ | $B_2O_3$ | $M_2O$ | M=Li | M=Na |
| 1 | 40 | 55 | 5 | 111 | 95·1 |
| 2 | 40 | 52 | 8 | — | 61·4 |
| 3 | 40 | 51 | 9 | 76·9 | 45·0 |
| 4 | 40 | 50 | 10 | — | 28·4 |
| 5* | 40 | 49 | 11 | 54·0 | 22·8 |
| 6 | 40 | 47 | 13 | 71·2 | 29·3 |
| 7 | 40 | 41 | 19 | 102 | 67·6 |
| 8 | 40 | 35 | 25 | 119 | 96·9 |

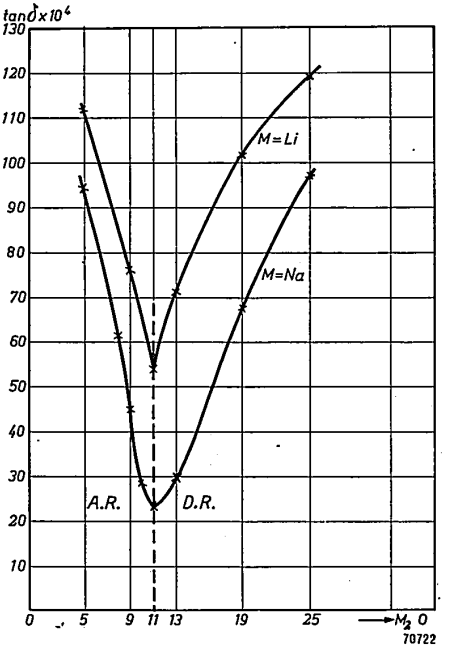



Fig. 3. The relation between the power factor $\tan\delta$ (for $f = 1{\cdot}5 \times 10^6$ c/s and 20 °C) and the composition of glasses of series *B* containing 40 mole % $SiO_2$, 60 mole % ($B_2O_3 + M_2O$) for the cases M = Na and M = Li.

Series *C*

| | mole percentages | | | $\tan\delta \times 10^4$ for | |
|---|---|---|---|---|---|
| | $SiO_2$ | $B_2O_3$ | $M_2O$ | M = Li | M = Na |
| 1 | 59 | 38·5 | 2·5 | 47·1 | 41·2 |
| 2 | 59 | 36 | 5 | 46·3 | 41·1 |
| 3* | 59 | 33·5 | 7·5 | 40·0 | 35·0 |
| 4 | 59 | 31 | 10 | 85·2 | 77·3 |
| 5 | 59 | 28·5 | 12·5 | 120 | 101 |

It is easy to find the compositions for which the transition line and the lines *A-C* intersect. These points (in the tables marked with an asterisk) are included in the series *A*, *B*, *C*, in which the various compositions and the corresponding power factors $\tan\delta$ at the frequency $f = 1{\cdot}5 \times 10^6$ c/s and for 20 °C are given.

The results for the frequency $f = 1{\cdot}5 \times 10^6$ c/s are also visualized in figs 2-4.

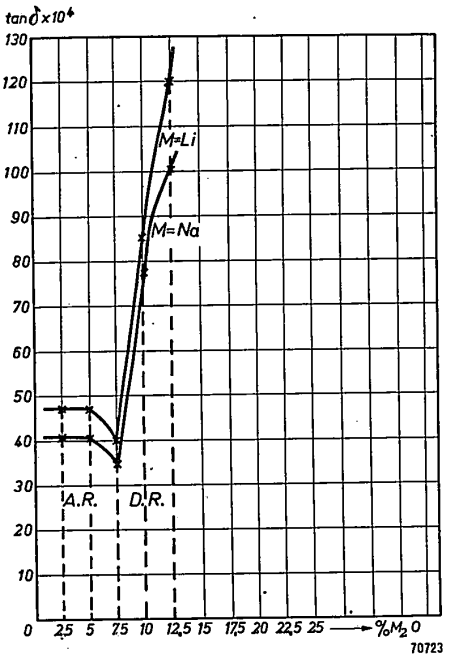



Fig. 4. The relation between the power factor $\tan\delta$ (for $f = 1{\cdot}5 \times 10^6$ c/s and 20 °C) and the composition of glasses of series *C* containing 59 mole % $SiO_2$ and 41 mole % ($B_2O_3$ + $M_2O$) for the cases M = Na and M = Li.

### 3. Discussion of the results

It is interesting to note that in all cases the power factor for all glasses containing lithium is higher than that for the corresponding sodium glasses. This may be easily understood because of the fact that at the given frequency the power factor is practically determined by the relaxation losses and, since the $Li^+$ ion in this respect is more mobile than the $Na^+$ ion [4]), the power factor is higher in all cases. Series *A* shows a very sharp minimum of the power factor for the concentration on the transition line. The minimum is somewhat less pronounced in series *B*, and diminishes considerably in series *C*.

The peculiar shape of the curves may be readily understood. Going from low $M_2O$ concentrations to higher $M_2O$ concentrations, the network in the *A.R.* becomes more and more rigid. Therefore the mobility of the $M^+$ ions is reduced and the power factor decreases in spite of the increasing concentration of $M^+$ ions. Having passed into the *D.R.*, the network breaks down more and more and for this reason, together with the increasing concentrations of mobile ions ($Na^+$ or $Li^+$), the power factor rises again.

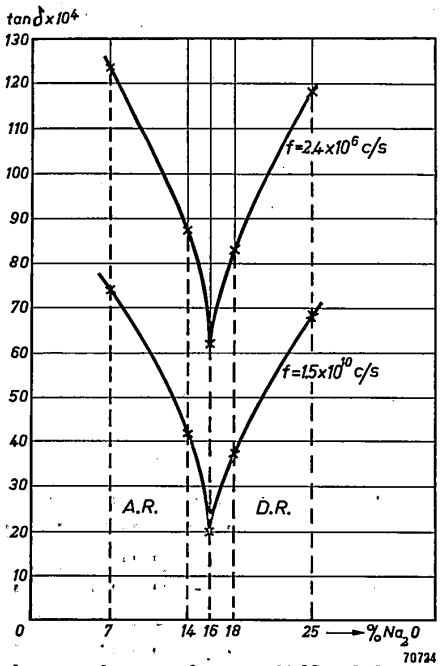



Fig. 5. The relation between the power factor at 20 °C and the composition of glasses of series *A* containing 10 mole % $SiO_2$ and 90 mole % ($B_2O_3$ + $Na_2O$) for two different frequencies, viz. $f = 1{\cdot}5 \times 10^6$ c/s and $f = 2{\cdot}4 \times 10^{10}$ c/s.

Figs 5 and 6 show the values of the power factor at the frequency $f = 2{\cdot}4 \times 10^{10}$ c/s for series $A$ and $B$ in the case of the sodium glasses. (As a comparison the values for $f = 1{\cdot}5 \times 10^{6}$ c/s (cf. figs 2 and 3) are added.)

For the frequency $f = 2{\cdot}4 \times 10^{10}$ c/s the power factor is due to the vibration losses [4]) *). The trend of the curves may be understood from the fact that for the intersection point on the transition line the strength of the network is maximum, the bonding of the $Na^+$ atoms is maximal,

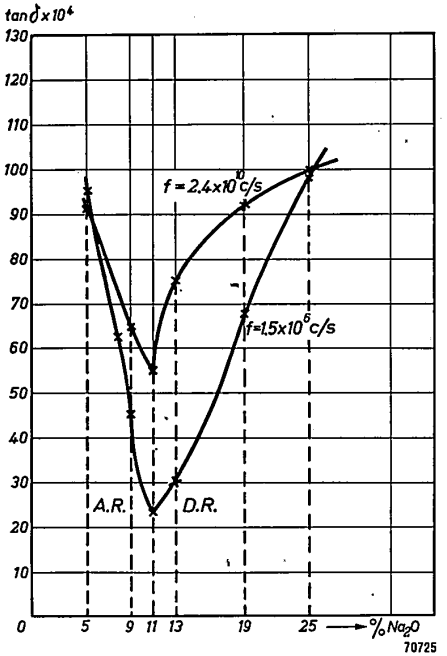



Fig. 6. The relation between the power factor at 20 °C and the composition of glasses of series $B$ containing 40 mole % $SiO_2$, 60 mole % ($B_2O_3$ + $Na_2O$) for two different frequencies, viz. $f = 1{\cdot}5 \times 10^{6}$ c/s and $f = 2{\cdot}4 \times 10^{10}$ c/s.

so that the resonance peak of the losses is found at very high frequencies (of the order of $f = 10^{13}$ c/s). The forerunner of this peak has already dropped to a rather low value at the frequency $f = 2{\cdot}4 \times 10^{10}$ c/s. For the compositions to the right and to the left of the transition line the peak has shifted to lower frequencies, that is to say nearer to the applied frequency $f = 2{\cdot}4 \times 10^{10}$ c/s and, therefore, the power factor is found to be higher.

*) The measurements at the frequency $f = 2{\cdot}4 \times 10^{10}$ c/s have been carried out by Mr H. G. Beljers of this laboratory.

The measurements (taken at greatly differing frequencies) furnish evidence for the fact that the transition line in the system $M_2O$-$B_2O_3$-$SiO_2$ is actually given by the straight line, at least in the cases where M = Li and M = Na.

For the glasses with relatively high $SiO_2$ content (at least 65%) the transition line is no longer a line of maximum strength.

It may be pointed out that recently similar measurements as described above have been taken at a "level" of 71% $SiO_2$ and of 75% $SiO_2$. In the former case a very small minimum still exists, but in the latter case a small maximum is found. This is due to the fact that in the region of about 75% $SiO_2$, 20·5% $B_2O_3$ and 4·5% $Na_2O$ (composition *D* on the transition line, see fig. 1) the glass structure — for some reason or other — is extra loose (causing a relative increase of the power factor). We have studied this subject in detail, taking into consideration various physical properties, and shall report on it in due course.

It is interesting to note that this peculiar region is the same as that where a separation into two phases, as a part of the well-known Vycor process, is very easy. It lies at hand to believe that these two facts are related to each other. The conclusion of this paper must, therefore, be confined to glasses with less than 65% $SiO_2$.

*Eindhoven, January 1952*

REFERENCES

1) J. Biscoe and B. E. Warren, J. Am. Ceram. Soc. **21**, 287-293, 1938.
2) J. M. Stevels, J. Soc. Glass Tech. **30**, 173-191, 1946.
3) J. M. Stevels, Ned. T. Natuurkunde **14**, 143-152, 1948.
4) J. M. Stevels, J. Soc. Glass Tech. **34**, 80-100, 1950; Philips Res. Rep. **5**, 23-36 1950; **6**, 34-53, 1951; Philips Tech. Rev., to be published shortly.

*Philips Res. Rep.* 7, 169-188, 1952

# ON THE SCATTERING MATRIX OF SYMMETRICAL WAVEGUIDE JUNCTIONS *)

by A. E. PANNENBORG

621.392.26:621.317.3

## CHAPTER III. SYMMETRY ANALYSIS OF WAVEGUIDE JUNCTIONS

If a waveguide junction exhibits a certain degree of symmetry, the number of constants needed to describe its behaviour as a transducer is less than for the general non-symmetrical junction with the same number of output leads.

A method to perform this reduction of constants has first been given, for lossless structures, by Dicke [1]). The scattering matrix of a lossless junction is unitary and can therefore always be brought into diagonal form by a suitable unitary similarity transformation. Dicke's method, accordingly, is one suitable to eigenvalue problems; he developed it largely by treating a number of specific examples.

In order to include dissipative structures, Kerns [8]) has lately put the theory on a more rigorous mathematical basis by extensive use of concepts derived from group theory. It is hardly necessary to point out that the methods used are for a large part identical with those developed for the analysis of symmetry problems in molecular physics [9]) and quantum mechanics [10]),[11]).

In this chapter we shall give an outline of Kerns' theory in conjunction with some aspects elaborated by Dicke. It is supposed that the reader is acquainted with the fundamental results of the theory of finite groups.

### III. 1. Representation of a symmetry group

From the vector notation of Maxwell's equations (I. 1) it is clear that their validity is not restricted to a particular choice of coordinate system. If a new coordinate system is introduced by a real linear transformation from the original system and the field components are transformed accordingly, the latter will again satisfy Maxwell's equations. Mathematically we can say that Maxwell's equations are invariant with respect to the real linear group. A particular subgroup of this linear group is the three-dimensional rotation reflection group, which in its turn comprises the point-symmetry groups (also called the crystallographic groups).

*) Continued from Philips Res. Rep. 7, 131-157, 1952.

The invariance of Maxwell's equations, though important, is in itself not very helpful. A new situation arises, however, if in a specific problem there can be found a transformation subject to which the boundary conditions also are invariant. Such a transformation is termed a covering operation for the structure under consideration. Because we deal only with structures of finite size, a translation cannot be a covering operation. The only possibilities left are reflections in a plane or a point and rotations about an axis. All possible covering operations of a structure, including the identity operator, constitute the symmetry group of the structure.

Let us consider now the electromagnetic field inside a waveguide junction. A covering operation will change the spatial position of the junction and the field, without altering their relative position. But as the junction remains by definition invariant under the operation, we may just as well assume the junction to be fixed in space, whilst the operation affects only the field within it.

In the previous chapters we have discussed how the electromagnetic behaviour of a junction can be described by the amplitudes of waves in the output leads. Once the amplitudes of the incoming waves, represented by the column matrix $\boldsymbol{A}$, are prescribed, the amplitudes $\boldsymbol{B}$ of the emergent waves are determined by the properties of the junction as expressed in terms of the scattering matrix. In Chap. I it has been explained that the elements $a_n$ and $b_n$ of $\boldsymbol{A}$ and $\boldsymbol{B}$ respectively are to be regarded as coefficients for the transverse electric field $E_{tn}$ in the various output leads of the junction. In other words we may state that the fields $E_{tn}$ constitute the coordinate system or basis for the description of the electromagnetic field; $\boldsymbol{A}$ comprises the coordinates of a possible incident electromagnetic field relative to this coordinate system and can as such be regarded as a vector.

If a covering operation $P$ of the symmetry group of the structure is applied to $\boldsymbol{A}$, we obtain another possible field $P\boldsymbol{A}$. The coefficient $a_n$, describing before the operation the field in lead $(n)$, will after the operation pertain to the field in lead $(m)$. The normalizing condition, eq. (I. 7), requires, that the basis fields $E_{tn}$ and $E_{tm}$ at corresponding points for all leads $(n)$ and $(m)$ respectively which interchange under the covering operation, are connected by

$$|E_{tn}|^2 = |E_{tm}|^2 . \qquad (1)$$

It should be noted that the one-mode assumption, condition (v), Sec. I.1, is essential here. The origin of the time scale can always be chosen so that $E_{tn}$ is real for all $n$. Then there are only two possible solutions of eq. (1), viz.

$$E_{tn} = \pm E_{tm} . \qquad (2)$$

It follows from eq. (2), that the matrix $\boldsymbol{D}(P)$, which expresses the vector $P\boldsymbol{A}$ in terms of $\boldsymbol{A}$, consists only of elements 0, +1 and −1. In each row and column there is only one non-zero element. The matrix $\boldsymbol{D}(P)$ is therefore real and orthogonal.

For all operators $P$ of the symmetry group the corresponding matrices $\boldsymbol{D}(P)$ can be formed. These matrices then constitute a representation of the group. The basis for this representation consists of the transverse electric fields in the output leads. It is evident that the dimension of the representation equals the number of output leads of the junction.

### III. 2. Relation between the symmetry of a junction and its scattering matrix

As our aim is the simplification of the scattering matrix of a junction, it is well to consider the connection between this quantity and the representation of the symmetry group.

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ specify a possible solution of Maxwell's equations for the junction; then

$$\boldsymbol{B} = \boldsymbol{S}\boldsymbol{A}. \tag{3}$$

By definition the fields

$$\boldsymbol{A}' = \boldsymbol{D}(P)\boldsymbol{A}, \qquad \boldsymbol{B}' = \boldsymbol{D}(P)\boldsymbol{B}, \tag{4}$$

obtained by a covering operation from $\boldsymbol{A}$ and $\boldsymbol{B}$, constitute also a possible solution of Maxwell's equations. Hence

$$\boldsymbol{B}' = \boldsymbol{S}\boldsymbol{A}'. \tag{5}$$

Inserting eq. (4) in eq. (3) we have

$$\boldsymbol{D}(P)\boldsymbol{B} = \boldsymbol{S}\boldsymbol{D}(P)\boldsymbol{A}, \tag{6}$$

or by virtue of eq. (3)

$$\boldsymbol{D}(P)\boldsymbol{S}\boldsymbol{A} = \boldsymbol{S}\boldsymbol{D}(P)\boldsymbol{A}. \tag{7}$$

Equation (7) is valid for any $\boldsymbol{A}$, so that for any $P$

$$\boldsymbol{D}(P)\boldsymbol{S} = \boldsymbol{S}\boldsymbol{D}(P). \tag{8}$$

Equation (8) tells us that the scattering matrix commutes with all matrices $\boldsymbol{D}(P)$ of the representation of the symmetry group. This relation is the key to the solution of our problem.

The choice of the transverse electric fields in the output leads as a basis follows in a natural way from the physical properties of the structure and is in this respect the simplest possible. However, this does not imply that it is, mathematically, the most logical choice for the analysis of eq. (8) and in general it will not be so. If then we want to analyse the influence of the symmetry properties of a junction on its electromagnetic behaviour, it will be wise to introduce a new basis, in which the symmetry operators appear in their most clear-cut form.

A change of basis can be achieved by the linear transformation

$$\boldsymbol{A} = \boldsymbol{T}\boldsymbol{A}', \quad \boldsymbol{B} = \boldsymbol{T}\boldsymbol{B}', \tag{9}$$

where the prime now denotes the quantities with respect to the new coordinate system. The matrices $\boldsymbol{D}'(P)$ representing the symmetry operators in the new basis are given by the similarity transformation

$$\boldsymbol{D}'(P) = \boldsymbol{T}^{-1}\,\boldsymbol{D}(P)\boldsymbol{T}. \tag{10}$$

Likewise the scattering matrix in the new basis becomes

$$\boldsymbol{S}' = \boldsymbol{T}^{-1}\boldsymbol{S}\boldsymbol{T}. \tag{11}$$

The quantity $\boldsymbol{S}'$ is a scattering matrix in so far that it relates outgoing waves to incident waves. Its physical meaning cannot always be visualized easily.

The specific transformation matrix $\boldsymbol{T}$ that is most useful here is the one for which $\boldsymbol{D}'(P)$ is a completely reduced representation. The matrices $\boldsymbol{D}'(P)$ will then consist of a set of square submatrices on the principal diagonal, all other submatrices being zero. Thus

$$\boldsymbol{D}'(P) = \begin{bmatrix} \boldsymbol{D}_1(P) & 0 & 0 & \cdots \\ 0 & \boldsymbol{D}_2(P) & 0 & \cdots \\ 0 & 0 & \boldsymbol{D}_3(P) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{12}$$

The dimension of each submatrix in eq. (12) is independent of $P$.

It is clear that in this specific basis, called symmetry basis, the coordinates $\boldsymbol{A}'$ fall apart in a number of subsets, each of which is invariant with respect to all operations $P$. The number of elements of any subset equals the dimension of the corresponding irreducible representation.

As $\boldsymbol{D}(P)$ and $\boldsymbol{S}$ are subject to the same similarity transformation, eq. (8) is equally valid in the new coordinate system as in the original basis, so that

$$\boldsymbol{D}'(P)\boldsymbol{S}' = \boldsymbol{S}'\boldsymbol{D}'(P), \tag{13}$$

where $\boldsymbol{D}'(P)$ is given by eq. (12). To evaluate the matrix products in eq. (13) we write

$$\boldsymbol{S}' = \begin{bmatrix} S'_{11} & S'_{12} & S'_{13} & \cdots \\ S'_{21} & S'_{22} & S'_{23} & \cdots \\ S'_{31} & S'_{32} & S'_{33} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}. \tag{14}$$

Here $S'$ has been split up into submatrices by identical dividing lines as have been used in eq. (12). The submatrices on the diagonal will therefore be square; the off-diagonal submatrices are in general rectangular, as the various irreducible representations may have different dimensions. Combining eqs (12)-(14) we obtain

$$\begin{bmatrix} D_1S'_{11} & D_1S'_{12} & D_1S'_{13}\cdots \\ D_2S'_{21} & D_2S'_{22} & D_2S'_{23}\cdots \\ D_3S'_{31} & D_3S'_{32} & D_3S'_{33}\cdots \\ \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} S'_{11}D_1 & S'_{12}D_2 & S'_{13}D_3\cdots \\ S'_{21}D_1 & S'_{22}D_2 & S'_{23}D_3\cdots \\ S'_{31}D_1 & S'_{32}D_2 & S'_{33}D_3\cdots \\ \vdots & \vdots & \vdots \end{bmatrix}. \quad (15)$$

Though for simplicity it has not been indicated explicitly, eq. (15) is valid for all $P$. The separate equations contained in eq. (15) all have the form

$$D_kS'_{kl} = S'_{kl}D_l. \quad (16)$$

Because the matrices $D'(P)$ are completely reduced, Schur's Lemma applies to eq. (16). This fundamental theorem states that a matrix $S'_{kl}$ satisfying eq. (16) for all $P$ is either (i) a multiple of the unit matrix, if $D_k$ and $D_l$ are the same irreducible representations, or (ii) identically zero, if the irreducible representations $D_k$ and $D_l$ are non-equivalent. To bring out most clearly the consequences of this theorem we rearrange the irreducible representations in eq. (12) so that

$$D'(P) = \left[\begin{array}{ccc|cc|c} D_1 & & & 0 & & \\ & D_1 & & & 0 & \\ & & D_1 & & & \\ \hline 0 & & & D_2 & & 0 \\ & & & & D_2 & \\ \hline & 0 & & 0 & & D_3 \\ & & & & & \ddots \end{array}\right], \quad (17)$$

where the subscript now denotes the species of the irreducible representation. Symbolically we can write $D'(P)$ as the direct sum of its irreducible components, viz.

$$D'(P) = n_1D_1 + n_2D_2 + \ldots + n_iD_i + \ldots. \quad (18)$$

The coefficient $n_i$ denotes the number of times the irreducible representation $D_i$ occurs in $D'(P)$.

If the submatrices in $S'$ are ordered in the same fashion as in $D'(P)$, eq. (17), $S'$ becomes, by Schur's Lemma,

$$S' = \left[\begin{array}{cc|c} S_{11}(1) & S_{12}(1)\cdots & \\ S_{21}(1) & S_{22}(1)\cdots & 0 \\ \vdots & \ddots\ S_{n_1 n_1}(1) & \\ \hline & & S_{11}(2)\cdots \\ 0 & & \vdots\ \ddots\ S_{n_2 n_2}(2) \\ \hline & & S_{11}(3) \end{array}\right] \qquad (19)$$

It is seen that a considerable number of submatrices in $S'$ necessarily vanishes, while all non-zero submatrices $S_{kl}(i)$ are multiples of the unit matrix. Thus

$$S_{kl}(i) = s_{kl}(i)\, \boldsymbol{I}\,, \qquad (20)$$

where $s_{kl}(i)$ is a number. Hence, the number of constants needed to specify $S'$ is reduced, by symmetry reasons alone, from $N^2$ ($N$ is the number of output leads of the junction) to

$$n_1^2 + n_2^2 + \ldots + n_i^2 + \ldots . \qquad (21)$$

No use has been made so far of any of the properties of $S'$ previously derived, such as the requirements for reciprocity ($S = \tilde{S}$) or for the absence of losses ($\tilde{S}\, S^* = \boldsymbol{I}$). Fulfilment of additional conditions like these will in general still further reduce the above number.

Summarizing we can state that the division of the coordinates $A'$ of the symmetry basis into invariant subsets is less far-reaching with respect to the scattering matrix $S'$ than with respect to the reduced-out representation of the symmetry group. In fact the number of invariant subsets equals in the former case the number of *different* irreducible representations, whilst in the latter case it is equal to the *total number* of irreducible representations contained in the representation of the group.

The construction of the reduced-out representation, eq. (17), and of the matrix $T$ needed in the similarity transformations, eqs (10) and (11), will now be outlined very briefly.

Character tables for the crystallographic groups can be found in many texts [9], [11]). The number of times each irreducible representation appears in the specific representation of the group for the problem on hand can be calculated by a well-known theorem of group theory, involving the characters. The irreducible representations themselves have not been tabulated,

but they can in most cases easily be constructed as they are implicitly determined by their known characters. In this way the completely reduced representation $\boldsymbol{D}'(P)$ can be found.

Once both $\boldsymbol{D}'(P)$ and $\boldsymbol{D}(P)$ are known, eq. (10) may be used to obtain a set of linear equations for the individual elements of $\boldsymbol{T}$. In all cases where one irreducible representation occurs more than once, this set of equations will be found to be insufficient for the determination of all elements of $\boldsymbol{T}$. In fact only the division of the symmetry basis into subsets that are invariant with respect to $\boldsymbol{S}'$, eq. (19), is fixed, whereas the choice of the specific coordinates within each subset is arbitrary. This indeterminacy is brought out very elegantly in the method given by Eyring, Walter and Kimball [11]) for the calculation of $\boldsymbol{T}$, which method furthermore is much less tedious than the solution of eq. (10).

The limited freedom in the choice of $\boldsymbol{T}$ may be used with advantage in those cases where it is known a priori that the scattering matrix can be transformed to a diagonal matrix. This is e.g. always possible for a lossless junction the scattering matrix of which has been proved to be unitary. The symmetry basis can then be chosen in such a way that the elements of $\boldsymbol{S}'$, eq. (19), obey the equations

$$|s_{kk}(i)| = 1\,; \qquad s_{kl}(i) = 0, \qquad k \neq l\,. \tag{21}$$

for all $i$.

### III. 3. Field distribution in a symmetry plane

By the adoption of the scattering matrix for the description of waveguide junctions we have deliberately relinquished the aim to obtain any information on the field distribution inside the junction. One is forced to accept this for most practical junctions because of the usually unsurmountable mathematical difficulties encountered in the process of solving Maxwell's equations subject to the boundary conditions imposed by the structure.

The symmetry of a junction can, however, provide us with some knowledge about the field in the interior; in particular some information can be obtained about the distribution in a symmetry plane, point or axis of those fields that correspond to the individual coordinates of the symmetry basis. The insight gained in this respect, though scanty, sometimes enables us to predict qualitatively the effect of alterations in the structure that change the boundary conditions in the symmetry plane, point or axis.

We shall now work out in detail the properties of the field in a symmetry plane, referring the reader for the results in a symmetry point or axis to Dicke[1]).

As reflection in a plane, applied twice in succession, leaves everything unchanged, it constitutes together with the identity operator a group.

This group (elements $I$ and $F$) has two one-dimensional irreducible representations, which obviously are the trivial and the alternating representations. Thus we have

$$D_1(I) = 1\,, \qquad D_1(F) = 1\,; \tag{22}$$

$$D_2(I) = 1\,, \qquad D_2(F) = -1\,. \tag{23}$$

The representations eqs (22) and (23) correspond to the familiar concept of solutions even and odd respectively with respect to the symmetry plane. In fig. 3.1 these two different types of solution are illustrated for the electric field vector. Let the field vector on the left-hand side of the symmetry plane be resolved into a component $E_n$ normal to the plane and a component $E_t$ parallel to it. The corresponding components of the field vector on the right-hand side at a symmetrical location are marked with a prime. The following properties can be read from fig. 3.1.

Even solution, $D_1(F) = +1$:

$$E_t = E_t'\,, \qquad E_n = -E_n'\,. \tag{24}$$

Odd solution, $D_2(F) = -1$:

$$E_t = -E_t'\,, \qquad E_n = E_n'\,. \tag{25}$$

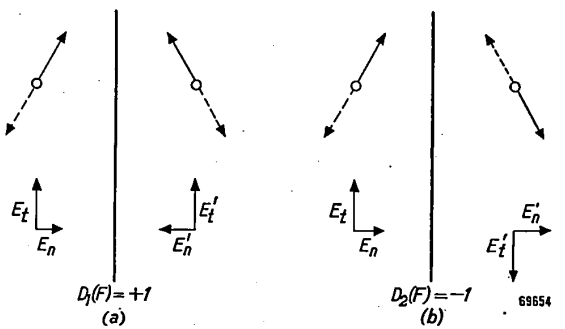



Fig. 3.1. Diagram illustrating an even electric field distribution ($a$) and an odd distribution ($b$).

The components of the magnetic field obey equations similar to eqs (24) and (25) but with opposite sign. It should be realized in this connection that the electric field vector is a polar vector, whereas the magnetic field vector is an axial vector (or pseudovector) [12]). Thus we have:

Even solutions, $D_1(F) = +1$:

$$H_t = -H_t'\,, \qquad H_n = H_n'\,. \tag{26}$$

IPR2025-00217 EXHIBIT 2010 Page 180 of 485

Odd solutions, $D_2(F) = -1$:

$$H_t = H_t' , \qquad H_n = -H_n' . \tag{27}$$

Equations (24) - (27) are valid for all pairs of points symmetrically located with respect to the symmetry plane, in particular also for the points lying in this plane. The field components with and without a prime become identical there. Hence the electromagnetic field must satisfy the following conditions in the symmetry plane.

Even solutions, $D_1(F) = +1$:

$$E_n = 0 , \qquad H_t = 0 . \tag{28}$$

Odd solutions, $D_2(F) = -1$:

$$E_t = 0 , \qquad H_n = 0 . \tag{29}$$

The requirements expressed in eq. (29) are identical with the boundary conditions on a perfect conductor. This means that the odd solution can be constructed with the aid of eqs (25) and (27) from the solution of Maxwell's equations for the space on one side only of the symmetry plane, which for this purpose is to be covered with a perfectly conducting layer. It may be said that the symmetry plane behaves in the odd solution as an electric wall. Analogously for the even solution the symmetry plane can be described as a magnetic wall, the properties of which are given by eq. (28).

The group $(I,F)$ will be a subgroup of any point-symmetry group in which the element $F$ occurs. In the one-dimensional irreducible representations of these groups the representation of $I$ and $F$ will necessarily be given either by eq. (22) or by eq. (23). Hence the results derived above apply also in this case.

## CHAPTER IV. JUNCTIONS OF FOUR WAVEGUIDES WITH TWO-FOLD PLANAR SYMMETRY

### IV. 1. The general case

In this and the following chapters we shall discuss junctions of four waveguides which have a structural symmetry that may be termed two-fold planar. The structures possess two mutually perpendicular symmetry planes; their intersection is a two-fold symmetry axis. The crystallographic notation for the group comprising these symmetry operators is $C_{2v}$.

Hollow waveguides used in microwave circuits have almost invariably a rectangular cross-section. The lowest mode they can support is the $TE_{10}$-mode. We may recall that the field of this mode has only one (transverse)

electric component, one transverse magnetic component and a longitudinal magnetic component. These components vary sinusoidally over the width of the guide and are independent of its height. In fig. 4.1 they are illustrated by arrows in the position of their respective maxima.

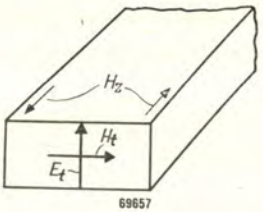



Fig. 4.1. Field components of the $TE_{10}$-mode in a rectangular waveguide.

A junction of rectangular waveguides displaying the required symmetry is shown in fig. 4.2. The transverse electric field components, which serve as a basis for the description, have been indicated by arrows. It is of course understood that the interior of the junction is consistent with the symmetry conditions. The two symmetry planes and the symmetry axis have been designated by the letters $F_1$, $F_2$ and $R$ respectively, which symbols will be used also to denote the corresponding symmetry operators.

The group table for the operators $I$, $F_1$, $F_2$, and $R$ is given in table 4.1. There are three non-trivial subgroups, viz. $(I, F_1)$, $(I, F_2)$ and $(I, R)$.

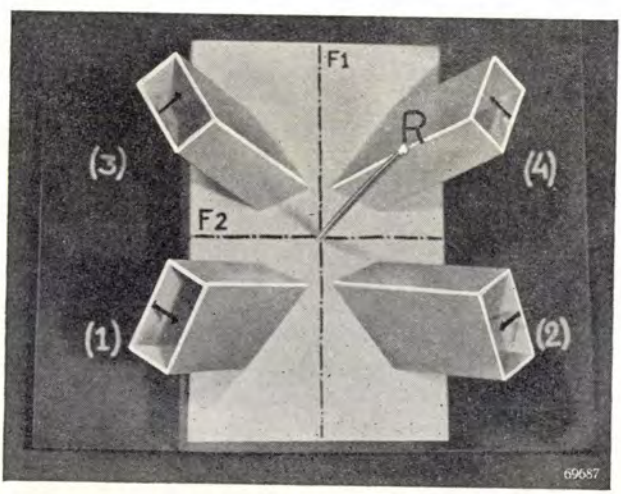



Fig. 4.2. Waveguide junction having two-fold planar symmetry.

Table 4.1 Group table

| | | | |
|---|---|---|---|
| $I$ | $F_1$ | $F_2$ | $R$ |
| $F_1$ | $I$ | $R$ | $F_2$ |
| $F_2$ | $R$ | $I$ | $F_1$ |
| $R$ | $F_2$ | $F_1$ | $I$ |

The matrices that indicate how the terminal fields interchange when subjected to a symmetry operator can be found by inspection from fig. 4.2. The operator $F_1$ e.g. permutes the fields in leads (1) and (2) and also the fields in leads (3) and (4). In this way we obtain the following representation of the group

$$\boldsymbol{D}(I)=\begin{pmatrix}1000\\0100\\0010\\0001\end{pmatrix},\ \boldsymbol{D}(F_1)=\begin{pmatrix}0100\\1000\\0001\\0010\end{pmatrix},\ \boldsymbol{D}(F_2)=\begin{pmatrix}0010\\0001\\1000\\0100\end{pmatrix},\ \boldsymbol{D}(R)=\begin{pmatrix}0001\\0010\\0100\\1000\end{pmatrix}. \quad (1)$$

Because the group is Abelian, each element forms a class by itself. The number of irreducible representations, which equals the number of classes, must hence be four. They are all one-dimensional and therefore identical with their characters.

It is to be noted that by a judicious choice of the numbering of the output leads and the direction of the terminal field eq. (1) can be identified with the regular representation of the group. In group theory it is shown that the regular representation contains all irreducible representations a number of times equal to their respective dimensions. The completely reduced form of the representation, eq. (1), must therefore consist of the four one-dimensional irreducible representations. The characters of the

Table 4.2 Character table

| | $I$ | $F_1$ | $F_2$ | $R$ |
|---|---|---|---|---|
| $D_1$ | 1 | 1 | 1 | 1 |
| $D_2$ | 1 | −1 | 1 | −1 |
| $D_3$ | 1 | 1 | −1 | −1 |
| $D_4$ | 1 | −1 | −1 | 1 |

group $C_{2v}$ are given in table 4.2. The reduced representation $\boldsymbol{D}'(P)$, which is given by

$$\boldsymbol{D}'(P)=\begin{pmatrix}D_1(P) & 0 & 0 & 0\\ 0 & D_2(P) & 0 & 0\\ 0 & 0 & D_3(P) & 0\\ 0 & 0 & 0 & D_4(P)\end{pmatrix}, \quad (2)$$

can be written down immediately from table 4.2. We have

$$\boldsymbol{D}'(I)=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix},\ \boldsymbol{D}'(F_1)=\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix},$$

$$\boldsymbol{D}'(F_2)=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix},\ \boldsymbol{D}'(R)=\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&1\end{pmatrix}. \quad (3)$$

Because the matrices $\boldsymbol{D}'(P)$ are diagonal, the scattering matrix transformed to the symmetry basis must also be diagonal. Hence let

$$\boldsymbol{S}'=\begin{pmatrix}s_1&0&0&0\\0&s_2&0&0\\0&0&s_3&0\\0&0&0&s_4\end{pmatrix}. \quad (4)$$

The matrix $\boldsymbol{S}'$ is purely diagonal as a consequence of symmetry only; in particular the validity of eq. (4) does not require that the junction should be lossless.

The transformation matrix $\boldsymbol{T}$ that connects $\boldsymbol{S}'$, eq. (4), with the scattering matrix $\boldsymbol{S}$ can in this specific problem be found in a very direct manner by the following considerations. As the representation, eq. (1), is regular, the symmetry basis for the field description is abstractly identical with the basis in which the group algebra appears as the sum of its invariant subalgebras. We can construct this basis by taking the sum over the group of the products of the complex conjugate of each element of the irreducible representations and the corresponding operator. It thus follows that the elements of $\boldsymbol{T}$ are proportional to the (real) elements of the irreducible representations, which are given in table 4.2. Contrary to the characters, which have been normalized to the order of the group, $\boldsymbol{T}$ is normalized to unity so that

$$\boldsymbol{T}=\tfrac{1}{2}\begin{pmatrix}1&1&1&1\\1&-1&1&-1\\1&1&-1&-1\\1&-1&-1&1\end{pmatrix}. \quad (5)$$

By the order in which the characters have been written down in table 4.2 $\boldsymbol{T}$ is a symmetrical matrix. Being orthogonal as well it has the useful property

$$\boldsymbol{T}=\tilde{\boldsymbol{T}}=\boldsymbol{T}^{-1}. \quad (6)$$

The calculation of $S$ is now a simple matter. By eq. (III.11) we have

$$S = T S' T^{-1}. \tag{7}$$

Substituting eqs (4) and (5) in eq. (7) we obtain after evaluating the matrix product

$$S = \begin{pmatrix} \alpha & \beta & \gamma & \delta \\ \beta & \alpha & \delta & \gamma \\ \gamma & \delta & \alpha & \beta \\ \delta & \gamma & \beta & \alpha \end{pmatrix}, \tag{8}$$

where

$$\left.\begin{aligned} \alpha &= \tfrac{1}{4}(s_1 + s_2 + s_3 + s_4), \\ \beta &= \tfrac{1}{4}(s_1 - s_2 + s_3 - s_4), \\ \gamma &= \tfrac{1}{4}(s_1 + s_2 - s_3 - s_4), \\ \delta &= \tfrac{1}{4}(s_1 - s_2 - s_3 + s_4). \end{aligned}\right\} \tag{9}$$

## IV. 2. The nature of the solution

Having completed the mathematical derivation we do well to investigate the physical meaning of the solution. The result that the scattering matrix $S$, eq. (8), contains only four different elements could have been deduced directly from inspection of fig. 4.2. The replacement of these four elements by four other unknown constants $s_n$ as given by eqs (9) is in itself trivial. It is possible, however, to obtain some information on the nature of these constants $s_n$. We note that they can be interpreted as reflection coefficients, because they are diagonal elements of the scattering matrix $S'$. Each coefficient $s_n$ belongs to one coordinate of the symmetry basis. The symmetry species of each coordinate can be found in the character table. Recalling the contents of Sec. III.3 we see, e.g., that the electromagnetic field which is invariant if operated on by the irreducible representation $D_2$ is odd with

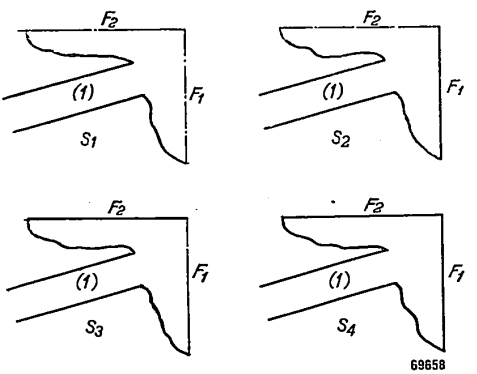



Fig. 4.3. Diagram illustrating the nature of the symmetry planes for the various symmetry coordinates; — · — · — = magnetic wall; ——— = electric wall.

respect to $F_1$ and even with respect to $F_2$. In other words, the quantity $s_2$ is the reflection coefficient that would appear in anyone of the output leads if an electric wall had been inserted in plane $F_1$ and a magnetic wall in plane $F_2$. Similar statements can be made for the other symmetry coordinates. The result is shown in diagrammatical form in fig. 4.3, where the quadrant of the junction containing lead (1) is depicted, in a cross-section perpendicular to the planes $F_1$ and $F_2$.

### IV. 3. The degenerate case

In the two preceding sections the only essential condition imposed on the structure was the two-fold planar symmetry. We shall now discuss a class of junctions which is often met in practical circuits. The geometry is particularly simple. The junction consists of two straight sections of rectangular waveguide which have one of their side walls in common. The two possible configurations, shown in figs 4.4 and 4.5, are clearly special cases of the more general junction shown in fig. 4.2. The common wall is

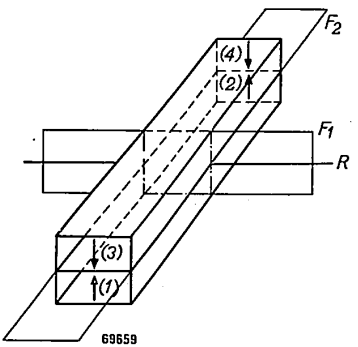



Fig. 4.4. Two sections of rectangular waveguide having one of their wide sides in common.

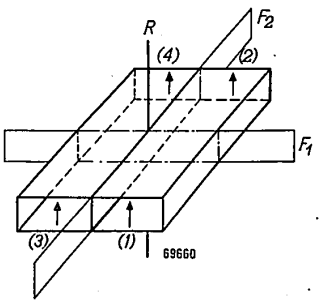



Fig. 4. 5.Two sections of rectangular waveguide having one of their narrow sides in common.

assumed to be infinitely thin; this condition is of course never fulfilled in practical structures, but it is often closely approximated. Only discontinuities lying wholly in the common wall are permitted. They can consist e.g. of holes or slots, possibly filled with some dielectric material. Their location and shape are arbitrary, but must be consistent with the symmetry requirements. This means simply that they must be symmetrical with respect to an axis $R$. Two possible configurations of this kind with circular holes in the common wall are shown in fig. 4.6.

It is possible to obtain more information on the junctions just described than on the more general types discussed in the previous sections, because one of the symmetry planes, viz. $F_2$, coincides with a side wall of the waveguides. A symmetry plane is known to behave with respect to the field distributions of the coordinates of the symmetry basis either as a perfect electric or a perfect magnetic conductor. The former case clearly imposes the same boundary conditions as the wall of an ideal waveguide. The effect of the discontinuities, lying in the symmetry plane, therefore vanishes for the solutions odd with respect to $F_2$; the junction then degenerates into two independent sections of waveguide. This happens for the irreducible representations $D_3$ and $D_4$, in which the operator $F_2$ is represented by $-1$,

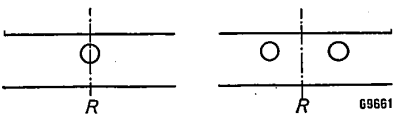



Fig. 4.6. Two examples of coupling holes in the common wall between two waveguides with the symmetry axis $R$.

as can be seen from table 4.2. The boundary conditions for the fields described by the reflection coefficients $s_3$ and $s_4$ are thus completely specified: an ideal waveguide terminated, for $s_3$, by a magnetic short-circuit ($D_3(F_1) = 1$) and, for $s_4$, by an electric short-circuit ($D_4(F_1) = -1$). To express the value of $s_3$ and $s_4$ numerically we must fix the reference planes; an obvious choice is the symmetry plane $F_1$ itself. With reference to condition (vi), Sec. I.1, it should be realized that in actual junctions the output leads should be made long enough so that higher modes excited by the discontinuities are sufficiently attenuated before reaching another discontinuity in one of the leads. As the definition of the scattering matrix is based on the transverse electric field, the reflection coefficient of an electric short-circuit is $-1$ if referred to the position of the short-circuit itself. Analogously for a magnetic short-circuit it equals $+1$. Hence

$$s_3 = +1\,, \qquad s_4 = -1\,. \tag{10}$$

It is thus seen that this special type of junction of four rectangular waveguides can be described by only two constants instead of the four needed for the more general structure with two-fold planar symmetry. Equations (10) inserted in eqs (9) yield

$$\left.\begin{aligned} \alpha &= \tfrac{1}{4}(s_1 + s_2) \quad , \\ \beta &= \tfrac{1}{4}(s_1 - s_2 + 2) \, , \\ \gamma &= \tfrac{1}{4}(s_1 + s_2) \quad , \\ \delta &= \tfrac{1}{4}(s_1 - s_2 - 2) \, . \end{aligned}\right\} \tag{11}$$

Between these four quantities there now exist two relations, viz.

$$\alpha = \gamma \quad \text{and} \quad \beta - \delta = 1 \, . \tag{12}$$

If the junction is lossless,

$$|s_n| = 1 \, . \tag{13}$$

The behaviour of the junction can in this case be specified completely by two real constants, preferably the phase angles of $s_1$ and $s_2$.

It is useful to determine the values of $s_1$ and $s_2$ in the limit of infinitely weak coupling between the two sections of waveguide. In the absence of discontinuities the junction simply consists of two sections of ideal waveguide. With the same arguments as used in the derivation of eqs (10) we can write immediately (cf. table 4.2)

$$\left.\begin{aligned} D_1(F_1) &= +1 \, , \qquad s_1 = +1 \, , \\ D_2(F_1) &= -1 \, , \qquad s_2 = -1 \, . \end{aligned}\right\} \tag{14}$$

As a check on the truth of this result we may note that insertion of eqs (14) in eqs (11) makes all elements vanish except $\beta$, which becomes equal to 1. As $\beta$ denotes the transmission coefficient between leads (1) and (2) and between leads (3) and (4), this result is seen to be correct.

If the junction is lossless and the coupling is weak, the first-order approximation will add small imaginary terms to the unperturbed values of $s_1$ and $s_2$, eqs (14), as follows from eq. (13). Then $\alpha$, $\gamma$ and $\delta$ will be small imaginary quantities. A method to calculate these quantities explicitly has been developed by Bethe in his well-known paper on the diffraction by small holes [13]) and in several reports [14]), [15]).

### IV. 4. The wave matrix

Close inspection of eqs (8) and (11) reveals that it is worth while to introduce a new basis in which the simple properties of the junction are particularly well illustrated. These new coordinates are proportional to the sum and the difference of the wave amplitudes in leads (1) and (3) and in leads (2) and (4). The matrix $\boldsymbol{U}$ that performs the transformation

from the original to the new coordinates is given by

$$\boldsymbol{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}. \tag{15}$$

It is seen to be orthogonal, symmetrical and normalized to unity. As a consequence all the general results derived in Chap. I will be valid for the scattering matrix as transformed to the new basis, which we will call two-mode basis (see below).

The incident waves in terms of the new coordinates are given by (cf. eq. (III.9))

$$\boldsymbol{A}_{tm} = \boldsymbol{U}^{-1}\boldsymbol{A} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} a_1 + a_3 \\ a_2 + a_4 \\ a_1 - a_3 \\ a_2 - a_4 \end{pmatrix}, \tag{16}$$

and analogously for the emergent waves $\boldsymbol{B}_{tm}$. The scattering matrix $\boldsymbol{S}_{tm}$ in the two-mode basis is connected to $\boldsymbol{S}$ by (cf. eq. (III. 11))

$$\boldsymbol{S}_{tm} = \boldsymbol{U}^{-1}\boldsymbol{S}\boldsymbol{U}. \tag{17}$$

Evaluating the matrix product we find with the aid of eqs (8) and (12)

$$\boldsymbol{S}_{tm} = \begin{pmatrix} \alpha + \gamma & \beta + \delta & 0 & 0 \\ \beta + \delta & \alpha + \gamma & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{18}$$

In terms of $s_1$ and $s_2$ as given by eq. (11) this becomes

$$\boldsymbol{S}_{tm} = \begin{pmatrix} \frac{1}{2}(s_1 + s_2) & \frac{1}{2}(s_1 - s_2) & 0 & 0 \\ \frac{1}{2}(s_1 - s_2) & \frac{1}{2}(s_1 + s_2) & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{19}$$

The matrix $\boldsymbol{S}_{tm}$ is seen to be decomposed. The new basis hence comprises two invariant subsets. The fields corresponding to the third and fourth coordinates are purely travelling waves as is evident from the submatrix on the lower right in eq. (19). Only the other subset, containing the first and second coordinates, is affected by the presence of the discontinuities. This suggests the following physical picture. Let us consider the junction as one section of waveguide which can support two modes of propagation instead of two sections of waveguide each capable of supporting only one mode. Each subset of the new coordinates corresponds to one mode. The fields of the first subset (the +mode) are proportional to the sum of

the original basis fields; the second subset (the —mode) deals with the difference of the original fields. The fields of these two modes are schematically shown in fig. 4.7. It is clear that the —mode is in no way affected by the presence or removal of the wall separating the two original guides. Discontinuities in this wall will influence only the + mode to an extent given by the submatrix in the upper left corner of $S_{tm}$, eq. (19). This submatrix is, indeed, seen to have the form that belongs to a symmetrical obstacle in a waveguide.

The introduction of the two-mode basis is especially useful in problems concerning junctions connected in cascade. Assume e.g. that the two waveguides are coupled by a number of holes spaced apart in the common wall, each of them satisfying the symmetry requirements. If then we want to express the combined effect in terms of the parameters of the individual discontinuities, the transformation to the two-mode basis reduces the problem to the treatment of two separate waveguides: one loaded by the discontinuities and the other completely unaffected by them [16]). The scattering matrix, which relates emergent to incident waves, is not suited for the description of junctions connected in cascade. For this purpose we want to express the wave amplitudes on one side of a junction in terms of the amplitudes on the other side. In our case this amounts to a relation between the amplitudes in guides (1) and (3) and the amplitudes in guides (2) and (4). The matrix describing the connection between these quantities has been termed the wave matrix [17]). The main point in the use of this matrix is that the wave matrix of a series of discontinuities equals simply the product of the wave matrices of the individual discontinuities taken in the proper order.

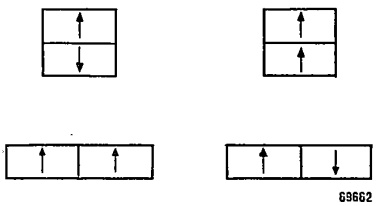



Fig. 4.7. Schematic representation of two modes of propagation.

We shall derive the wave matrix only in the two-mode basis. We first reproduce the scattering description explicitly. By eqs (16) and (19) we have

$$\frac{1}{\sqrt{2}}\begin{pmatrix} b_1+b_3 \\ b_2+b_4 \\ b_1-b_3 \\ b_2-b_4 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(s_1+s_2) & \frac{1}{2}(s_1-s_2) & 0 & 0 \\ \frac{1}{2}(s_1-s_2) & \frac{1}{2}(s_1+s_2) & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \frac{1}{\sqrt{2}}\begin{pmatrix} a_1+a_3 \\ a_2+a_4 \\ a_1-a_3 \\ a_2-a_4 \end{pmatrix}. \quad (20)$$

By an elementary calculation the above four linear equations can be rearranged so as to bring all wave amplitudes with subscripts 2 and 4 on the left-hand side. We find

$$\frac{1}{\sqrt{2}}\begin{pmatrix} b_2+b_4 \\ a_2+a_4 \\ b_2-b_4 \\ a_2-a_4 \end{pmatrix} = \begin{pmatrix} -\frac{2s_1s_2}{s_1-s_2} & \frac{s_1+s_2}{s_1-s_2} & 0 & 0 \\ -\frac{s_1+s_2}{s_1-s_2} & \frac{2}{s_1-s_2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \frac{1}{\sqrt{2}}\begin{pmatrix} a_1+a_3 \\ b_1+b_3 \\ a_1-a_3 \\ b_1-b_3 \end{pmatrix}. \tag{21}$$

Like the scattering matrix the wave matrix in eq. (21) is, of course, decomposed. The two elements 1 on the diagonal clearly demonstrate again that the —mode is unaffected by the discontinuities.

### IV. 5. Shift of the reference planes

If it is desirable to describe the behaviour of the junction with reference to planes other than the symmetry plane $F_2$, it should be realized that the new planes must all be equidistant from $F_2$ for reasons of symmetry. The matrix $\Phi$ needed in the transformation of the scattering matrix $\boldsymbol{S}$ (cf. Sec. I.4) therefore has the almost-trivial form

$$\Phi = e^{i\varphi}\boldsymbol{I}, \tag{22}$$

where $\varphi = \beta l$ is the phase angle corresponding to an outward shift over a distance $l$. Denoting with a prime the quantities with respect to the new reference planes we simply have

$$\boldsymbol{S}' = e^{2i\varphi}\boldsymbol{S}. \tag{23}$$

To investigate the required transformation for the wave matrix we rewrite eq. (21). We have

$$\boldsymbol{C}_{24} = \boldsymbol{W}\boldsymbol{C}_{13}, \tag{24}$$

where the column matrices of eq. (21) have been denoted by $\boldsymbol{C}_{24}$ and $\boldsymbol{C}_{13}$; $\boldsymbol{W}$ represent the wave matrix. The column matrices transform according to

$$\boldsymbol{C}'_{24} = \Psi\boldsymbol{C}_{24}, \qquad \boldsymbol{C}'_{13} = \Psi^{-1}\boldsymbol{C}_{13}, \tag{25}$$

where

$$\Psi = \begin{pmatrix} e^{i\varphi} & 0 & 0 & 0 \\ 0 & e^{-i\varphi} & 0 & 0 \\ 0 & 0 & e^{i\varphi} & 0 \\ 0 & 0 & 0 & e^{-i\varphi} \end{pmatrix}, \tag{26}$$

as follows from the properties of the individual elements of $\boldsymbol{C}_{24}$ and $\boldsymbol{C}_{13}$. Substitution of eq. (25) in eq. (24) yields

$$\boldsymbol{W}' = \Psi\boldsymbol{W}\Psi.$$

With the aid of eqs (21), (26) and (27) we can construct the wave matrix $W'$ with respect to the new terminal planes explicitly; it is found to be

$$W' = \begin{pmatrix} -\dfrac{2s_1's_2'}{s_1'-s_2'} & \dfrac{s_1'+s_2'}{s_1'-s_2'} & 0 & 0 \\ -\dfrac{s_1'+s_2'}{s_1'-s_2'} & \dfrac{2}{s_1'-s_2'} & 0 & 0 \\ 0 & 0 & e^{2i\varphi} & 0 \\ 0 & 0 & 0 & e^{-2i\varphi} \end{pmatrix}, \tag{28}$$

where

$$s_n' = s_n e^{2i\varphi},\ n = 1, 2\,. \tag{29}$$

*(To be continued)*

Philips Res. Rep. 7, 189-196, 1952

# INSTABILITY OF BELLOWS SUBJECTED TO INTERNAL PRESSURE *)

by J. A. HARINGX

539.371

**Summary**

Like thin-walled cylinders, dealt with in a previous paper, also bellows may become unstable when loaded by internal pressure. The critical value of this pressure, which in accordance with Euler's well-known formula is governed by the rigidity of the bellows with respect to bending, has been computed only for rectangularly shaped corrugations, and has been checked experimentally.

**Résumé**

Tout comme les cylindres à minces parois, dont il a été question dans un article précédent, les soufflets, soumis à une pression intérieure, peuvent devenir instables. La valeur critique de cette pression, qui est régie, suivant la formule bien connue de Euler, par la résistance des soufflets à la flexion, est calculée uniquement pour des ondulations de section rectangulaire. Le résultat a été vérifié expérimentalement.

**Zusammenfassung**

Ebenso wie dünnwandige Zylinder, die in einem früheren Artikel behandelt wurden, können auch Faltenbälge unstabil werden, wenn sie einem inneren Druck unterworfen sind. Der kritische Wert dieses Druckes, welcher gemäß der bekannten Eulerschen Formel durch die Steifigkeit der Bälge in bezug auf Biegung bestimmt wird, ist nur für rechteckig gewellte Federungskörper berechnet worden. Die gefundenen Resultate wurden experimentell bestätigt.

## 1. Introduction

A previous paper [1]) dealt with the instability of thin-walled cylinders subjected to internal pressure. There it was shown that the total axial force exerted externally on the cylinder is to be substituted in Euler's well-known formula to produce the condition of buckling, it being immaterial, as regards the instability, whether the force is applied directly on the wall or, entirely or partly, via a piston on the liquid contained in the cylinder.

Owing to the great rigidity of the cylinder with respect to bending, the indicated case of instability will occur only at high liquid pressures and/or great lengths of the cylinder. Furthermore special precautions have to be taken to unload the cylinder wall entirely or partly in the axial direction, for if the cylinder is closed at its ends the axial tension is just sufficient to cancel the buckling effect resulting from the radial pressures.

*) First published in Dutch in De Ingenieur **63**, O42-O44, 1951.

In the case, however, of elastic elements, such as bellows, whose rigidity with respect to bending is a great deal less than that of thin-walled cylinders, the buckling effect described will occur at more normal pressures and lengths. The buckling of bellows subjected to internal pressure has actually been observed and the investigation into the cause of this — under the given conditions more or less surprising — effect put the author on the track of the similar behaviour of thin-walled cylinders.

## 2. Bellows built up from elastic flat rings subjected to internal pressure

The elementary calculation given in section 2 of the previous paper [1]), where the cylinder was regarded as a profiled straight rod, can also be used in the case of bellows subjected to internal pressure. All that is needed is to know their rigidity with respect to bending more precisely. This rigidity depends to a large extent upon the shape of the corrugations of the bellows and its calculation is generally a problem in itself.

In this section our considerations will be confined to the case where the bellows are composed of a number of elastic flat rings $a$ (see fig. 1) interconnected by completely rigid cylinders $d$ of radii $r_u$ and $r_i = r_u - b = \varrho r_u$ respectively. Only then can the rigidity with respect to bending be easily calculated, since the problem of the flat ring fixed at the edges and loaded by a bending moment has already been solved by Reissner [2]). If $h$ is the plate thickness and $E'$ the plate modulus [compare $E/(1-\nu^2)$ in the previous paper [1])] the moment $M_0$ required to turn the plane of the inner edge with respect to the plane of the outer edge over the unit angle amounts to

$$M_0 = \pi\varepsilon E' h^3 r_u^3 / b^3 ,$$

where

$$3\,\varepsilon = -\frac{(1+\varrho^2)\,(1-\varrho)^3}{1-\varrho^2+(1+\varrho^2)\ln\varrho} . \qquad (1)$$

The quantity $\varepsilon$ is represented diagrammatically in fig. 2 as a function of $\varrho$.

When the bellows of length $l$ have the number of corrugations $n$ (thus the number of flat rings is $2n$) the rigidity $c_0$ of the bellows with respect to bending and per unit of length is equal to $M_0 l/2n$, thus

$$c_\theta = \pi\varepsilon E' h^3 l r_u^3 / 2b^3 n .$$

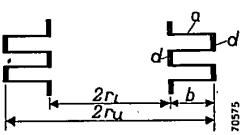



Fig. 1. Transverse cross-section of bellows built up from elastic flat rings $a$ interconnected by completely rigid cylinders $d$.

Replacing $EI$ in Euler's formula [compare eq. (1) in the previous paper [1])] by the rigidity $c_0$ we find as a condition of instability

$$P = \pi^3 \varepsilon E' h^3 r_u^3 / 2b^3 nl \ldots \quad (2)$$

By $P$ is meant here, as described in the introduction, the total axial force applied at the ends, it being immaterial whether the force is transmitted to the wall of the bellows directly as an "axial" load or, entirely or partly, by means of a liquid as a "radial" load.

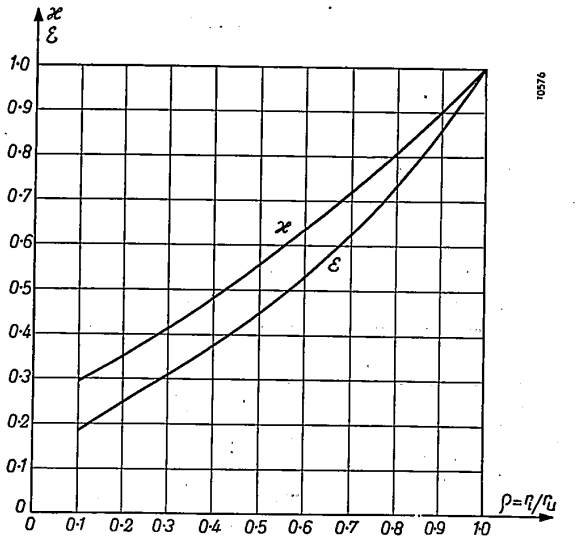



Fig. 2. The quantities $\varepsilon$ and $\varkappa$ as functions of $\varrho = r_i/r_u$.

The axial force exerted depends on the nature of the ends of the bellows. If (except for a connection to the pressure piping) the bellows are closed at the ends and further left to themselves, no axial force is exerted. The liquid pressure $q'$ against the ends then causes an axial tension in the bellows, resulting in the latter being lengthened over a distance which can be calculated with the aid of the theory of bending for flat rings:

$$f_0 = \frac{3nq' r_u^4}{8E'h^3} (1 - \varrho^4 + 4\varrho^2 \ln \varrho) . \quad (3)$$

On the other hand, if the bellows are allowed to undergo only an elongation $f (f < f_0)$ the compressive force to be exerted in the axial direction is equal to

$$P = \pi \varkappa q' r_u^2 (1 - f/f_0) , \quad (4)$$

where $f_0$ is given by (3) and $\varkappa$ is given by

$$\varkappa = \frac{(1-\varrho^2)\,(1-\varrho^4+4\varrho^2\ln\varrho)}{4(1-\varrho^2+2\varrho\ln\varrho)\,(1-\varrho^2-2\varrho\ln\varrho)}. \tag{5}$$

This quantity $\varkappa$ is plotted in fig. 2 as a function of $\varrho$.

According to eqs (2) and (4) the buckling of the bellows considered here appears to take place at the internal pressure $q'$, equal to

$$q' = \frac{\zeta E' h^3 r_u}{b^3 n l\,(1-f/f_0)}, \tag{6}$$

with

$$\zeta = \pi^2\,\varepsilon/2\varkappa\,. \tag{7}$$

According to eqs (1), (5) and (7) the quantity $\zeta$ is a function of $\varrho$, and it has been plotted as such in fig. 3.

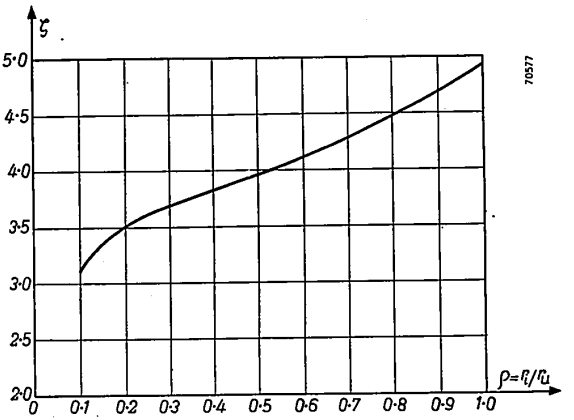



Fig. 3. The quantity $\zeta = \pi^2\varepsilon/2\varkappa$ as a function of $\varrho = r_i/r_u$.

### 3. Bellows with rectangularly shaped corrugations subjected to internal pressure

Whereas the bellows in the previous section were built up from a number of elastic flat rings interconnected by completely rigid cylinders, in the case of the bellows with rectangularly shaped corrugations we again have to do with a number of flat rings but these are now connected to each other by means of short elastic cylinders.

The bellows being compressed rotation-symmetrically about the axis by an axial force $P$, the load along the edges of the cylinders per unit

length of arc consists not only of the axial forces $P/2\pi r_u$ and $P/2\pi r_i$ respectively, but also of the bending moments $M$ and the radial shear forces $Q$ as shown in fig. 4. For the thin-walled cylinder of radius $r(r = r_u$ and $r = r_i$ respectively), height $2c$ and wall thickness $h$, we find with [3])

$$\left.\begin{aligned} \mu c &= \sqrt[4]{3(1-\nu^2)}\; c/\sqrt{hr}\,, \\ \lambda &= \frac{\sinh 2\,\mu c + \sin 2\,\mu c}{2\mu c\,(\cosh 2\mu c + \cos 2\mu c)}\,, \end{aligned}\right\} \tag{8}$$

that the angular rotation $\vartheta$ at the edge of the cylinder is equal to

$$\vartheta = 12\,\lambda Mc/E'h^3\,,$$

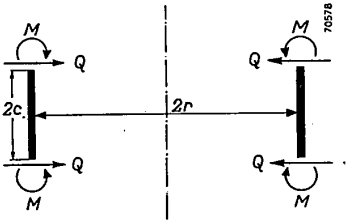



Fig. 4. The thin-walled cylinder loaded at the edges.

when for the sake of simplicity the radial displacement along the edge is taken to be zero. With the notations

$$\lambda_u = \lambda_{r=r_u}, \qquad \lambda_i = \lambda_{r=r_i},$$

this means that the edges of the flat rings have to be regarded as being clamped elastically with the coefficients of constraint equal to $E'h^3/12\lambda_u c$ and $E'h^3/12\lambda_i c$ respectively. The theory of bending of the flat plate teaches that the ratio $\chi$, indicating in what ratio the axial rigidity of the bellows with rectangularly shaped corrugations is smaller than that of the similar bellows built up from elastic rings, in the case $c \ll r_u$ is approximately equal to

$$\chi \approx 1 + F(\varrho, c/\sqrt{hr_u})c/r_u\,, \tag{9}$$

with

$$F(\varrho, c/\sqrt{hr_u}) = 2\,\frac{(1-\varrho^2+2\varrho^2\ln\varrho)^2\,\lambda_u + \varrho(1-\varrho^2+2\ln\varrho)^2\,\lambda_i}{(1-\varrho^2)\{(1-\varrho^2)^2 - 4\varrho^2(\ln\varrho)^2\}}. \tag{10}$$

The reciprocal value of the function $F(\varrho, c/\sqrt{hr_u})$, with Poisson's ratio $\nu = 0{\cdot}3$, is represented graphically in fig. 5.

If we put the effect of the elasticity of the cylinders upon the rigidity

with respect to bending as being equal to that upon the axial rigidity *), the critical pressure $q'$ will have to be reduced in the ratio $\chi$. As a result we thus find that the internal pressure at which the buckling of the bellows with rectangularly shaped corrugations occurs is given by

$$q' \approx \frac{\zeta' E' h^3 r_u}{b^3 n l (1 - f/f_0')}, \tag{11}$$

with

$$\zeta' = \zeta/\chi = \pi^2 e/2\varkappa\chi, \qquad f_0' \approx \chi f_0 . \tag{12}$$

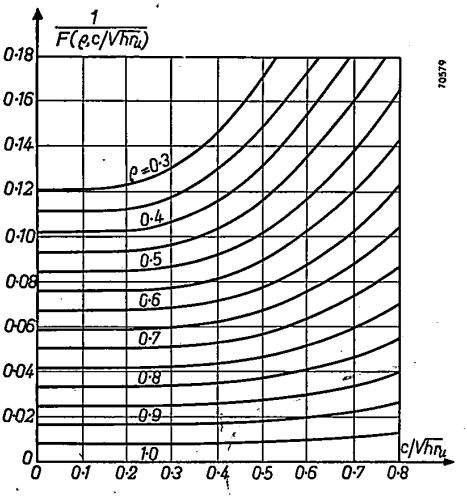



Fig. 5. The reciprocal value of $F(\varrho, c/\sqrt{hr_u})$ plotted as a function of $c/\sqrt{hr_u}$ for different values of $\varrho = r_i/r_u$.

Here it is reiterated that by the length $l$ is meant the so-called buckling length and that the number of corrugations $n$ refers to this length. In the case of two hinged ends or two ends constrained parallel, the buckling length is equal to the length of the bellows themselves, in the case of two clamped ends it is equal to half the length of the bellows, and in the case of one clamped and one free end it is equal to twice the length of the bellows.

Formula (11) has been put to the test with an example. The bellows with a length of 166 mm fixed at the ends have 55 corrugations. The dimensions of these corrugations follow from fig. 6. The rounded shape of the corruga-

*) The cylindrical part of the wall of the bellows behaves in principle less rigidly in the case of bending than in the case of compression. This difference, however, is of minor importance, as proved by experiments, and in the case of bending, too, $\chi$ is a correction factor that can be used fairly well.

IPR2025-00217 EXHIBIT 2010 Page 198 of 485

tions is imagined as being replaced by the circumscribed rectangle, so that the following dimensions have to be reckoned with

$$r_u = 19{\cdot}7 \text{ mm}, \qquad 2c = 1{\cdot}51 \text{ mm},$$
$$r_i = 13{\cdot}3 \text{ mm}, \qquad 2l = 166 \text{ mm},$$
$$b = 6{\cdot}4 \text{ mm}, \qquad 2n = 55.$$
$$h = 0{\cdot}16 \text{ mm},$$

Therefore $\varrho$ amounts to 0·675 and $c/\sqrt{hr_u} = 0{\cdot}425$. From the graphs of figs 3 and 5 we read $\zeta = 4{\cdot}24$ and $F(\varrho, c/\sqrt{hr_u}) = 17{\cdot}1$, so that $\chi$ becomes equal to 1·66 and $\zeta'$ equal to 2·56.

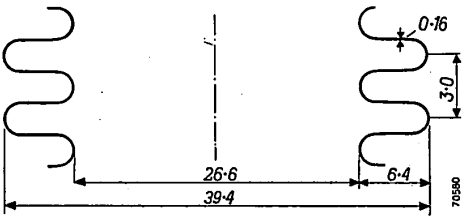



Fig. 6. Transverse cross-section of the bellows investigated for instability.

The internal pressure at which, according to formula (11), the bellows constrained between two flat end plates ($f = 0$) should buckle amounts to 0·34 kg/cm², the plate modulus $E'$ for tombac being taken equal to 10 000 kg/mm². The experiment (see figures 7 and 8) gives for this pressure 0·29 kg/cm². Considering the approximations introduced and the unavoidable variation of the wall thickness, this result may be considered to agree fairly well with the calculation.

## 4. Summary

In a previous paper [1]) the possibility of buckling of a thin-walled cylinder subjected to internal pressure has been pointed out. As a rule, on account of the great rigidity of the cylinder with respect to bending, this occurs only at high pressures and/or large lengths of the cylinder. Furthermore special precautions have to be taken to unload the wall of the cylinder entirely or partly in the axial direction.

When, however, one has to do with elastic elements of less rigidity with respect to bending such as bellows, this phenomenon of instability occurs at more normal lengths and pressures. Here too, the (corrugated) wall has to be unloaded in the axial direction, which can be done by constraining

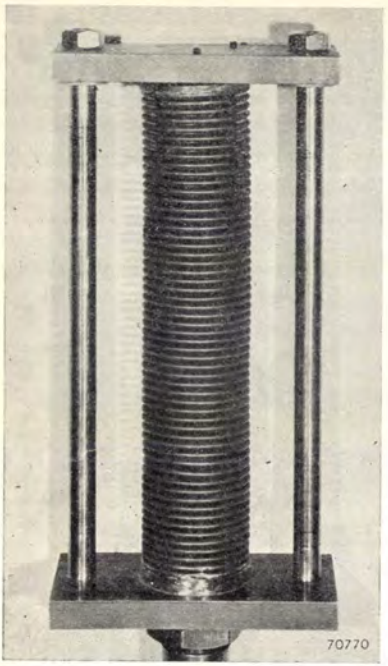



Fig. 7. Picture of the bellows, fixed at the ends, at an internal pressure of 0·28 kg/cm².

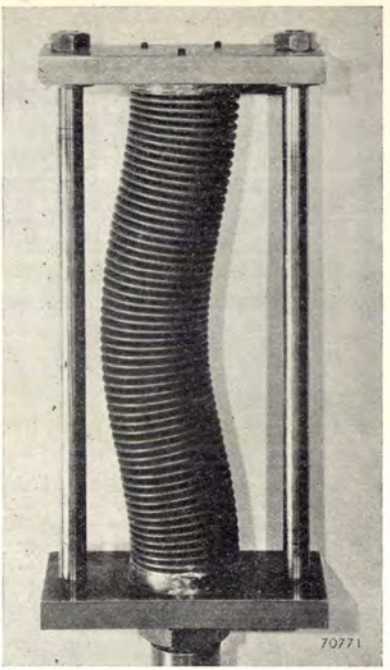



Fig. 8. Picture of the bellows, fixed at the ends, at a pressure of 0·29 kg/cm².

the bellows between two flat end plates. In the case where the bellows are provided with corrugations with a — possibly approximately — rectangular shape, the internal pressure $q'$ required for the instability follows from eqs (11) and (12) for $f = 0$. An example quoted from practice shows that these equations give a fairly good result.

*Eindhoven, October 1951*

## REFERENCES

1) J. A. Haringx, The instability of thin-walled cylinders subjected to internal pressure, Philips Res. Rep. **7**, 112-118, 1952.
2) H. Reissner, Über die unsymmetrische Biegung dünner Kreisringplatten, Ing. Archiv **1**, 72, 1929.
3) See, e.g., C. B. Biezeno and R. Grammel, Die Technische Dynamik, Berlin, Springer, 1939, p. 461.

R 192 *Philips Res. Rep. 7, 197-235, 1952*

# TRANSFER AND TRANSPORT OF ENERGY BY RESONANCE PROCESSES IN LUMINESCENT SOLIDS *)

by Th. P. J. BOTDEN 537.37:539.2

CONTENTS

1. Explanations given for the phenomena in sensitized phosphors.
2. Theoretical ratios of the fluorescence efficiencies as a function of sensitizer and activator concentrations according to the theories of section 1.
3. Experimental results on sensitized phosphors.
   3.1 Calcium orthophosphate activated with cerium and manganese.
   3.2 Strontium orthophosphate activated with tin and manganese.
   3.3 Calcium fluorophosphate activated with antimony and manganese.
   3.4 Calcium pyrophosphate activated with tin and manganese.
   3.5 Calcium carbonate activated with lead or thallium and manganese.
   3.6 Magnesium pyrophosphate activated with cerium, thorium and manganese according to the investigations of Froelich.
   3.7 Conclusion.
4. Explanation of the energy transfer in sensitized systems.
   4.1 Energy transfer from one excited sensitizer to another.
   4.2 Theoretical relation between the quantum efficiencies of fluorescence of the activator and sensitizer as a function of their concentrations, based on the assumptions of section 4.1.
   4.3 Calculation of $k$, $l$ and $p$ from the experiments.
   4.4 Discussion of results.
   4.5 Summary of section 4.
5. Theoretical treatment of the transfer processes in sensitized phosphors on the basis of the theory of Mott and Seitz.
   5.1 The model of Mott and Seitz applied to sensitized phosphors.
   5.2 Calculation of the activation energy $\varepsilon_t$ in calcium orthophosphate activated with cerium and manganese.
   5.3 Summary of section 5.

*) Thesis, Utrecht University, 1952. For Part I, see these Reports 6, 425-473, 1951.

# PART II

## TRANSFER AND TRANSPORT OF ENERGY IN SENSITIZED SYSTEMS; EXCITATION IN THE SENSITIZER

### Introduction

It is well known that a number of activated phosphors which can be excited by cathode rays or X-rays cannot be excited by ultraviolet radiation of wavelengths longer than about 2200 Å because this radiation is not absorbed. In some of these systems it is possible to extend the absorption to this ultraviolet region by the incorporation of suitable metal ions. Irradiation into this absorption band gives rise to fluorescence in a band due to the original activator. This phenomenon is called sensitization; the incorporated ion is called the sensitizer. In most cases excitation gives rise to fluorescence also in a band due to the sensitizer.

Because the absorption is localized at the sensitizer, energy has to be transferred from the sensitizer to the activator in order to obtain emission of fluorescence by the activator. It will be seen that transfer of energy takes place by resonance processes, but that, contrary to the systems described in part I, energy cannot be transported via constituents of the base lattice.

Sensitization was found already by Tomaschek [41]) (1924) in SrS-Gd-Sm and by Rothschild [42]) (1934) in (Ca, Sr)S-Sm-(Pb,Bi). It is only in the last few years that more attention has been paid to the sensitized systems.

Now a large number of systems are known, viz.: NaCl-Pb-Mn [43]), [44]), [45]); KCl-Pb-Mn [46]); $CaCO_3$-(Ce,Tl,Pb)-Mn [47]); $CaSiO_3$-Pb-Mn [45]), [48-52]); $Ca_3(PO_4)_2$-Ce-Mn [52]); $Ca_5P_3O_{12}F$-Sb-Mn [53]), [54]), [55]); $Mg_2P_2O_7$-Ce-Th-Mn [56]).

In this part II the mechanism of transfer and transport of energy will be discussed. Investigations on $Ca_5P_3O_{12}F$-Sb-Mn, $Sr_3(PO_4)_2$-Sn-Mn, $Ca_2P_2O_7$-Sn-Mn, $CaCO_3$-(Tl,Pb)-Mn and especially on $Ca_3(PO_4)_2$-Ce-Mn will be dealt with.

### 1. Explanations given for the phenomena in sensitized phosphors

Several theories have been proposed to explain the phenomena of sensitization. Some of them could be rejected because of later experiments.

(1) Transport of energy from the sensitizer to the activator by electrons via the conduction band [53]). For this the sensitizer ion should be ionized when excited by ultraviolet radiation. This transition of the electron is schematically given by transition *a* in fig. 1, which gives the energy band scheme. The hole in the sensitizer centre *S* may be liberated by transition *b* of an electron. The hole will then move through the lattice and, when meeting an activator centre *A*, may be trapped (*c*). Recombination of an electron from the conduction band with the hole (*d*) can give rise to emission

of activator fluorescence. This is the mechanism for the transfer of energy according to Klasens and Schön [13]), which is applicable to zinc-sulphide phosphors.

For this mechanism energy is needed both for transition *a* and for transition *b*. The first energy can be provided by the exciting radiation. Transition *b* requires an energy of more than 1 eV in most sensitized phosphors, as has been shown in an earlier paper [52]). This energy cannot be obtained from the lattice as vibration energy because sensitization occurs also at low temperatures (100 °K and even lower). The assumption that transition *b* is caused by a separate excitation quantum is in contradiction with the experiments. Then the quantum efficiency of the fluorescence could be at most 50 % and should depend quadratically on the intensity of the exciting radiation, neither of which is the case [45]), [52]). Therefore, although excitation may consist in the ionisation of the sensitizer centre, this centre must be brought back into its original unexcited state, and accordingly the fluorescence takes place at the same centre, so that the energy cannot be transferred.

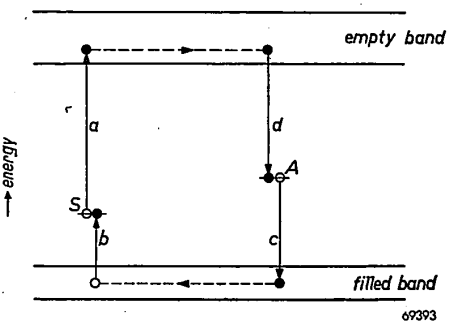



Fig. 1. The energy band scheme of a phosphor, with transfer of energy from one centre to the other according to Klasens and Schön.

(2) Transport of vibration energy from an excited sensitizer to the activator via the base lattice. The excitation energy is considered to produce a change of the potential-energy curve of the activator so that the activator can now be excited by another excitation quantum (Butler [57])). This theory has been rejected by Schulman and coworkers [45]), because the quantum efficiency should depend quadratically on the intensity of the exciting radiation, contrary to the experiments.

(3) The formation of complex centres of sensitizer and activator (according to Murata-Smith called "coactivators" [43])) is also in contradiction with the experiments [45]), [52]).

(4) Transport of energy via an excited state of the base lattice in a similar way as described in part I. However, from the position of the absorption edge it can be concluded that excitation of the base lattice to its lowest excited state requires an energy that is about 1 eV higher than the energy of an excitation quantum. This energy of 1 eV is too large to be provided by the lattice as vibration energy and cannot be taken from a second excitation quantum for the reason stated in (1).

Remaining possible explanations:

(1) Transport of the energy from the sensitizer to the activator via a so-called long-wave exciton band. Franck and Teller [58]) suggested that the matrix lattice may have an excited state which is lower than the lowest one found on the basis of absorption measurements; the lower excited state (long-wave exciton) is not indicated by absorption because transition between this state and the ground state is optically forbidden. When the energy of the exciton is close to that of the excited state of a sensitizer, transport of energy might occur via the exciton band. Although this explanation is not very attractive, it could not be excluded by the experiments.

(2) Direct transfer of energy from the sensitizer to the activator by a resonance process, similar to that occurring in gases and solutions. When the sensitizer and activator have excited states of nearly equal energy it is possible that the activator is brought into an excited state by a resonance process even when the transition in the activator is optically forbidden [6]). The probability of the energy transfer decreases rapidly with increasing distances between the sensitizer and the activator. The question arising is over what distance transfer may take place. With respect thereto we may distinguish two theories:

(a) Botden and Kröger [52]) suggested that transfer will take place only when the sensitizer and activator are situated at adjoining metal ion sites ("pairs"). With a statistical distribution of sensitizer and activator ions the fraction of pairs is too low to explain the high efficiencies of the activator fluorescence. Therefore it was assumed that an influence exists stimulating the formation of pairs.

This stimulation might be explained as follows: when the radius of the sensitizer ion is larger and the ionic radius of the activator is smaller than that of the metal ion of the matrix lattice, the sensitizer and activator tend to occupy adjacent sites in order to reach a position with a minimum of potential energy.

Evidence supporting this theory has been given by Kröger [59]). From the emission spectra and the temperature dependence of the intensity and the decay of the manganese fluorescence in $Ca_3(PO_4)_2$-(Sn,Ce)-Mn and $Sr_3(PO_4)_2$-(Sn,Ce)-Mn it followed that the sensitizer strongly influences radiation-less transitions in the activator, and this can only be explained if the sensitizer and the activator are situated close together. Yet the following

experiments carried out to prove the formation of pairs by chemical methods were unsuccessful *). Calcium orthophosphate activated with tin and manganese was prepared in a normal way; a part of the product was cooled suddenly and the other part cooled slowly from 1100 °C to room temperature. In the first case the distribution of the sensitizer and the activator ions is expected to be nearly at random over the lattice, whereas in the product cooled slowly more pairs of sensitizer and activator ions will be found if there is an influence stimulating the formation of pairs. However, no difference in the optical properties was observed.

Another possibility to explain the high efficiencies of the activator fluorescence was mentioned to be the transfer of energy from an excited sensitizer via neighbouring sensitizers to the activator. This mechanism will be considered in section 4.

(b) Schulman a.o. [45]), assuming a random distribution of the sensitizer and the activator ions over the lattice, supposed that the energy is transferred directly from the sensitizer to the activator by a resonance process over distances which may be larger than those between a sensitizer and an activator on adjacent cation sites. The theory was based on measurements of absorption in $CaSiO_3$-Pb-Mn. According to Froelich [48]), however, the $\alpha$-wollastonite structure of $CaSiO_3$-Pb-Mn with low concentrations of lead and manganese transforms into the $\beta$-modification at higher concentrations of lead and manganese. Although this effect possibly may be neglected according to Schulman a.o. their conclusions are not entirely convincing.

## 2. Theoretical ratios of the fluorescence efficiencies as a function of sensitizer and activator concentrations according to the theories of section 1

In order to test the theories mentioned above, new experiments have been carried out on sensitized systems with varying concentrations of sensitizer and activator ions. In this section relations will be derived between the quantum efficiencies of the fluorescence of the sensitizer and the activator on the basis of each of the theories. A sensitizer ion will be denoted by $S$ and an activator ion by $A$. In all cases concentrations are defined as the ratio of the number of sensitizer or activator ions to the number of cation sites of the lattice.

(1) Transport of energy from $S$ to $A$ via a long-wave exciton band. In this case an excited $S$ can dissipate the energy:

(a) through a transition to the ground state with emission of $S$ fluorescence with a transition probability $p_s$;

(b) by forming an exciton with a transition probability $p_e$; dissipation of energy by a quenching process will be left out of account, thus $p_s + p_e = 1$.

Then the probability that the excited state of $S$ is transformed into an

*) Dr F. A. Kröger, private communication.

exciton is $p_e/(p_e + p_s)$. The exciton may move through the lattice until it is captured by an $S$ with a transition probability $q_s$ or captured by an $A$ with transition probability $q_a$. Once the energy is transferred to $A$ it is assumed that emission of $A$ fluorescence follows, whereas $q_s$ is so defined that capturing of the exciton in $S$ leads to emission of $S$ fluorescence and not to formation of a new exciton. Let $X_s$ be the concentration of $S$ and $X_a$ that of $A$. Then the probability that the exciton gives rise to $A$ fluorescence is

$$X_a q_a/(X_a q_a + X_s q_s)$$

and the ratio $B$ of the quantum efficiencies of the $A$ fluorescence to the $S$ fluorescence is

$$B = N/(1-N), \text{ where } N = \{p_e/(p_e+p_s)\}\{X_a q_a/(X_a q_a + X_s q_s)\}.$$

Hence

$$B = c_1/(c_2 + c_3 X_s), \tag{1}$$

in which $c_1$, $c_2$ and $c_3$ are constant if $X_a$ is constant.

From (1) it follows that $B$ must decrease with increasing $X_s$ for constant $X_a$.

(2) Formation of $S$-$A$ pairs is stimulated. It may be assumed that absorption in an $S$ that is one of a pair, gives rise to $A$ fluorescence, whereas absorption in a "free" $S$ (not belonging to an $S$-$A$ pair) leads to $S$ fluorescence.

Let $X_s$ and $X_a$ be the concentrations of all $S$ and $A$; and let $X_p$ be the concentration of $S$-$A$ pairs, so that $(X_s - X_p)$ is the concentration of the free $S$ and $(X_a - X_p)$ the concentration of the free $A$.

The combinations of one $S$ with two $A$, one $A$ with two $S$, etc., will be neglected because the concentrations of $S$ and $A$ are low.

An equilibrium will exist between the free atoms and the pairs, which may be written in the form $S + A \rightleftarrows SA$, for which the law of mass action gives

$$X_p/(X_s - X_p)(X_a - X_p) = K,$$

in which $K$ is the equilibrium constant. Now the ratio $B$ of the quantum efficiencies of the $A$ fluorescence to the $S$ fluorescence is

$$B = X_p/(X_s - X_p) = K(X_a - X_p), \tag{2}$$

if it is assumed that the absorption strength of a free $S$ and that of an $S$ belonging to a pair are equal. When this is not the case a numerical factor has to be added in (2), which is unimportant here.

When $X_a$ is kept constant $X_p$ will increase with increasing $X_s$ and therefore $B$ will decrease with increasing $X_s$ for constant $X_a$.

(3) Random distribution of $S$ and $A$ over the lattice and direct transfer of energy from $S$ to $A$ over distances which may be greater than those between $S$ and $A$ on adjacent metal ion sites. Let $k$ be the number of cation sites round an $S$ such that the energy is transferred from $S$ to an $A$ when this

$A$ is situated on one of the $k$ sites. As will be seen in section 4.2 the ratio $B$ is

$$\{1-(1-X_a)^k\}/(1-X_a)^k \qquad (3)$$

(i.e. $P_2/P_1$ in section 4.2). Thus $B$ is independent of $X_s$ *).

According to the above theories the ratio $B$ will not increase with increasing $X_s$ and constant $X_a$. In the next section it will be seen how far the experiments are in agreement with this.

## 3. Experimental results on sensitized phosphors

In all products the concentrations of the sensitizer or activator ions are given as percentages of the number of these ions to the total number of cation sites of the lattice.

### 3.1. *Calcium orthophosphate activated with cerium and manganese*

The products were prepared **) according to a description given elsewhere [60], [61].

$Ca_3(PO_4)_2$-Mn does not show an absorption band in the ultraviolet for

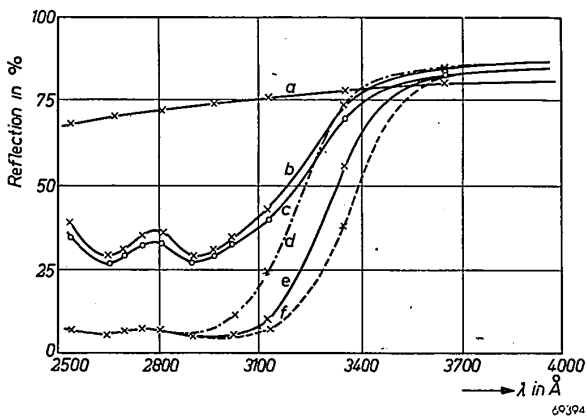



Fig. 2. Reflection spectra of: (*a*) $Ca_3(PO_4)_2$-3/2 % Mn at 300 °K; (*b*) $Ca_3(PO_4)_2$-1/3 % Ce at 300 °K; (*c*) $Ca_3(PO_4)_2$-1/3% Ce-5/3 % Mn at 300 °K; (*d*) $Ca_3(PO_4)_2$-5/3 % Ce-2/3 % Mn at 110 °K; (*e*) $Ca_3(PO_4)_2$-5/3 % Ce-2/3 % Mn at 300 °K; (*f*) $Ca_3(PO_4)_2$-5/3 % Ce-2/3 % Mn at 575 °K. All concentrations of Ce and Mn are given as percentages of the number of cation sites.

*) In formula (3) are included the probabilities of finding more than one $A$ and/or one $S$ (or more) on the $k$ sites around a given $S$, with the previous assumptions that there is equal absorption in all $S$ and that transfer of energy from $S$ to $A$ leads to $A$ fluorescence. When only one $A$ around an $S$ is taken into account an expression similar to formula (2) is found.

**) The preparation of all phosphors mentioned in this part was carried out under the supervision of Dr F. A. Kröger and Mr P. Zalm.

wavelengths longer than 2500 Å, as may be seen in fig. 2, which gives the spectral distributions of diffuse reflections for various samples. However, excitation is possible by cathode rays and then a red fluorescence is emitted.

$Ca_3(PO_4)_2$-Ce shows an ultraviolet absorption with a long-wavelength limit at about 3400 Å, which must be ascribed to the $Ce^{3+}$ion. The reflection spectrum shows two minima, situated at $\lambda$ 2650 Å and $\lambda$ 2900 Å. Excitation of $Ca_3(PO_4)_2$-Ce in one of the absorption bands gives rise to an ultraviolet emission. In figs 3 and 4 the emission spectra of $Ca_3(PO_4)_2$-Ce are given at various temperatures for excitation by $\lambda$ 2537 Å. The spectrum of

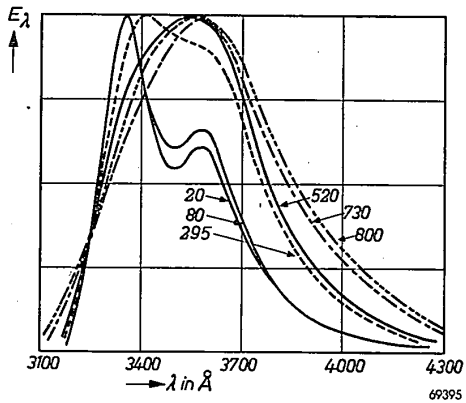



Fig. 3. Emission spectra of $Ca_3(PO_4)_2$-1/9%Ce at 20, 80, 295, 520, 730 and 800 °K excitation by $\lambda$ 2537 Å.

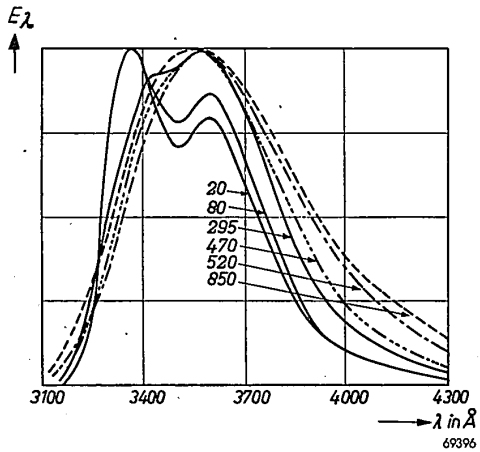



Fig. 4. Emission spectra of $Ca_3(PO_4)_2$-2/3%Ce at 20, 80, 295, 470, 520 and 850 °K; excitation by $\lambda$ 2537 Å.

the cerium fluorescence consists of two bands with maxima at $\lambda$ 3350 Å and $\lambda$ 3600 Å. This can be seen in the spectra at low temperatures; at higher temperatures the intensity of the short-wavelength emission band decreases, that of the long-wavelength band increases and the bands are broadened. The double peak is due to electronic transitions from an excited $Ce^{3+}$ level to the levels $^2F_{5/2}$ and $^2F_{7/2}$ of the ground state [62]).

The emission spectra are found to be the same for excitation by $\lambda$ 3130 Å as for excitation by $\lambda$ 2537 Å *).

Fig. 5 shows the emission spectra of $Ca_3(PO_4)_2$-$x$ % Ce, with $x = 1/36$, 1/9, 2/3 and 16/9, at room temperature for excitation by $\lambda$ 2537 Å; the intensity of the long-wavelength band increases with respect to the short-wavelength band with increasing cerium content.

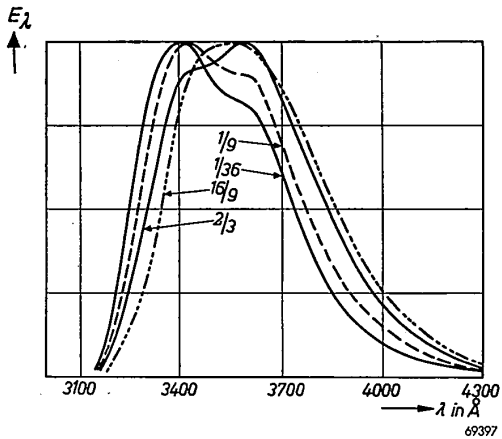



Fig. 5. Emission spectra of $Ca_3(PO_4)_2$-$x$%Ce, with $x$ = 1/36, 1/9, 2/3 and 16/9, at room temperature; excitation by $\lambda$ 2537 Å.

This relative increase of the long-wavelength emission band of cerium, both with increasing cerium content and with increasing temperature, is accompanied by a shift of the long-wavelength side of the absorption spectrum. The absorption side shifts to longer wavelengths with increasing temperatures and also, though less markedly, with increasing cerium content (fig. 2). As a result of this shift, the fluorescence radiation of the short-wavelength part of the fluorescence spectrum, which becomes overlapped by the absorption spectrum, can be absorbed by the phosphor itself (self-absorption) and again give rise to

*) To measure the emission spectra, the exciting radiation of $\lambda$ 2537 Å was obtained from a low-pressure mercury discharge lamp in combination with a Bäckström filter [63]) (a cell filled with a saturated solution of $CoSO_4$-$NiSO_4$); the radiation of $\lambda$ 3130 Å was obtained from a high-pressure mercury lamp in combination with a $K_2CrO_4$ filter and a WG 7 (Schott) filter. The measurements of quantum efficiencies are carried out with monochromatic exciting radiation (obtained with the aid of a quartz double monochromator).

fluorescence, so that the intensity of the long-wavelength emission band will increase. (That the absorbed short-wavelength fluorescence radiation can again excite the phosphor follows from the excitation spectrum of $Ca_3(PO_4)_2$-Ce in fig. 6: excitation by $\lambda$ 3342 Å shows a quantum efficiency of 55 %.) In this way the change of the relative intensities of the emission bands both with temperature and with cerium concentration might be explained.

However, it is unlikely that the strong temperature dependence of each emission spectrum is caused only by the change in absorption, because:

(*a*) the fluorescence intensity of about $\lambda$ 3200 Å and shorter increases with temperature, while it is expected to decrease owing to the strong increase of absorption;

(*b*) the quantum efficiency of $Ca_3(PO_4)_2$-Ce is low for excitation by wavelengths in the long-wavelength limit of the absorption spectrum (fig. 6). When this absorption limit is shifted to longer wavelengths the self-absorbed fluorescence radiation will only partly be converted into fluorescence radiation and the remaining part is dissipated without radiation, so that the quantum efficiency is decreased. Therefore the strong change of the emission spectrum with temperature ought to be accompanied by a fairly strong decrease of the quantum efficiency, but only a slight temperature dependence of the cerium fluorescence in $Ca_3(PO_4)_2$-Ce is found, as will be seen in fig. 8.

Probably a temperature dependence of the emission spectrum is also caused by a variation of the probabilities of transition from the excited level to the $^2F_{5/2}$ and $^2F_{7/2}$ levels.

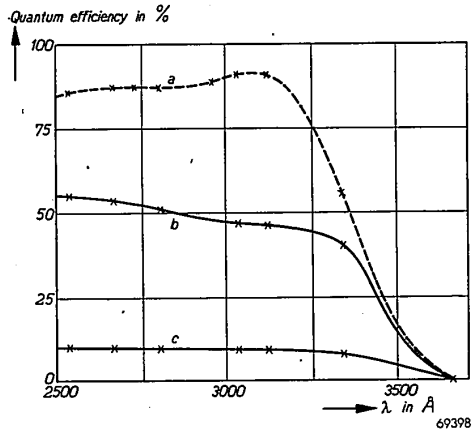



Fig. 6. Excitation spectra at room temperature:
(*a*) $Ca_3(PO_4)_2$-5%Ce; (*b*) $Ca_3(PO_4)_2$-5%Ce-2·5%Mn (Mn emission);
(*c*) $Ca_3(PO_4)_2$-5%Ce-2·5%Mn (Ce emission).

$Ca_3(PO_4)_2$-Ce-Mn shows the same absorption behaviour as $Ca_3(PO_4)_2$-Ce, as may be seen in fig. 2 for the phosphors with 1/3%Ce (curves *b* and *c*); with other cerium concentrations too only slight changes in the spectrum are found when manganese is added.

Excitation of $Ca_3(PO_4)_2$-Ce-Mn in this absorption region now gives rise to a red manganese fluorescence in addition to the cerium emission: part of the excitation energy is transferred from the cerium to the manganese ions. The excitation spectrum for the cerium and manganese fluorescence is similar to that for the cerium fluorescence in $Ca_3(PO_4)_2$-Ce (cf. fig. 6).

The spectral distribution of the cerium emission is found to be nearly the same whether manganese is present or not.

The emission spectrum of the manganese fluorescence at various temperatures is given in fig. 7. With increasing temperatures the maximum of the emission band shifts to shorter wavelengths and the band is broadened; apparently the Mn emission consists of two overlapping sub-bands. At 20 °K the emission spectrum is nearly the same as at 110 °K, with a maximum at $\lambda$ 6730 Å. The spectrum is only dependent on temperature and does not change when the concentrations of cerium or manganese are varied. The Mn emission spectrum is about the same as that given in fig. 7 when $Ca_3(PO_4)_2$-Ce-Mn or $Ca_3(PO_4)_2$-Mn is excited by cathode rays [52]).

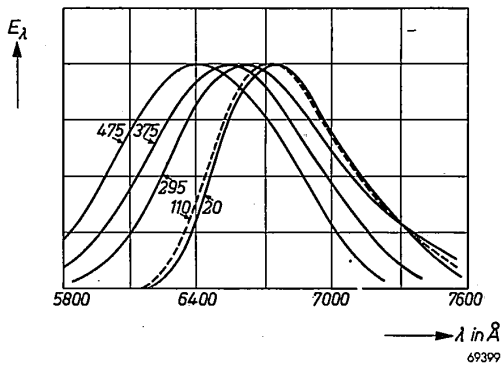



Fig. 7. Emission spectra of $Ca_3(PO_4)_2$-Ce-Mn (Mn emission) at 20, 110, 295, 375 and 475 °K; excitation by $\lambda$ 2537 Å.

Fig. 8 shows the temperature dependence of the quantum efficiencies of the cerium fluorescence in $Ca_3(PO_4)_2$-4/3%Ce and of both the cerium and manganese fluorescence in $Ca_3(PO_4)_2$ activated with respectively 1/3%Ce-1/3%Mn, 10/3%Ce-1/3%Mn, 1/3%Ce-5/3%Mn and 5/3%Ce-5/3%Mn.

The relative efficiencies at temperatures below 100 °K are determined from the emission spectra of the phosphors measured at various temperatures, while the intensity of exciting radiation is kept constant. To carry out these measurements the phosphor is fixed at the outside of a thick lead tube and this tube placed in a quartz Dewar vessel. The vessel can be evacuated to prevent condensation onto the powder. The ultraviolet exciting radiation reaches the phosphor via a quartz window in the vessel and the fluorescence radiation can come out also through the window. The phosphor is cooled to about 20 °K by filling the lead tube with liquid hydrogen. When the hydrogen is evaporated the temperature rises slowly and is measured with a thermo-couple, fixed near the phosphor *).

In $Ca_3(PO_4)_2$-Ce the efficiency of the cerium fluorescence is nearly constant between 20 °K and 450 °K. In $Ca_3(PO_4)_2$-Ce-Mn the efficiency of

*) I wish to thank Dr J. Volger for his help with the measurements in liquid hydrogen.

the cerium fluorescence is constant from 20 °K to about 150 °K but then the efficiency decreases with rising temperatures; above about 350 °K the efficiency is again constant (until quenching sets in at high temperatures). The efficiency of the manganese fluorescence is also constant below 150 °K and above 350 °K but increases with rising temperatures in the range from 150 °K to 350 °K, in which the efficiency of the cerium fluorescence decreases. The temperature dependence is found to be the same for excitation by $\lambda$ 3130 Å as for excitation by $\lambda$ 2537 Å. This temperature dependence of the efficiencies of the cerium and manganese fluorescences cannot be explained by the shift of the long-wavelength limit of the absorption spectrum with temperature, because above 350 °K this absorption limit still shifts whilst the efficiencies are constant. Probably we are dealing with a temperature-dependent transfer mechanism. A further discussion will be given in section 5. It appears from the above experiments that addition of manganese in $Ca_3(PO_4)_2$-Ce does not influence the excitation spectra and the spectral distribution of the Ce emission and reflection. Therefore the system $Ca_3(PO_4)_2$-Ce-Mn seems to be very suitable to test the formulae derived in section 2.

For this purpose the quantum efficiencies of the cerium and the manganese fluorescence at room temperature have been measured on some

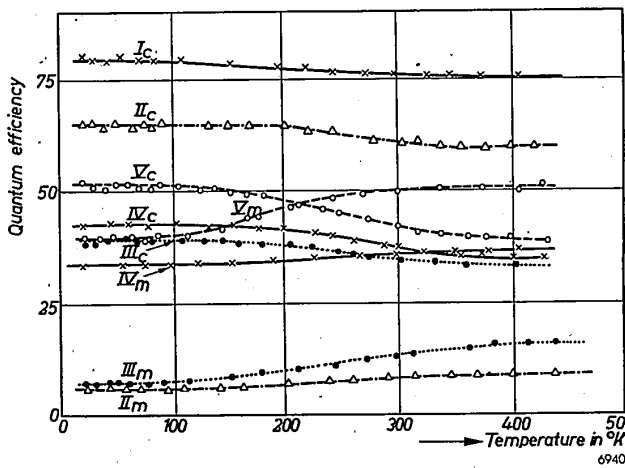



Fig. 8. Temperature-dependence of the quantum efficiencies of the cerium fluorescence (index $c$) and the manganese fluorescence (index $m$) for excitation by $\lambda$ 2537 Å.

$I_c$ : $Ca_3(PO_4)_2$-4/3 % Ce
$II_c$ and $II_m$ : $Ca_3(PO_4)_2$-1/3 % Ce-1/3 % Mn
$III_c$ and $III_m$: $Ca_3(PO_4)_2$-10/3 % Ce-1/3 % Mn
$IV_c$ and $IV_m$ : $Ca_3(PO_4)_2$-1/3 % Ce-5/3 % Mn
$V_c$ and $V_m$ : $Ca_3(PO_4)_2$-5/3 % Ce -5/3 % Mn.

IPR2025-00217 EXHIBIT 2010 Page 212 of 485

TABLE I

| Phosphor *) $Ca_3(PO_4)_2$-$x$ % Ce - $y$% Mn | | Quantum efficiency in % at room temperature; excitation $\lambda$ 2537 Å. | | Ratio of quant. eff. of Mn fluor. to Ce fluor. | Ratio of quant. eff. of Mn fluor. to (Ce+Mn) fluor. |
|---|---|---|---|---|---|
| $x$ | $y$ | Ce fluor. | Mn fluor. | | |
| 1/3 | 0 | 79·5 | 0 | — | — |
| 2/3 | 0 | 81 | 0 | — | — |
| 5/3 | 0 | 86·5 | 0 | — | — |
| 10/3 | 0 | 88 | 0 | — | — |
| 1/3 | 1/6 | 78·5 | 5·3 | 0·067 | 0·063 |
| 2/3 | 1/6 | 75·5 | 6·5 | 0·086 | 0·079 |
| 5/3 | 1/6 | 82 | 7·8 | 0·095 | 0·087 |
| 10/3 | 1/6 | 79·5 | 11·0 | 0·138 | 0·121 |
| 1/3 | 1/2 | 64·5 | 12·0 | 0·185 | 0·158 |
| 2/3 | 1/2 | 70 | 19·0 | 0·27 | 0·21 |
| 5/3 | 1/2 | 68 | 20·0 | 0·29 | 0·23 |
| 10/3 | 1/2 | 54·5 | 23·0 | 0·42 | 0·30 |
| 1/3 | 1 | 52·5 | 24·5 | 0·47 | 0·32 |
| 2/3 | 1 | 53·5 | 32·5 | 0·61 | 0·37 |
| 5/3 | 1 | 58 | 37·5 | 0·65 | 0·39 |
| 10/3 | 1 | 44·5 | 41 | 0·92 | 0·48 |
| 1/3 | 5/3 | 37 | 36 | 0·98 | 0·49 |
| 2/3 | 5/3 | 37·5 | 40 | 1·06 | 0·52 |
| 5/3 | 5/3 | 42 | 49·5 | 1·18 | 0·54 |
| 10/3 | 5/3 | 33 | 52 | 1·58 | 0·61 |
| 10/3 | 10/3 | 15·5 | 60 | 3·87 | 0·80 |

*) All concentrations are given as percentages of the number of cation sites.

phosphors with various cerium and manganese concentrations; from this the ratios $B$ of the quantum effiiciencies of the manganese to the cerium fluorescence are calculated. The phosphors were excited by $\lambda$ 2537 Å.

The results are given in table I, which shows that for a given Mn concentration the ratio $B$ increases with increasing cerium concentration. This is also shown in table II, in which the results are given of a set of phosphors $Ca_3(PO_4)_2$-$x$%Ce-1/3%Mn prepared all together, under the same conditions *). The results are in contradiction with the formulae (1), (2) and (3).

*) The cerium and manganese contents in these products have been analysed; those of manganese were found to be constant within a few percent whilst the cerium contents showed deviations up to 15 % from the intended concentrations. The products mentioned in table I could not be analysed because the quantity left over after the experiments

TABLE II

| Phosphor $Ca_3(PO_4)_2$-$x$%Ce-1/3%Mn | Quantum efficiency at roomtemp.; excitation $\lambda$ 2537 Å | | Ratio of quant. eff. of Mn fluor. to Ce fluor. | Ratio of quant. eff. of Mn fluor. to (Ce+Mn) fluor. |
|---|---|---|---|---|
| $x$ | Ce fluor. | Mn fluor. | | |
| 0·11 | 39·8 | 4·6 | 0·115 | 0·104 |
| 0·35 | 60·9 | 8·5 | 0·140 | 0·122 |
| 0·57 | 55·8 | 9·6 | 0·172 | 0·147 |
| 1·10 | 53·5 | 11·1 | 0·208 | 0·175 |
| 1·82 | 55 | 15·1 | 0·275 | 0·215 |
| 3·05 | 37·5 | 14·1 | 0·375 | 0·275 |

It may be thought that the increase of the ratio of the quantum efficiencies of the manganese fluorescence to the cerium fluorescence is due to self-absorption: with increasing Ce content the long-wavelength side of the absorption spectrum is shifted to longer wavelengths so that the overlapped short-wavelength tail of the Ce fluorescence can be absorbed and part of the self-absorbed energy may be transferred to manganese. But a rough calculation shows that this effect can only partly explain the experimental results.

Let us assume that all the Ce fluorescence with $\lambda < 3350$ Å is self-absorbed in the products with high Ce content and that all this energy is transferred to manganese. Then from fig. 4 it follows that about 20 % of the quantum efficiency of the Ce fluorescence in the phosphor with low Ce content would be transferred to manganese in the phosphor with high Ce content. Comparing $Ca_3(PO_4)_2$-1/3 % Ce-1 % Mn with $Ca_3(PO_4)_2$-10/3 % Ce-1 % Mn (table I) the increase of the quantum efficiency of the Mn fluorescence would be 0·20.52·5 = 10·5%; this has to be corrected by a factor 85·5/77 (because the total quantum efficiency is increased from 77 to 85·5%), thus giving 11·5 %. But this increase is found to be 41—24·5 = 16·5%. Thus the calculated increase is lower than the experimental one. Furthermore only a part of the short-wavelength Ce fluorescence is self-absorbed and this energy is transferred only partly to manganese. Then it is clear that, although through self-absorption the ratio of efficiencies of the Mn to the Ce fluorescence may increase with the Ce content, the most important part of the increase is due to another process (considered in section 4).

It may be noticed that a little self-absorption must occur according to the theory of Förster [11]) in order to explain the energy transport, as will be seen in section 4.

### 3.2. *Strontium orthophosphate activated with tin and manganese*

The method of preparation is given elsewhere [60]), [64]).

In fig. 9 the reflection spectra are given of $Sr_3(PO_4)_2$, $Sr_3(PO_4)_2$-Sn and $Sr_3(PO_4)_2$-Sn-Mn. $Sr_3(PO_4)_2$ does not show an absorption band for $\lambda > 2500$ Å. This is also the case with $Sr_3(PO_4)_2$-Mn, but this phosphor can be excited by cathode rays and then shows fluorescence in a green band with maximum at $\lambda$ 5250 Å [60]).

was too small, but the deviations of the cerium and manganese concentrations from the numerical values should be about the same as those found in the products of table II. The phosphors of table I were prepared in series with constant cerium and various manganese contents.

$Sr_3(PO_4)_2$-Sn shows an absorption band in the ultraviolet with a long-wavelength limit at about $\lambda$ 3000 Å. Excitation in this band gives rise to an ultraviolet fluorescence. The emission spectrum is temperature dependent and probably consists of two bands with maxima at about $\lambda$ 3350 Å and $\lambda$ 3850 Å (cf. fig. 10).

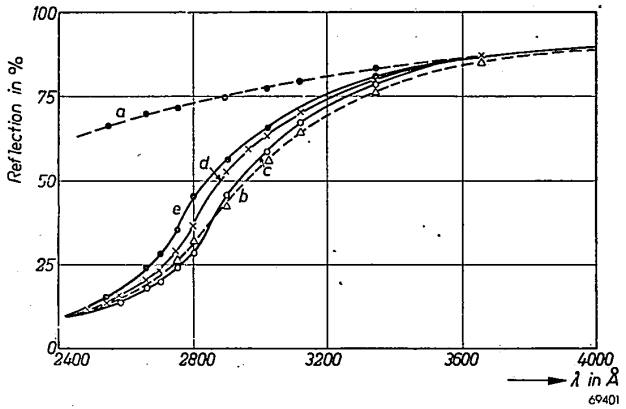



Fig. 9. Reflection spectra of (*a*) $Sr_3(PO_4)_2$ at 295 °K; (*b*) $Sr_3(PO_4)_2$-2%Sn at 295 °K; (*c*) $Sr_3(PO_4)_2$-2%Sn-1%Mn at 430 °K; (*d*) $Sr_3(PO_4)_2$-2%Sn-1%Mn at 295 °K; (*e*) $Sr_3(PO_4)_2$-2%Sn-1%Mn at 120 °K.

, The spectrum is independent of the tin concentration up to about 4 % Sn; at higher concentrations a faint blue emission band appears, with maximum at $\lambda$ 4500 Å, in addition to the ultraviolet bands.

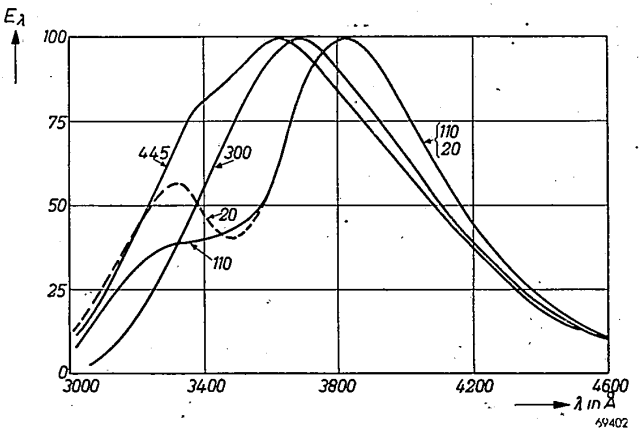



Fig. 10 Emission spectra of $Sr_3(PO_4)_2$-1/3 % Sn at 20, 110, 300 and 445 °K; excitation by $\lambda$ 2537 Å.

The absorption band of $Sr_3(PO_4)_2$-Sn-Mn is similar to that of $Sr_3(PO_4)_2$-Sn; the long-wavelength edge shifts to longer wavelengths with increasing temperatures. Excitation of $Sr_3(PO_4)_2$-Sn-Mn in the tin absorption band now gives rise to manganese fluorescence in addition to the tin fluorescence. The spectrum of the tin emission is the same as in the phosphors without manganese. With increasing temperatures the manganese emission band is broadened and shifts to shorter wavelengths, as can be seen in fig. 11.

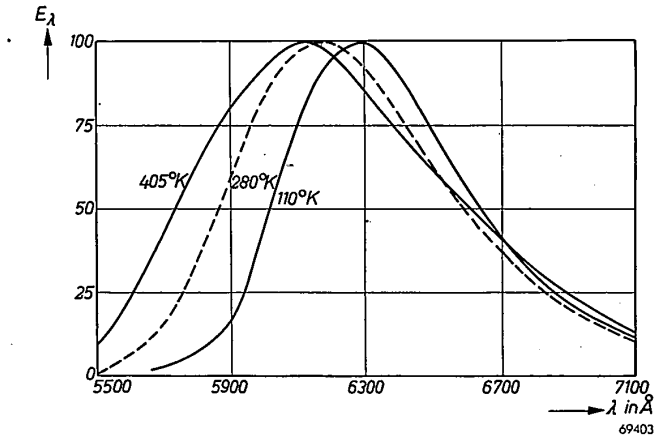



Fig. 11. Emission spectra of $Sr_3(PO_4)_2$-1 % Sn-1 % Mn (Mn emission) at 110, 280 and 405 °K; excitation by $\lambda$ 2537 Å.

At 20 °K the maximum of emission lies at $\lambda$ 6280 Å. Probably the manganese emission consists of two overlapping sub-bands. The spectrum is found to be independent of the manganese concentrations and of the tin concentration up to about 3%Sn. At higher tin concentrations the manganese emission spectrum shifts to shorter wavelengths.

In fig. 12 the temperature dependence is given of the quantum efficiencies of the tin fluorescence in $Sr_3(PO_4)_2$-Sn and of both the tin and manganese fluorescence in $Sr_3(PO_4)_2$-Sn-Mn; excitation by $\lambda$ 2537 Å.

The efficiency of the manganese fluorescence is independent of temperature up to 320 °K, then quenching begins; at 550 °K the fluorescence is totally quenched.

The efficiency of the tin fluorescence is constant between 20 °K and 50 °K and then increases with temperature; above 260 °K the efficiency is again constant until quenching begins at about 650 °K.

The behaviour of the efficiency of the tin fluorescence in $Sr_3(PO_4)_2$-Sn-Mn is similar to that in the phosphor without manganese except within the

temperature range between 450 °K and 550 °K, where part of the tin fluorescence is quenched.

Thus the temperature dependence of the tin fluorescence is rather complicated.

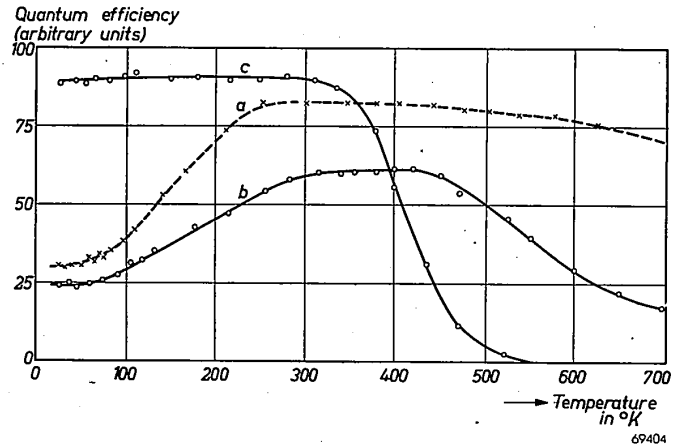



Fig. 12. Temperature dependence of the quantum efficiencies (in arbitrary units). (*a*) tin fluorescence in $Sr_3(PO_4)_2$-1/3%Sn; (*b*) tin fluorescence in $Sr_3(PO_4)_2$-1%Sn-1%Mn; (*c*) manganese fluorescence in $Sr_3(PO_4)_2$-1%Sn-1%Mn. Excitation by $\lambda$ 2537 Å.

The increase of the efficiency of the tin fluorescence between 50 °K and 260 °K, which occurs whether manganese is present or not, may be explained in the same way as has been done for the similar temperature dependence of the efficiency of the manganese fluorescence in $Cd_2B_2O_5$-Mn [25]). Thus it is assumed that excitation of $Sr_3(PO_4)_2$-Sn leads to a higher excited state of tin than the one responsible for the tin fluorescence and that a radiation-less transition exists from the higher excited state to the ground state. At low temperatures the probability of this transition has to be great enough to compete with the probability of the transition from the higher excited state to the lower state from which the fluorescence transition takes place. With increasing temperatures the first probability may increase less strongly than the second one, so that the efficiency of the tin fluorescence will increase.

The quenching of a part of the tin fluorescence between 450 °K and 550 °K in $Sr_3(PO_4)_2$-Sn-Mn is caused by manganese; a further discussion will be given in section 5.2.

Near room temperature the efficiencies of both the tin and the manganese fluorescence are nearly independent of temperature. Therefore the ratios $B$ of the quantum efficiencies of the manganese to the tin fluorescence are measured at room temperature, to test formulae (1), (2) and (3). The results are given in table III; excitation by $\lambda$ 2537 Å. From this table it follows that for the phosphors with constant manganese content the ratio $B$ increases with increasing tin concentration when the tin concentrations are lower than about 4 %. At higher concentrations the ratio $B$ decreases.

This decrease is accompanied by a decrease of the total quantum efficiency of the tin and manganese fluorescences, while also the emission spectra change at high concentrations. These effects are probably caused by a change of the crystal structure at high tin content, which was proved by the X-ray diffraction patterns *). For this reason the results in the system $Sr_3(PO_4)_2$-Sn-Mn are less valuable. But at low tin content, at which the emission spectra are not changed and the total quantum efficiency is nearly constant, the increase of the ratios $B$ with rising tin content for constant manganese content may be considered to be in contradiction to the formulae (1), (2) and (3).

TABLE III

Ratio of the quantum efficiencies of the manganese to the tin fluorescence in $Sr_3(PO_4)_2$-$x$%Sn-$y$%Mn

| $x$ ↑ \ $y$ → | 0·2 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| 0·2 | 0·053 | — | — | — | — |
| 0·5 | 0·18 | — | — | — | — |
| 1 | 0·37 | 3·3 | 4·2 | 9·7 | 3·4 |
| 2 | — | 5·7 | 5·5 | 17 | 8·4 |
| 4 | — | 7·0 | — | 26 | — |
| 5 | — | 7·2 | 7·7 | 26 | 8·5 |
| 6 | — | 7·7 | 10·0 | 10 | 6·3 |
| 8 | — | 6·5 | 11·4 | 3·2 | 2·3 |

3.3. *Calcium fluorophosphate activated with antimony and manganese*

It is well known that $3Ca_3(PO_4)_2.CaF_2$-Sb-Mn shows a blue-green antimony fluorescence with maximum at $\lambda$ 4850 Å and a red manganese fluorescence with maximum at $\lambda$ 5700 Å when excited by ultraviolet radiation ($\lambda < 3000$ Å) [53]), [54]), [55]).

The two emission bands overlap each other strongly as may be seen in fig. 13. Therefore the ratios $B$ of the quantum efficiencies at room temperature of the manganese to the antimony fluorescence were calculated from the emission spectra. In table IV the results are given for excitation by $\lambda$ 2537 Å. It must be pointed out that the antimony absorption band is influenced by the addition of manganese **). This may affect the quantum efficiencies; therefore the values of the ratios $B$ in table IV will be less accurate.

*) Thanks are due to Mr P. B. Braun for carrying out the X-ray work.
**) Dr F. A. Kröger, private communication.

## TABLE IV

Ratio of the quantum efficiencies at room temperature of the manganese to the antimony fluorescence in $3Ca_3(PO_4)_2.CaF_2$-$x$%Sb-$y$%Mn; excitation by $\lambda$ 2537 Å

| $x$ \ $y$ | 0·1 | 0·3 | 0·5 | 0·9 |
|---|---|---|---|---|
| 0·01 | — | — | 0·16 | — |
| 0·1 | 0·023 | 0·09 | 0·19 | 0·265 |
| 0·3 | 0·037 | 0·08 | — | 0·30 |
| 0·7 | — | — | 0·26 | — |
| 2·7 | 0·076 | 0·14 | 0·34 | 0·28 |
| 5·4 | 0·105 | 0·17 | — | 0·32 |

But it is clear that in this system also the ratio $B$ increases with increasing antimony content *) for constant manganese concentrations. The increase is small for the set of phosphors with 0·9 % Mn.

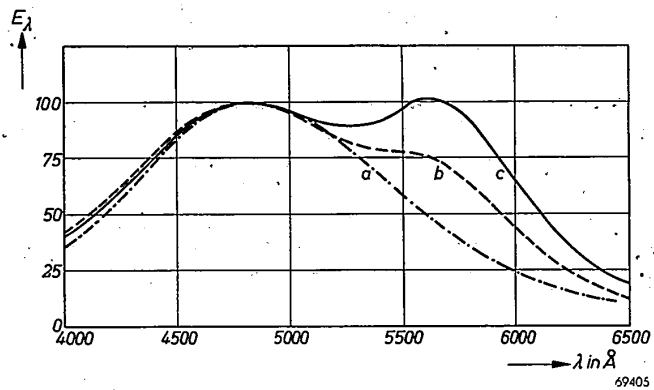



Fig. 13. Emission spectra at room temperature.
(*a*) $3Ca_3(PO_4)_2.CaF_2$-Sb; (*b*) $3Ca_3(PO_4)_2.CaF_2$-0·01%Sb-0·5%Mn;
(*c*) $3Ca_3(PO_4)_2.CaF_2$-2%Sb-0·5%Mn. Excitation by $\lambda$ 2537 Å.

### 3.4. *Calcium pyrophosphate activated with tin and manganese*

$Ca_2P_2O_7$ activated with manganese cannot be excited by ultraviolet radiation. $Ca_2P_2O_7$-Sn shows an absorption band with long-wavelength limit at about $\lambda$ 3000 Å (cf. fig. 14). Excitation in this absorption band gives rise to emission of tin fluorescence. The spectrum consists of various bands; the strongest one has a peak at $\lambda$ 3600 Å, while weaker bands are found in the blue and green.

*) The values are based on the amount of Sb added before the firing.

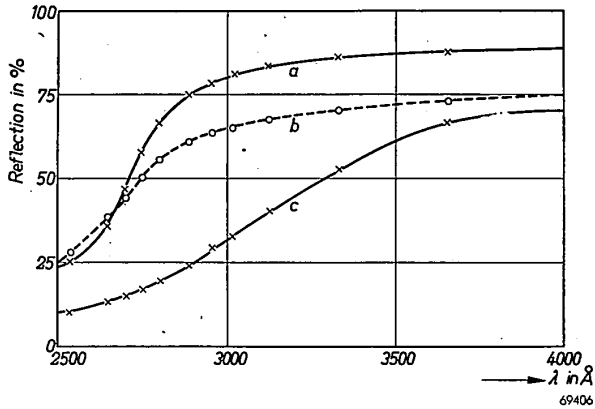



Fig. 14. Reflection spectra at room temperature.
(*a*) $Ca_2P_2O_7$-1%Sn; (*b*) $Ca_2P_2O_7$-1%Sn-1%Mn; (*c*) $Ca_2P_2O_7$-5%Sn-0·5%Mn.

$Ca_2P_2O_7$-Sn-Mn shows an absorption similar to that of the product without manganese. Excitation in this absorption band now leads to emission of both the tin and the manganese fluorescence. The emission spectra are given in fig. 15. The spectrum of the tin fluorescence is nearly the same as that for $Ca_2P_2O_7$-Sn. The position of the manganese emission band depends on the crystal structure of the phosphor. The phosphor with the $\alpha$-modification shows a band with a peak at $\lambda$ 6000 Å, the $\beta$-modification at $\lambda$ 5650 Å [65]), while a third modification shows a band with maximum at $\lambda$ 6500 Å. Since the position of the peaks is very different, the various modifications of the phosphors can roughly be determined from the emission spectra.

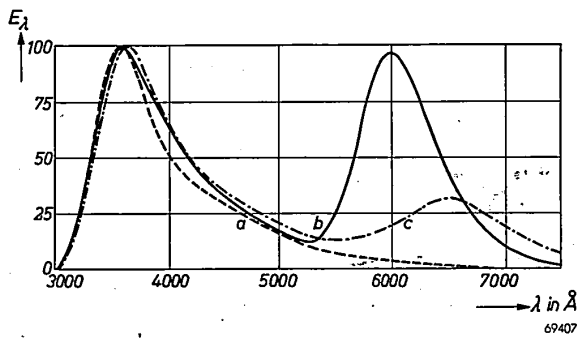



Fig. 15. Emission spectra at room temperature.
(*a*) $Ca_2P_2O_7$-1%Sn; (*b*) $\alpha$-$Ca_2P_2O_7$-10%Sn-1%Mn; (*c*) $Ca_2P_2O_7$-10%Sn-0·5%Mn.
Excitation by $\lambda$ 2537 Å.

The ratios $B$ of the quantum efficiencies of the manganese to the tin fluorescence for some phosphors with the $\alpha$-modification and the third modification are measured at room temperature; excitation by $\lambda$ 2537 Å. For $\alpha$-$Ca_2P_2O_7$-$x$%Sn-1%Mn the ratio $B$ with $x = 0{\cdot}5$ and $x = 5$ was found to be 0·57 and 1·14 respectively. For $Ca_2P_2O_7$-$x$%Sn-0·5%Mn with a peak of the manganese fluorescence band at $\lambda$ 6500 Å the ratio $B$ with $x = 1{\cdot}5$ and 10 was respectively 0·30, 0·35 and 0·53.

Thus in these cases too the ratio $B$ increases with increasing tin and constant manganese concentrations, in contradiction to formulae (1), (2) and (3).

### 3.5. *Calcium carbonate activated with lead or thallium and manganese*

$CaCO_3$ activated with manganese cannot be excited by ultraviolet radiation, but $CaCO_3$ activated with lead, thallium or cerium and manganese (prepared according to Fonda [66]) shows an orange-red manganese fluorescence when excited by ultraviolet radiation at room temperature [47]).

Fig. 16 shows the emission spectra of the manganese fluorescence in $CaCO_3$-(Pb or Tl)-Mn at various temperatures for excitation by $\lambda$ 2537 Å. The emission spectra are found to be the same for sensitization by lead as by thallium. Probably the emission consists of at least three sub-bands with peaks at about 6050 Å, 6600 Å and 7000 Å.

In fig. 17 the emission spectra of $CaCO_3$-Pb are given at various temperatures; excitation by $\lambda$ 2537 Å. The maximum of the emission band shifts from $\lambda$ 3300 Å at 20 °K to $\lambda$ 3050 Å at 300 °K.

In $CaCO_3$-Pb-Mn the lead emission band is shifted slightly towards shorter wavelengths compared with that in $CaCO_3$-Pb; however, the intensity of the lead emission in $CaCO_3$-Pb-Mn was very low, so that the spectrum is not accurate.

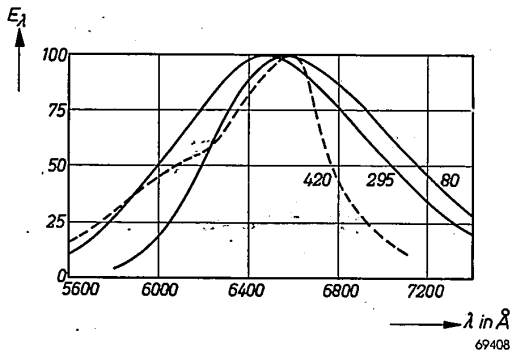



Fig. 16. Emission spectra of the manganese fluorescence in $CaCO_3$-(Pb, Tl)-Mn at 80, 295 and 420 °K; excitation by $\lambda$ 2537 Å.

The system $CaCO_3$-Pb-Mn was found to be unsuitable for testing formulae (1), (2) and (3) because in $CaCO_3$-$x$%Pb-0·5%Mn the intensity of the lead fluorescence was very low and with lower manganese content the intensity of the manganese fluorescenec was too low to be measured with sufficient accuracy.

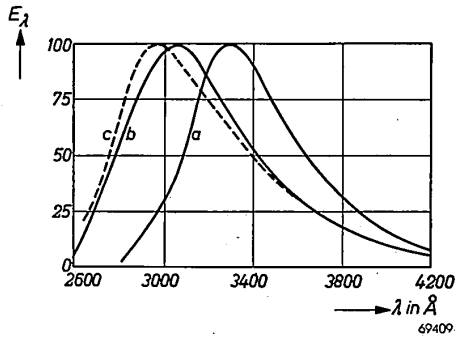



Fig. 17. Emission spectra of the lead fluorescence of (*a*) $CaCO_3$-Pb at 80 °K; (*b*) $CaCO_3$-Pb at 300 °K; (*c*) $CaCO_3$-Pb-Mn at 300 °K. Excitation by $\lambda$ 2537 Å.

Neither was the system $CaCO_3$-Tl-Mn suitable for testing (1), (2) and (3) because the intensity of the thallium fluorescence was very low. It is remarkable that in $CaCO_3$-Tl, with various thallium contents up to 4%, the intensity of the thallium fluorescence was very low, whereas in $CaCO_3$-4%Tl-3%Mn the intensity of the manganese fluorescence was found to be much stronger than the intensity of the thallium fluorescence in $CaCO_3$-Tl. Presumably thallium acts fairly well as sensitizer but is a bad activator.

### 3.6. *Magnesium pyrophosphate activated with cerium, thorium and manganese according to the investigations of Froelich*

Froelich [56]) showed that $Mg_2P_2O_7$-Th-Mn emits neither ultraviolet nor visible fluorescence, while $Mg_2P_2O_7$-Ce-Mn emits a long-wavelength ultraviolet cerium fluorescence and a very faint red manganese fluorescence when excited by $\lambda$ 2537 Å. For both phosphors the results are the same at room temperature as at liquid-air temperature.

Excitation of $Mg_2P_2O_7$-Ce-Th-Mn by ultraviolet radiation gives rise to an ultraviolet fluorescence due to cerium or thorium (or both) and a red fluorescence due to manganese. Froelich proposed that in this system cerium acts as sensitizer for thorium and thorium in its turn as sensitizer for manganese. Thus the energy absorbed in cerium will be transferred to thorium and then from thorium to manganese.

## TABLE V

Ratio of the intensities of the manganese to the ultraviolet fluorescence in $Mg_2P_2O_7$-Ce-Th-Mn. Intensities of both the manganese and the ultraviolet fluorescence in arbitrary units *); excitation by $\lambda$ 2537 Å

| nr of phosphor | moles per 100 moles $Mg_2P_2O_7$ of | | | intensity Mn fluor. | intensity *UV* fluor. | ratio $\frac{\text{Mn fluor.}}{\text{UV fluor.}}$ |
|---|---|---|---|---|---|---|
| | $Ce_2O_3$ | $ThO_2$ | MnO | | | |
| P 270 | 1·46 | 8·4 | 4·2 | 73 | 32 | 2·3 |
| P 280 | 1·50 | 12·6 | 4·3 | 91 | 27 | 3·4 |
| P 251 | 1·46 | 8·4 | 8·4 | 70 | 14 | 5·0 |
| P 271 | 3·2 | 8·4 | 8·4 | 69 | 10 | 6·9 |
| P 240 | 13·4 | 1·74 | 28·2 | 37 | 56 | 0·7 |
| P 245 | 13·6 | 8·6 | 28·2 | 80 | 14 | 5·7 |
| P 237 | 13·4 | 17·0 | 28·2 | 67 | 1 | 67 |

*) Results copied from Froelich.

A number of results were taken from Froelich's table 2 and the ratios of intensities of the manganese to the ultraviolet fluorescence have been calculated from them; the results are given in table V. From the last column of this table it follows that for constant manganese content the ratio of the intensities of the manganese to the ultraviolet fluorescence increases with increasing cerium and constant thorium contents and also with increasing thorium and constant cerium contents. Thus the ratio of intensities increases whether the ultraviolet fluorescence arises from cerium or from thorium. Although the values of the ratios in table V have to be considered with some reserve, because the transfer mechanism is rather complex and the emission spectra vary a little with the concentrations (Froelich [56])), it may be concluded that the results in $Mg_2P_2O_7$-Ce-Th-Mn give a strong indication that the variation of the ratio of the activator to the sensitizer fluorescence does not agree with formulae (1), (2) and (3).

The experiments of Froelich are the only ones published about sensitized solid systems that could be used to test formulae (1), (2) and (3).

The system $CaSiO_3$-Pb-Mn, showing transfer of energy from lead to manganese, was found to be unsuitable for this purpose. For, according to Schulman [45]), an extra absorption band appears both with increasing lead and increasing manganese concentrations, while Froelich [48]) showed that at higher lead and manganese concentrations the $\beta$-wollastonite structure appears in addition to the $\alpha$-modification. The results would therefore be very doubtful.

### 3.7. *Conclusion*

From the experiments carried out on $Ca_3(PO_4)_2$-Ce-Mn, $Sr_3(PO_4)_2$-Sn-Mn, $3Ca_3(PO_4)_2.CaF_2$-Sb-Mn, $Ca_2P_2O_7$-Sn-Mn and those carried out by Froelich on $Mg_2P_2O_7$-Ce-Th-Mn it is proved, in the one system less convincingly than in the other, that the ratio of the efficiencies of activator to sensitizer fluorescence increases with increasing sensitizer concentration at constant activator content. These results are in contradiction to the formulae (1), (2) and (3). The theories from which the formulae are deduced are, therefore, either incorrect or incomplete. In the next section an attempt is made to meet this difficulty.

## 4. Explanation of the energy transfer in sensitized systems

First the possibility will be considered of explaining the energy transfer by resonance processes. The possibility with the long-wave exciton (cf. section 1) will be discussed later.

So far only resonance between sensitizer and activator ions has been considered. Botden and Kröger [52]) mentioned that transfer of energy from a sensitizer to an activator could perhaps take place via other sensitizer ions through resonance between sensitizer ions.

Förster [11]) proved that resonance exists between equivalent molecules in solutions. Here it will be discussed whether resonance between sensitizer ions plays a rôle in the energy transfer in solid systems.

### 4.1. *Transfer of energy from one excited sensitizer ion to another*

Förster [11]) deduced a relation between the critical distance $R_0$, over which energy transfer through resonance can take place between equivalent molecules, and the half-value width $\Delta\nu$ of the overlap of the fluorescence and the absorption spectrum of a fluorescent material (cf. fig. 18). This relation is:

$$R_0 \approx \frac{\lambda_0}{2\pi n}\left\{\frac{3}{8}\frac{\eta_0^2}{\tau_0 c\Delta\nu}\right\}^{1/6}, \qquad (4)$$

in which $\lambda_0$ = the average value of the wavelengths of the maxima for the absorption and emission band; $n$ = index of refraction; $\eta_0$ = quantum efficiency; $\tau_0$ = lifetime of the excited state; $c$ = velocity of light.

Formula (4) may be applied to solid systems as well. When resonance between sensitizer ions is considered, the critical distance $R_0$ between sensitizer ions can be calculated. In $Ca_3(PO_4)_2$-Ce-Mn the $R_0$ for cerium ions is calculated to be about 35 Å (with $n = 1$; $\eta_0 = 1$; $\lambda_0 = 3200$ Å; $\Delta\nu = 1000$ cm$^{-1}$; $\tau_0 = 10^{-7}$ sec; $c = 3.10^{10}$ cm sec$^{-1}$).

In $CaSiO_3$-Pb-Mn (from the experiments of Schulman [45]) and of Botden-Kröger [52])) the $R_0$ for lead ions is found to be about 30 Å (with $n = 1$; $\eta_0 = 1$;

$\lambda_0 = 2850$ Å, $\Delta\nu = 1200$ cm$^{-1}$, $\tau_0 = 10^{-6}$ sec). In the other systems mentioned in section 3 the overlap of the absorption and emission spectra is smaller and rather difficult to determine, but because of the exponent 1/6 in (4) the $R_0$ does not change rapidly with $\Delta\nu$.

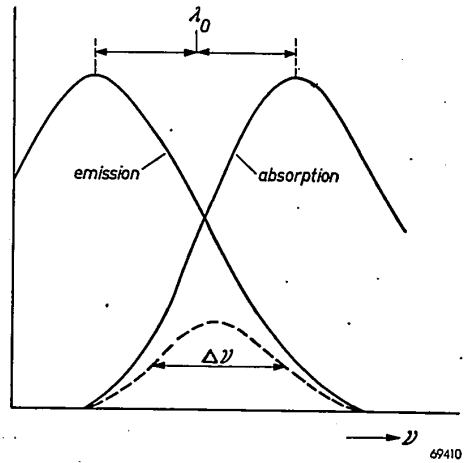



Fig. 18. Spectral distributions of emission, absorption and the product of emission and absorption (dashed curve), showing the $\Delta\nu$.

Thus according to the theory of Förster, transfer of energy from one sensitizer to another may take place over distances of the order of 30 Å. This means that around an excited sensitizer some hundreds of cation sites will be found, such that the energy may be transferred from the sensitizer to another sensitizer on one of those cation sites. According to the theory of Kalmann-London [6]) the critical distance over which transfer of energy from the one atom to another may take place increases when the difference in energy of the excited levels of the atoms is lowered. This difference in energy is expected to be low for the excited levels of two sensitizer ions and probably is lower than between the excited levels of a sensitizer and an activator ion.

It will, therefore, be assumed that energy transfer takes place from an excited sensitizer ion to an activator ion not only directly through resonance between these ions but also via other sensitizer ions. In this way the energy absorbed in a sensitizer at great distances from an activator may reach this activator via other sensitizer ions ("energy transport").

A random distribution of the sensitizer and activator ions will be assumed because no indication could be found that the formation of sensitizer-

activator pairs is stimulated (section 1). On the basis of these assumptions a relation between the quantum efficiencies of the sensitizer and activator fluorescences will now be deduced.

4.2. *Theoretical relation between the quantum efficiencies of fluorescence of the activator and sensitizer as a function of their concentrations, based on the assumptions in section 4.1*

A sensitizer will be denoted by $S$ and an activator by $A$. The contribution to the $A$ fluorescence will be calculated for absorption of excitation energy in:

(a) the $S$ that can transfer the energy directly to an $A$ (process $S_0 \rightarrow A$);
(b) the $S$ from which the energy can reach an $A$ only via one other $S$ (process $S_0 \rightarrow S_1 \rightarrow A$);
(c) the $S$ from which the energy can reach an $A$ only via two other $S$ ions (process $S_0 \rightarrow S_1 \rightarrow S_2 \rightarrow A$), etc.

Let $k$ be the number of cation sites around an $S$ such that transfer of energy takes place from $S$ to $A$ when this $A$ is situated on one of the $k$ sites. The probability of the energy transfer is assumed to be 1, so that absorption of excitation energy in the $S$ with an $A$ on one of the $k$ surrounding sites leads to $A$ fluorescence, while absorption in the remaining $S$ may lead to $S$ fluorescence (absorption in quenching centres is left out of account, and it is assumed that no temperature-quenching of the fluorescence takes place).

That the probability of direct energy transfer is close to 1 follows from:
(1) The high fraction of absorbed energy that is transferred from $S$ to $A$ at sufficiently high concentrations of $A$; the ratio of quantum efficiencies of $A$ fluorescence to $S$ fluorescence in, for example, $Ca_3(PO_4)_2$-5%Ce-2·5%Mn is found to be 5·5, in $CaSiO_3$-0·7%Pb-2%Mn [52]) it is 6, while in $Ca_3(PO_4)_2$-Sn-Mn, $Sr_3(PO_4)_2$-Sn-Mn and $3Ca_3(PO_4)_2 \cdot CaF_2$-Sb-Mn even higher values are found.
(2) The decay measurements on some sensitized phosphors carried out by Kröger [59]).

Let $X_a$ and $X_s$ be the respective ratios of the numbers of $A$ and $S$ to the number of cation sites. Finally let $Z$ be the efficiency of energy transfer i.e. the ratio of the quantum efficiencies of $A$ fluorescence to $(S + A)$ fluorescence *). For the sake of simplicity this $Z$ will be calculated instead of the ratio of quantum efficiencies of $A$ to $S$ fluorescence.

With a random distribution of $S$ and $A$ the probability of finding an $A$ on a particular site of the $k$ sites around an $S$ is $X_a$. Thus the probability that an $A$ is not found on that particular site is $(1 - X_a)$. Then the probability $P_1$ that there is no $A$ on any one of the $k$ sites is $P_1 = (1 - X_a)^k$ and the probability $P_2$ of finding at least one $A$ on the $k$ sites is $P_2 = 1 - (1 - X_a)^k$. This derivation was given by Schulman [45]).

*) Thus $Z$ includes both direct energy transfer and energy transport.

The contribution to $Z$ through the process $S_0 \rightarrow A$, mentioned in (a), is now given by the ratio of $P_2$ to $(P_1 + P_2)$, thus by

$$1 - (1 - X_a)^k. \tag{5}$$

Let $l$ be the number of cation sites around an $S$ such that transfer of energy can take place from that $S$ to another $S$ on one of the $l$ sites and let $p$ be the probability of transfer. Then the contribution to $Z$ through process $S_0 \rightarrow S_1 \rightarrow A$ (b) is found to be

$$\underbrace{(1 - X_a)^k}_{\text{I}} \quad \underbrace{p \{1 - (1 - X_s)^l\}}_{\text{II}} \quad \underbrace{\{1 - (1 - X_a)^{k'}\}}_{\text{III}}, \tag{6}$$

in which I = the probability that there is no $A$ on any of the $k$ sites around the primarily excited $S_0$; II = the probability that an $S_1$ is found on one of the $l$ sites around $S_0$ (analogous to $P_2$) and the energy is transferred from $S_0$ to $S_1$; III = the probability that an $A$ is found on one of the $k'$ sites around $S_1$ so that the energy reaching $S_1$ is transferred to $A$. III is similar to (5), but $k$ is replaced by $k'$ because some of the $k$ sites around $S_1$ may be among the $k$ sites around $S_0$; the number of overlapping sites is $(k - k')$.

In a similar way the contribution to $Z$ through process $S_0 \rightarrow S_1 \rightarrow S_2 \rightarrow A$ (c) is found to be

$$\underbrace{(1 - X_a)^k \, p \{1 - (1 - X_s)^l\} \, (1 - X_a)^{k'}}_{\text{I}} \quad \underbrace{p \{1 - (1 - X_s)^l\}}_{\text{II}} \quad \underbrace{\{1 - (1 - X_a)^{k''}\}}_{\text{III}}, \tag{7}$$

in which: I = the probability that energy is transferred from $S_0$ to $S_1$ on one of the $l$ sites around $S_0$ and that an $A$ is not found on one of the $k$ or $k'$ sites around $S_0$ or $S_1$; II = the probability that an $S_2$ is found on one of the $l$ sites around $S_1$ and that the energy is transferred from $S_1$ to $S_2$; III = the probability that an $A$ is situated on one of the $k''$ sites around $S_2$, $k''$ being $k$ reduced by the number of sites belonging to the $k$ and $k'$ sites around $S_0$ and $S_1$.

In a similar way the contributions to $Z$ through the processes $S_0 \rightarrow S_1 \rightarrow S_2 \rightarrow S_3 \rightarrow A$, etc., can be calculated. Addition of (5), (6), (7) gives

$$= k X_a - \frac{k(k-1)}{2} X_a^2 + plk' X_a X_s + \frac{k(k-1)(k-2)}{3!} X_a^3 - \\ - X_a^2 X_s plk' \left(\frac{k'-1}{2} + k\right) - X_a X_s^2 pl \left(\frac{k'(l-1)}{2} - plk''\right) + \cdots. \tag{8}$$

When $X_a$ is small compared with $1/k$, and $X_s$ is small compared with $1/l$, the terms in (7) become smaller with increasing powers of $X_s$ and $X_a$. Let $Z_1$, $Z_2$ and $Z_3$ be the successive approximations to $Z$, where the index

is the highest value of the sum of exponents of $X_a$ and $X_s$. Then we find:

$$Z_1 = k\,X_a\,, \tag{9}$$

$$Z_2 = k\,X_a - \frac{k\,(k-1)}{2}\,X_a^2 + plk'\,.\,X_aX_s, \tag{10}$$

$$Z_3 = kX_a - \frac{k(k-1)}{2}\,X_a^2 + \frac{k(k-1)\,(k-2)}{3!}\,X_a^3 + \\ + plk'X_aX_s\left\{1 - X_a\left(\frac{k'-1}{2} + k\right) + X_s\left(\frac{plk''}{k'} - \frac{l-1}{2}\right)\right\}. \tag{11}$$

According to (9) $Z_1$ is independent of $X_s$, but in (10) we see that $Z_2$ increases with $X_s$ when $X_a$ is kept constant, while according to (11) this increase is not linear with $X_s$.

In all the experiments mentioned in section 3 an increase of $Z$ with increasing $X_s$ and constant $X_a$ is found. This is illustrated for example in fig. 19, where $Z$ is plotted against cerium concentration for $Ca_3(PO_4)_2$-Ce-Mn with constant manganese concentrations (values of $Z$ are given in table I), and in fig. 20 for the phosphors $Ca_3(PO_4)_2$-Ce-Mn mentioned in table II.

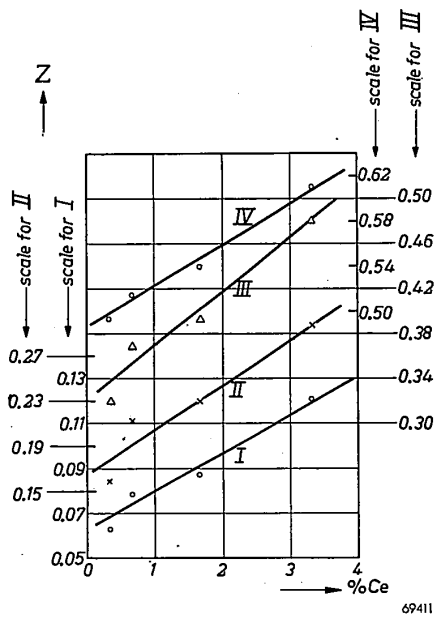



Fig. 19. Efficiency of energy transfer $Z$ as a function of the cerium concentration for $Ca_3(PO_4)_2$-Ce-y%Mn. I for $y = 1/6$; II for $y = 1/2$; III for $y = 1$; IV for $y = 5/3$. Excitation by $\lambda$ 2537 Å.

In fig. 20 it can be seen that the slope of the curve becomes flatter with rising cerium content. This is not the case with the curves for the phosphors in fig. 19 because these phosphors were prepared in sets with constant cerium and varying manganese contents, so that the relative values of $Z$ belonging to each curve are less accurate than those in fig. 20.

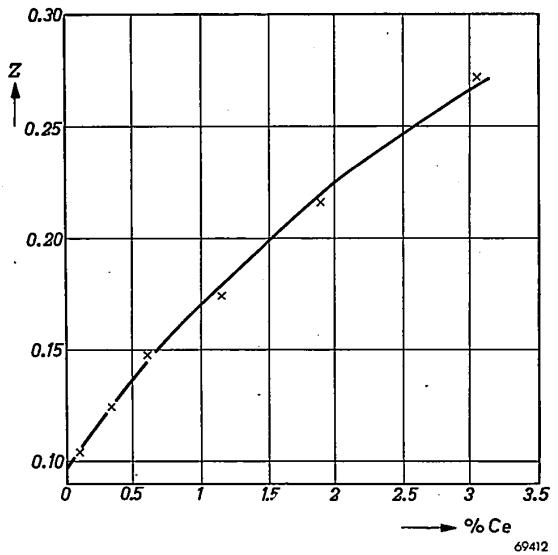



Fig. 20. Efficiency of energy transfer $Z$ as a function of the cerium concentration for $Ca_3(PO_4)_2$-Ce-1/3%Mn. Excitation by $\lambda$ 2537 Å.

### 4.3. *Calculation of k, l and p from the experiments*

Because the $X_s$ and $X_a$ are too high to use formula (9), since the $Z$ is found to be dependent on $X_s$, we shall start with formula (10). Then for $X_s = 0$ we obtain

$$Z_{(X_s=0)} = kX_a - \frac{k(k-1)}{2} X_a^2,$$

and hence

$$k = (1/X_a)\{1 - (1 - 2Z_{(X_s=0)})^{1/2}\}, \qquad (12)$$

in which $(\frac{1}{2} + 1/X_a)$ is replaced by $1/X_a$. Thus the experimental value of $Z_{(X_s=0)}$ for a given $X_a$ determines $k$. Then the number of overlapping sites of the $k$ sites around an $S$ and a neighbouring $S$ can be calculated at a given $X_s$; thus $k'$ and in a similar way $k''$, can be found. Neglecting derivatives of $k'$, $k''$ and higher powers of $X_s$, the first and second derivatives of $Z$ to

$X_s$ are approximately

$$\frac{dZ}{dX_s} = plk'X_a \tag{13}$$

and

$$\frac{d^2Z}{dX_s^2} = 2\,plk'X_a\left\{\frac{plk''}{k'} - \frac{l-1}{2}\right\}. \tag{14}$$

Thus the slope of the experimental curves as given in figs 19 and 20, determines the value of $plk'$ for a given $X_a$, and then from the calculated $k'$ the value of $pl$ is found. The $p$ and $l$ can be found separately from the change of the slope $d^2Z/dX_s^2$.

Calculations have been carried out for $Ca_3(PO_4)_2$-Ce-Mn from the figs 19 and 20. In table VI the values of $k$ and $plk'$ are given; the value of $plk'$ for the phosphor with 1/3%Mn is calculated from the part of the curve in fig. 20 at low cerium concentrations.

TABLE VI

Values of $k$ and $plk'$ for sets of phosphors with varying cerium and constant manganese concentrations

| phosphor | $k$ | $plk'$ |
|---|---|---|
| $Ca_3(PO_4)_2$-Ce-1/6%Mn | 37 | 1370 |
| $Ca_3(PO_4)_2$-Ce-1/2%Mn | 34 | 870 |
| $Ca_3(PO_4)_2$-Ce-1%Mn | 40 | 480 |
| $Ca_3(PO_4)_2$-Ce-5/3%Mn | 50 | 270 |
| $Ca_3(PO_4)_2$-Ce-1/3%Mn from fig. 20 | 30 | 2500*) |

*) Value at low cerium concentration.

The values obtained from fig. 19 will be less accurate because they are calculated from an average slope (the curves are given as straight lines). This may be the reason why the value of $k$ at high manganese content (5/3 %) is increased; the correction for $k$ is only 10% when $k$ is calculated by using formula (11).

From table VI it may be concluded that in $Ca_3(PO_4)_2$-Ce-Mn the value of $k$ is 30 to 35. The value of $k'$ decreases with increasing cerium content because the overlap of the $k$ sites around two neighbouring cerium ions becomes larger. For a random distribution of the cerium ions and a value of $k = 35$ this overlap $(k - k')$ is calculated to be about 0, 2, 8 and 14 for $X_s = 1/3, 2/3, 5/3$ and $10/3$ respectively, so that $k' = 35, 33, 27$ and 21.

Owing to this decrease of $k'$ with increasing cerium content the slope of the curves, giving $Z$ as a function of $X_s$ at constant $X_a$, will decrease according to (13) and (14). But at low cerium concentrations this effect may be neglected. Then the calculation of $p$ and $l$ from fig. 20 with the aid of (14) gives $p \approx 0{\cdot}5$ and $l \approx 170$.

The decrease of the values $plk'$ in table VI obtained from fig. 19 is probably not only due to a decrease of $k'$ with increasing Mn content but also to the inaccuracy of the relative values of $Z$ for each set of phosphors with varying cerium and constant manganese concentrations, so that only an average slope of the curve in fig. 19 could be given; this slope decreases with increasing Mn content. Apparently the values of $Z$, and thus also of $k$, $l$ and $p$, depend on the preparation of the products. Therefore the calculated values of $k$, $l$ and $p$ have to be considered as giving only the order of magnitude.

It will be clear that similar calculations can be given for constant $X_s$ and varying $X_a$. For this in $Ca_3(PO_4)_2$-Ce-Mn the values of $Z$ in table I can be used. From (8) it follows that for constant $X_s$ the $Z$ increases linearly with $X_a$, at low $X_a$ and $X_s$. At higher values of $X_a$ the increase is lower and this effect is stronger at high values of $X_s$. This may be seen in fig. 21, where $Z$ is given as a function of $X_a$ for $Ca_3(PO_4)_2$-Ce-Mn. In this figure the curves are more accurate than in fig. 19, because the sets of

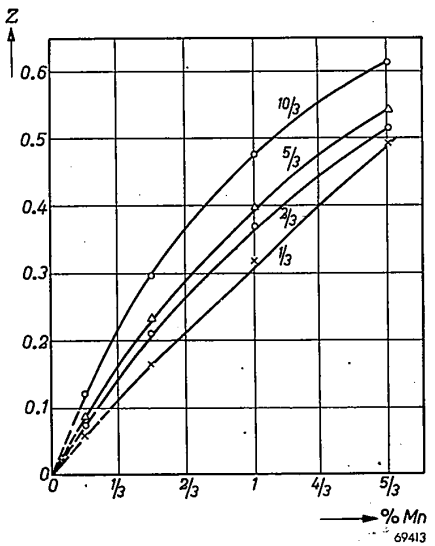



Fig. 21. Efficiency of energy transfer $Z$ as a function of the manganese concentration for $Ca_3(PO_4)_2$-$x$%Ce-Mn with $x = 1/3$, $2/3$, $5/3$ and $10/3$. Excitation by $\lambda$ 2537 Å.

phosphors with constant manganese and varying cerium concentrations were prepared together. From fig. 21 the values of $k$ can be calculated in a similar way as above and are found to be about the same.

Calculations have been carried out also for $3Ca_3(PO_4)_2.CaF_2$-Sb-Mn, with the result $k = 26$ and $pl = 60$, and for $Ca_2P_2O_7$-Sn-Mn with the result $k = 43$ and $pl = 10$. The accuracy of the experimental results was not high enough to calculate $p$ and $l$ separately. If $p$ is assumed to be of the order of 0·5 the value of $l$ is low in $Ca_2P_2O_7$-Sn-Mn; apparently transfer of energy from the excited $S$ to another $S$ is less probable than in the other systems.

4.4. *Discussion of the results*

The results of calculations are: about 30 cation sites are found around an $S$ such that transfer of energy takes place from $S$ to an $A$ situated on one of the 30 sites; about 150 sites are found around an $S$ such that energy transfer may take place from the excited $S$ to another $S$ situated on one of the 150 sites, and the probability of transfer is about 0·5.

The 150 sites around an $S$ available for transfer to another $S$ are of the order of magnitude expected according to the theory of Förster (cf. section 4.1). Schulman [45]) calculated the number of sites around an $S$ available for transfer of energy from $S$ to an $A$, from absorption measurements in $CaSiO_3$-Pb-Mn, and he found it to be 28. This is about the same as the values calculated above for other phosphors. Therefore on the basis of the assumptions of section 4.1 not only the experiments can be explained qualitatively, but also the calculated values of the number of sites around an $S$ available for energy transfer to another $S$ and to an $A$ are very plausible.

The explanation of the sensitization, given above, is an extension of the theory of Schulman. The formation of pairs seems to be unlikely. The experiments of Kröger [59]) (compare section 1), however, are not in contradiction with the explanation described here if it is assumed that a sensitizer and an activator may influence each other when the activator is situated on one of (about) 30 cation sites around the sensitizer. An $S$ at larger distances from an $A$ can transfer the energy to an $A$ without being directly influenced by the $A$.

If transport of energy takes place through long-wave excitons (section 1) then the efficiency of energy transfer cannot increase with increasing sensitizer and constant activator concentrations; no agreement with the experiments can be obtained. Therefore it can be excluded that transfer of energy in sensitized phosphors is caused by long-wave excitons.

When "killers" (quenching centres) are added instead of activator ions

a similar mechanism of energy transfer and energy transport may take place: quenching of the fluorescence will increase with increasing concentration of the "sensitizer" ions (in these systems called "activator" ions) owing to the energy transport via "sensitizer" ions to the killer. An indication for this may be found in the paper of Johnson and Williams [67]) on concentration quenching in various systems.

Probably in organic fluorescent solutions a similar mechanism of sensitization will occur as in solids: owing to the energy transport the efficiency of energy transfer will increase with increasing sensitizer and constant activator concentrations. Recently Moodie and Reid [68]) showed that the intensity of the naphthacene fluorescence, sensitized by anthracene, increases strongly when the concentration of anthracene is increased, whereas the intensity of the anthracene fluorescence is nearly the same. This can be explained in the same way as described here for solids.

More experiments suitable for this purpose could not be found in the literature.

### 4.5. *Summary of section 4*

An explanation of the phenomena of sensitization in solids is given on the basis of the assumptions that:
(a) around a sensitizer $k$ cation sites are found such that excitation energy is transferred from the sensitizer to an activator ion, if this ion is situated on one of the $k$ sites;
(b) when no activators are found on the $k$ sites around an excited sensitizer the energy may reach an activator through energy transport via other sensitizers, $l$ sites being available for energy transfer from the excited sensitizer to another one, situated on one of the $l$ sites. Transfer of energy from a sensitizer both to an activator and to another sensitizer takes place through resonance processes.

Some formulae are deduced to calculate the $k$ and $l$ from the experimental results. The value of $l$ is found to be of the order of magnitude expected according to the theory of Förster. The value of $k$ is in agreement with that given by Schulman, who calculated the $k$ from absorption measurements.

Similar explanations may be given for systems in which killers are added and for organic sensitized systems.

## 5. Theoretical treatment of the transfer processes in sensitized phosphors on the basis of the model of Mott and Seitz

In this section we shall discuss the possibility of applying the model of Mott and Seitz to the sensitized systems. In this way the temperature dependence of the efficiency of energy transfer will be explained.

5.1. *The model of Mott and Seitz applied to sensitized phosphors*

This will be done in a similar way as in part I. In fig. 22 a schematic representation is given of the potential-energy curves as a function of the configurational coordinate. Excitation of an $S$ with an $A$ on one of the neighbouring cation sites (transition $e$) can either give rise to $S$ fluorescence (transition $s$) or to transfer of energy to $A$ through resonance at $R$. According to section 4.2 the probability of the resonance transition is higher than that of transition $s$. Transfer to $A$ will be followed by emission of $A$ fluorescence (transition $a$). When excitation energy is absorbed in an $S$ at larger distances from an $A$, then the energy has to be transported to an $S$ with an $A$ in its neighbourhood, before transfer to $A$ can take place. In fig. 22 this is schematically given by a transition from $S^*_{far}$ to $S^*_{next}$ at $T$, which transition may be followed by a transition to $A^*$ at $R$ and finally by transition $a$.

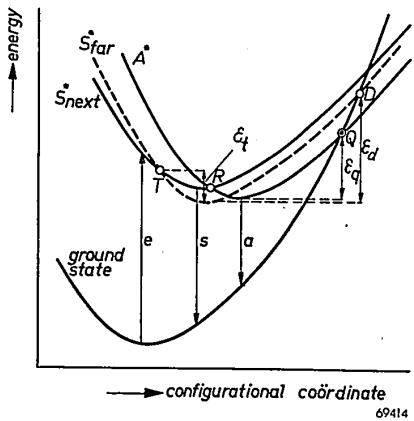



Fig. 22. Schematic representation of the potential-energy curves of the ground state and excited states (marked by asterisks) of a sensitized phosphor. $S$ = sensitizer, $A$ = activator. The curve $S^*_{next}$ belongs to an $S$ with an $A$ on one of the neighbouring cation sites, the curve $S^*_{far}$ to an $S$ that has not an $A$ in its neighbourhood.

The excited state of manganese responsible for manganese emission will not appear in the resonance process between cerium and manganese ions in $Ca_3(PO_4)_2$-Ce-Mn. For, the probability of the fluorescence transition in manganese (decay time $\approx 10^{-2}$ sec) is much lower than that in cerium (decay time $\approx 10^{-7}$ sec), so that the probability of energy transfer would be low, contrary to the experiments. Therefore resonance will take place between an excited cerium level and an excited manganese level higher than that responsible for manganese fluorescence: the probability of the transition from the higher to a lower excited state of manganese may be large compared with the probability of the fluorescence transition in cerium. Whether resonance occurs between this higher state of manganese and the excited state of cerium responsible for the cerium fluorescence — or a higher excited state

of cerium — cannot be concluded from the known experiments *). The same holds for other sensitized phosphors with short decay times of the sensitizer fluorescence.

Only one excited state of manganese is given in fig. 22.

Transfer of energy from $S$ to $A$ takes place even at 20 °K, as was shown in figs 8 and 12 for $Ca_3(PO_4)_2$-Ce-Mn and $Sr_3(PO_4)_2$-Sn-Mn; also in $CaSiO_3$-Pb-Mn and $Ca_2P_2O_7$-Sn-Mn the intensity of the manganese fluorescence was visually the same at 20 °K as at room temperature. Thus in these cases the transfer of energy from $S$ to $A$ needs no activation energy (or a very low one).

In fig. 22 this is expressed by placing $R$ near the minimum of the curve $S^*_{next}$.

It is shown that in $Ca_3(PO_4)_2$-Ce-Mn from about 150 °K up to 350 °K the efficiency of the cerium fluorescence decreases whilst the manganese fluorescence increases, the total efficiency being nearly constant. Thus in this temperature range the efficiency of energy transfer is temperature dependent. It is natural to assume that the temperature dependent part of energy transfer is due to energy transport: the energy absorbed in a cerium ion in the neighbourhood of a manganese ion will be transferred to manganese without an activation energy, while the energy transport from a cerium ion at larger distances from a manganese ion, via other cerium ions to manganese will need an activation energy ($\varepsilon_t$ in fig. 22). Consequently the efficiency of energy transfer has to be independent of the cerium concentration at low temperatures. In table VII it may be seen that the efficiencies of energy transfer at 20 °K for the phosphors with constant manganese and various cerium contents, although not equal in all cases, are indeed much less dependent on the cerium content than the values of transfer efficiencies at 295 °K. The deviations at 20 °K are probably due to the effect of self-absorption (section 3.1), while small differences in the preparation of the products and the less accurate measurements at 20 °K may influence the results. The activation energy $\varepsilon_t$ will be calculated in the next section.

It may well be that $\varepsilon_t$ in other systems is reached only at high temperatures, at which the activator fluorescence is already quenched. In this case a temperature range may be found in which the efficiency of the sensitizer fluorescence decreases with temperature while the increase of the activator fluorescence is not found because of the quenching. In this way the decrease of the efficiency of the tin fluorescence in $Sr_3(PO_4)_2$-Sn-Mn between 450 °K and 550 °K (fig. 12) could be explained. But then the increase of the

*) Possibly this can be concluded from measurements of the decay of the cerium fluorescence. If the decay is faster when manganese is added this suggests that resonance occurs between the lowest excited state of cerium and an excited state of manganese. These experiments are made difficult by the very short decay time of the cerium fluorescence.

efficiency of energy transfer with increasing tin and constant manganese contents at room temperature cannot be understood. Either these values of efficiencies of energy transfer are incorrect owing to changes in the crystal structure (cf. section 3.2) or the decrease of the tin fluorescence between 450 °K and 550 °K is caused by another effect. It is difficult to decide.

TABLE VII

| $Ca_3(PO_4)_2$ - $x$ % Ce - $y$ % Mn | | Efficiency of energy transfer | |
|---|---|---|---|
| $x$ | $y$ | at 295 °K | at 20 °K |
| 5/3 | 5/3 | 0·54 | $0{\cdot}43^5$ |
| 1/3 | 5/3 | 0·49 | 0·44 |
| 5/3 | 1 | 0·39 | $0{\cdot}27^5$ |
| 1/3 | 1 | 0·32 | 0·24 |
| 5/3 | 1/6 | 0·087 | 0·052 |
| 1/3 | 1/6 | 0·063 | 0·040 |
| 3·05 | 1/3 | $0{\cdot}27^5$ | 0·15 |
| 1·10 | 1/3 | 0·175 | $0{\cdot}14^5$ |
| 0·35 | 1/3 | 0·122 | $0{\cdot}10^5$ |

At high temperatures the sensitizer and the activator fluorescence are quenched through radiationless transitions from the excited states to the ground state at $D$ and $Q$ respectively (fig. 22). To reach the energy at point $D$ an activation energy $\varepsilon_d$ is needed and for $Q$ an activation energy $\varepsilon_q$. Whether the sensitizer or the activator fluorescence is first quenched with rising temperatures depends on $\varepsilon_q$, $\varepsilon_d$ and the transition probabilities to the ground state in $Q$ and $D$. It may well be that the sensitizer fluorescence is quenched at lower temperatures than the activator fluorescence. But the opposite is found in $Ca_3(PO_4)_2$-Ce-Mn, $Sr_3(PO_4)_2$-Sn-Mn, $CaSiO_3$-Pb-Mn and $Sr_3(PO_4)_2$-Ce-Mn, as may be seen from figs 12 and 23. In $Ca_3(PO_4)_2$-Ce-Mn the cerium and the manganese fluorescence are quenched in the same temperature range.

The temperature dependence of the intensities of fluorescence, shown in fig. 23, was obtained by measuring at various temperatures the relative photo-currents due to the fluorescence. The fluorescence of the sensitizer and of the activator is measured separately with the aid of suitable filters. As photocell an I.P.28 R.C.A. photo-multiplier was used. Because its sensitivity changes with the wavelength of the radiation, a variation of the spectral distributions of the fluorescence with temperature causes a variation of the photo-currents, even when the efficiency of the fluorescence is constant. The curves of fig. 23 are not corrected for this variation because the measurements could not be carried out on the calibrated apparatus, owing to the high quenching temperatures. Therefore the

curves are not accurate, but the qualitative conclusions from fig. 23 given above will be correct. During the high-temperature measurements the phosphors were kept in an atmosphere of nitrogen and hydrogen (12%) to prevent oxidation.

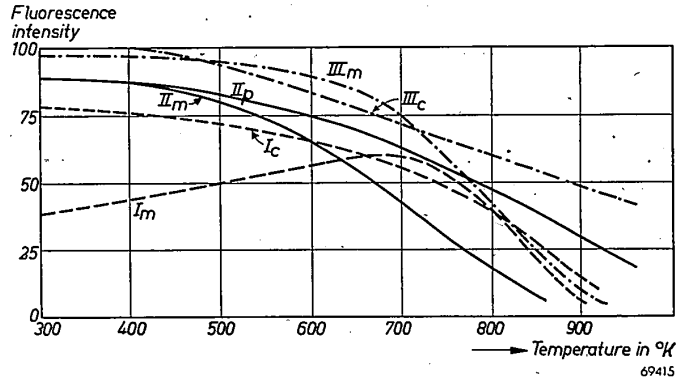



Fig. 23. Temperature dependence of the intensity of fluorescence. $I_c$: Ce, $I_m$: Mn, in $Ca_3(PO_4)_2$-Ce-Mn; $II_p$: Pb, $II_m$: Mn, in $CaSiO_3$-Pb-Mn; $III_c$: Ce, $III_m$: Mn, in $Sr_3(PO_4)_2$-Ce-Mn. Excitation by $\lambda$ 2537 Å.

## 5.2. *Calculation of the activation energy* $\varepsilon_t$ *in* $Ca_3(PO_4)_2$-Ce-Mn

In section 5.1 it is assumed that the temperature dependence of the efficiency of energy transfer between say 100 °K and 400° K in $Ca_3(PO_4)_2$-Ce-Mn is due to energy transport from an excited cerium via other cerium ions to manganese ions. This energy transport will need an activation energy $\varepsilon_t$. An excited cerium ion near a manganese ion transfers the energy to manganese without needing an activation energy. The $\varepsilon_t$ will now be calculated in a way similar to that followed in part I for other systems.

Excitation in a cerium ion not near a manganese ion gives rise either to cerium fluorescence with probality $K_c$ or to manganese fluorescence by energy transport with probability $K_t$.

Let $\eta_T$, $\eta_{100}$ and $\eta_{400}$ be the quantum efficiencies of the cerium fluorescences at temperature $T$, 100 °K and 400 °K. Then the relative decrease of the quantum efficiency of the cerium fluorescence $\eta_c$ with temperature can be defined as

$$\eta_c = (\eta_{100} - \eta_T)/(\eta_{100} - \eta_{400}) .$$

$\eta_c$ is determined by $K_c$ and $K_t$ according to

$$\eta_c = K_c/(K_c + K_t) . \qquad (15)$$

$K_c$ may assumed to be constant and $K_t = s_t \exp(-\varepsilon_t/kT)$, in which $\varepsilon_t$ is the activation energy and $s_t$ is a constant. Substitution in (15) gives after some calculation

$$-\varepsilon_t/kT + \ln(s_t/K_c) = \ln\{(1/\eta_c) - 1\}. \quad (16)$$

In the same way a formula can be deduced for the relative increase of the quantum efficiency of the manganese fluorescence $\eta_m$, giving:

$$\varepsilon_t/kT - \ln(s_t/K_c) = \ln\{(1/\eta_m) - 1\}. \quad (17)$$

When $\ln\{(1/\eta_c) - 1\}$ and $\ln\{(1/\eta_m) - 1\}$ are plotted against $1/T$ we ought to find straight lines, the slopes of which define $\varepsilon_t$ and the intercepts at $1/T = 0$ define $s_t/K_c$. The results of plotting $\ln\{(1/\eta_c) - 1\}$ and $\ln\{(1/\eta_m) - 1\}$ against $1/T$ for some phosphors are shown in fig. 24; the values of $\eta_c$ and $\eta_m$ were obtained from fig. 8. In fig. 24 it may be seen that fairly good straight lines are obtained. From this figure the values of $\varepsilon_t$ and $s_t/K_c$ are calculated and the results are given in table VIII.

TABLE VIII

| $Ca_3(PO_4)_2$-$x$ % Ce-$y$ % Mn | | applied formula | $\varepsilon_t$ in eV | $s_t/K_c$ |
|---|---|---|---|---|
| $x$ | $y$ | | | |
| 10/3 | 1/3 | (10) | 0·05 | 13 |
| 1/3 | 1/3 | (10) | 0·08 | 50 |
| 5/3 | 5/3 | (10) | 0·08 | 150 |
| 5/3 | 5/3 | ( 9) | 0·08 | 50 |
| 1/3 | 5/3 | ( 9) | 0·14 | 400 |

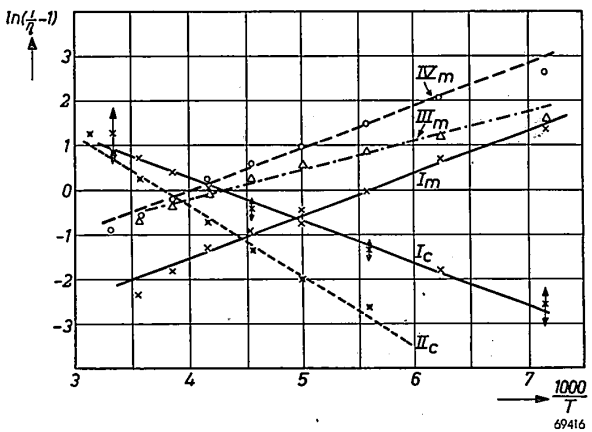



Fig. 24. Relation between $\ln[(1/\eta_c) - 1]$ and $1/T$ (lines marked with index $c$) and between $\ln[(1/\eta_m) - 1]$ and $1/T$ (lines marked with index $m$) for $Ca_3(PO_4)_2$-$x$ % Ce-$y$ % Mn.
I for $x = 5/3, y = 5/3$; II for $x = 1/3, y = 5/3$;
III for $x = 10/3, y = 1/3$; IV for $x = 1/3, y = 1/3$.

For $\varepsilon_t$ a plausible value is found. Substituting the value of $K_c \approx 10^7$ sec$^{-1}$ we obtain $s_t \approx 10^9$ sec$^{-1}$, a value not unlikely according to the theory of Mott and Seitz. This is in favour of the assumption of section 5.1.

5.3. *Summary of section 5*

Starting from the model of Mott and Seitz it is possible to give a schematic picture of the energy transfer in sensitized phosphors and of the temperature dependence of the efficiency of energy transfer. The observed temperature dependence in $Ca_3(PO_4)_2$-Ce-Mn could be explained assuming that the temperature-dependent part of the efficiency of energy transfer is due to energy transport from an excited cerium via other cerium ions to a manganese ion.

**Acknowledgement**

The author is greatly indebted to Prof. H. B. G. Casimir for fruitful discussions, and to Drs F. A. Kröger, H. A. Klasens, and Mr P. Zalm for the preparation of the phosphors as well as for many valuable discussions.

*Eindhoven, September 1951*

REFERENCES

41) R. Tomaschek, Ann. Phys. **75**, 561-597, 1924.
42) S. Rothschild, Physik. Z. **35**, 557-560, 1934; **37**, 757-763, 1936.
43) K. J. Murata and R. L. Smith, Am. Mineral. **31**, 527-538, 1946.
44) J. H. Schulman, E. Burnstein, L. W. Evans, R. J. Ginther and M. White, Bull. Am. phys. Soc. **24**, (No. 4), 25, 1949.
45) J. H. Schulman, R. J. Ginther and C. C. Klick, J. electrochem. Soc. **97**, 123-132,
46) R. J. Ginther, J. electrochem. Soc. **98**, 74-82, 1951. [1950.
47) J. H. Schulman, L. W. Evans, R. J. Ginther and K. J. Murata, J. app. Phys.
48) H. C. Froelich, J. electrochem. Soc. **93**, 101-113, 1948. [**18**, 732-739, 1947.
49) G. R. Fonda and H. C. Froelich, J. electrochem. Soc. **93**, 114-122, 1948.
50) F. J. Studer and G. R. Fonda, Bull. Am. phys. Soc. **23**, (No. 3), 48, 1948; J. opt. Soc. Am. **39**, 655-660, 1949.
51) J. B. Merrill and J. H. Schulman, J. opt. Soc. Am. **38**, 471-479, 1948.
52) Th. P. J. Botden and F. A. Kröger, Physica, 's-Grav., **14**, 553-566, 1948.
53) H. G. Jenkins, A. H. Mc. Keag and R. W. Ranby, J. electrochem. Soc. **96**, 1-12, 1949.
54) R. Nagy, R. W. Wollentin and C. K. Lui, J. electrochem. Soc. **95**, 187-193, 1949.
55) K. H. Butler and C. W. Jerome, J. electrochem. Soc. **97**, 265-270, 1950.
56) H. C. Froelich, J. electrochem. Soc. **95**, 254-266, 1949.
57) K. H. Butler, J. electrochem. Soc. **95**, 267-281, 1949.
58) J. Franck and E. Teller, J. chem. Phys. **6**, 861-872, 1938.
59) F. A. Kröger, Physica, 's-Grav., **15**, 801-806, 1949.
60) M. Hüniger and H. Pankc, U. S. Patents 2,241,950 and 2,241,951,1941.
61) H. C. Froelich, Electrochem. Soc. preprint **91**, 11, 1947.
62) F. A. Kröger and J. Bakker, Physica, 's-Grav., **8**, 628-645, 1941.
63) H. L. J. Bäckström: Arkiv Kemi Mineral. Geol. 13 A, No. 24, 1-15, 1940.
64) H. C. Froelich, U.S.A. Patent Applic. 498.731.
65) F. A. Kröger, J. Th. G. Overbeek, J. Goorissen and J. v. d. Boomgaard, J. electrochem. Soc. **96**, 132-141, 1949.
66) G. R. Fonda, J. phys. Chem. **44**, 435-439, 1939.
67) P. D. Johnson and F. E. Williams, J. chem. Phys. **18**, 1477-1483, 1950; Phys. Rev. **81**, 146, 1951.
68) M. M. Moodie and C. Reid, J. chem. Phys. **19**, 986, 1951.

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

Reprints of the majority of these papers can be obtained on application to the Administration of the Research Laboratory, Kastanjelaan, Eindhoven, Netherlands. Those papers of which no reprints are available are marked with an asterisk (*).

**1970:** J. Haantjes and W. F. de Vrijer: Flicker in television pictures (Wireless Engr **28**, 40-42, 1951).

The flicker of television pictures can be compared with the flicker of motion-picture projection. In both cases the perceptibility of flicker depends upon the frequency at which the pictures follow each other, the luminance of the picture, and some other factors. In the case of motion-picture projection an important factor is the ratio of the intervals during which the illumination is on and off, and taking the place of this in the case of television is the decay time(s) of the phosphor mixture used for the screen of the tube. This article first deals with the origin of flicker and the properties of the eye giving rise to this phenomenon. Some laws of relationship are mentioned. Finally some experiments with television pictures are discussed from which it appears that, with a suitable choice of the fluorescent substance and working with the European frame frequency of 50 c/s, a high-light luminance of 200 cd/m$^2$ is permissible without causing troublesome flicker. This value is ample for good viewing of the picture in surroundings which are not blacked out.

**1971:** N. W. H. Addink: Quantitative spectrochemical analysis by means of the direct current carbon arc (Rec. Trav. chem. Pays-Bas **70**, 155-167, 1952).

In view of the fact that in some spectrochemical laboratories all kinds of materials have to be analysed, a general method of complete evaporation by means of the direct-current carbon arc has been worked out. This method has been succesfully applied for over two years, during which time 2500 analyses have been carried out, 30% of which have yielded quantitative results. Difficulties encountered in putting the method to the test (sputtering, decreasing or increasing the speed of vaporization of materials to be analysed) are described in this paper. The method of successive addition has proved to be of importance for the determination of the exponent $n$ in the equation $i = k.c^n$ ($i$ = intensity of a spectrum line; $k$ and $n$ are constants and $c$ = the concentration of the element to

be analysed). Some values of $n$ are listed at the end of the paper. The method of successive addition enables one to determine element concentrations independently of the chemical analysis, by adding successive amounts of that element to the material under consideration and by measuring line intensities of mixtures prepared in this manner. The method of measuring line intensities by means of "standard paper density scales", which gives rise to the choice of lines corresponding to electron transitions in the higher parts of the electron-level diagram of an atom, has proved that the intensities of these lines show an $n$-value of about 1. S.p.d. scales contain a set of lines of well-known increasing density (and corresponding relative intensity); the image of a spectrum line is compared with them on the screen of a spectrum projector. By this method the influence of background on line-intensities is eliminated. The accuracy of results obtained is discussed.

**1972:** N. W. H. Addink: Quantitative spectrochemical analysis by means of the direct current carbon arc, II. Biological materials. A possible correlation between the zinc content of liver and blood and the cancer problem (Rec. Trav. chim. Pays-Bas **70**, 168-181, 1951).

For the contents of this paper, see **1947**.

**1973:** N. W. H. Addink: The degree of imperfection of crystals (Rec. Trav. chim. Pays-Bas **70**, 202-208, 1951).

An attempt is made to arrange crystals of various materials according to their degree of imperfection, derived from the apparent value of $N$ (the Avogadro number). $N$ has been determined for each kind of crystal according to the formula $N = cM/Vd$, where $M$ is the molecular weight, $c$ a factor related to the number of molecules per unit cell, $V$ the volume of the unit cell, and $d$ the density of the crystal. Metals, in accordance with their mosaic structure, show a relatively high value of $N$ and therefore a high degree of imperfection (0·06 % of their volume), whilst the lowest value of $N$ has been found for diamond, quartz, calcite and potassium chloride (prepared from a solution), namely $N\mathrm{Cm} = (6{\cdot}0228 \pm 0{\cdot}0014)\ 10^{23}$, as compared with Cohen and Dumond's value $(6{\cdot}02378 \pm 0{\cdot}00011)\ 10^{23}$. Special attention is drawn to the behaviour of perfect crystals of potassium chloride (from a solution) and imperfect crystals of the same substance (from the melt): adsorption phenomena of water vapour enable the thickness of mosaic blocks to be estimated. A value of $6 \times 10^{-4}$ cm has been found, assuming a monomolecular layer of water molecules on the internal surface.

**1974:** R. van der Veen: Influence of day-length on the dormancy of some species of the genus Populus (Physiologica Plantarium **4**, 35-40, 1951).

Experiments show that dormancy of poplar trees (Populus) is induced solely by day length. In long days (14 hours and longer) growth and development of new leaves is maintained; in short days (12 hours and shorter) the trees go into dormancy irrespective of the temperature. At 20°C and 29°C all eight species of Populus investigated had stopped their growth in less than 45 short days. After 102 short days all these trees had shed their leaves and had gone into complete dormancy.

**1975*:** P. J. Bouma: Farbe und Farbwahrnehmung. Einführung in das Studium der Farbreize und Farbempfindungen. (N.V. Philips' Gloeilampenfabrieken, Dept. tech. and scientific literature, 1951, XVI + 358 pp., 113 figs).

German translation of "Physical aspects of colour" by the same author; see **1768***, **1842***.

**1976*:** J. M. Stevels: Expériences et théories sur le facteur de puissance des verres en fonction de leur composition, II (Verres et Réfractaires **5**, 4-14, 1951).

French translation of **R 158.**

**1977:** A. J. W. M. van Overbeek: Voltage-controlled secondary-emission multipliers (Wireless Engnr **28**, 114-125, 1951).

Difficulties with life in secondary-emission amplifier valves have been overcome by employing a layer of caesium oxide kept at a temperature below 180°C. Several constructions of experimental valves are shown and a description is given of their characteristic properties. A variable-mu valve and a very-high-slope valve employing four stages of multiplication are shown. The use of grid-shaped dynodes (secondary-emission electrodes) is discussed. Some circuits in which secondary-emission valves offer typical advantages are described, in particular generators of sinusoidal and non-sinusoidal oscillations and monostable and bistable trigger circuits.

**1978*:** W. Elenbaas; The high pressure mercury vapour discharge. (Noord Holl. Uitg. Mij, Amsterdam, 1951, XII + 173 pp., 80 figs).

This book contains a survey of about 30 papers by the author and some hundred papers by others writers on the high-pressure mercury discharge.

After a short historical introduction and a description of the mercury lamps now in use, the suppositions on which the theory of this discharge is founded are exposed, viz. that in every point of the discharge there is thermodynamic equilibrium as regards excitation and ionization and that the power developed is dissipated by radiation and conduction. Different methods of measuring the temperature of the discharge are discussed. From the differential equation governing the radial temperature distribution similarity laws are derived, and it is shown how the equation may be solved graphically. Then there follows a discussion about the influence of a certain number of phenomena, such as convection, external magnetic fields, addition of rare gasses and of metals with an ionization potential lower than that of mercury. The above considerations mostly pertain to the discharge in mercury vapour of about 1 atm. A special chapter is devoted to small mercury discharges with higher pressure and high luminance, which may be artificially cooled. Elaborate discussions are devoted to the spectrum and the gradual transition from a pure line spectrum to a spectrum with broadened lines and continuous background at higher vapour densities. A number of miscellaneous questions are discussed, such as the relation between the pressure and the potential gradient, the role of positive ions, the conditions near the electrodes, the influence of self-absorption on the intensity of spectral lines, the question of the dimensioning of mercury lamps. Finally it is tried to give account of the quasi-thermodynamic equilibrium as found experimentally, by means of kinetic considerations using known data about probabilities of excitation and ionization. The result is rather convincing, taking into account the uncertainty regarding the functions governing the elementary processes.

**1979:** G. W. Rathenau and J. L. Meijering: Rapid oxidation of metals and alloys in the presence of $MoO_3$ (Metallurgia **42**, 167-172, 1950).

Alloys which are protected against corrosion by a dense skin of oxide, and thus may serve as heat-resistant constructional materials, sometimes show a strongly accelerated oxidation when brought into contact with foreign substances, especially foreign oxides. With a view to accounting for this phenomenon, the oxidation of a series of metals and alloys, i.a. with copper, iron and silver as base, has been investigated for the case where these are brought into contact with molybdenum oxide ($MoO_3$ in powder form). The degree of oxidation ("penetration depth") was determined by measuring the breaking strength of wires drawn from the respective material and heated in air with and without contact with $MoO_3$. When the oxidation is thus determined as a function of temperature a curve is obtained showing one or more discontinuities where the oxidation is suddenly accelerated. These discontinuities appear to lie at the

eutectic temperatures of the systems of the respective oxides (lowest temperatures at which the formation of liquid occurs). From this it can be concluded that the accelerated oxidation is due to formation of liquid in the protective oxide skin of the alloy. In some cases oxides of the basic metal not essentially belonging to the protecting skin, such as $Cu_2O$ in the case of aluminium bronze, may considerably contribute towards a lowering of the eutectic temperature. Sometimes in the presence of $MoO_3$ perceptible oxidation already takes place below the lowest eutectic temperature. This is accounted for by the absorption of molybdenum ions in the crystal lattice of the oxide skin. A third way in which $MoO_3$ may promote oxidation is found in the case of silver. Silver, as a precious metal, is not oxidizable above 150 °C in air of normal pressure, but in the presence of $MoO_3$ double oxides may be formed, making it possible for oxidation to take place at much higher temperatures; then again melting-phenomena in the oxide skin make their influence felt. Metallographic examination shows peculiarities in the attack of some alloys.

**1980:** H. Bremmer: On the theory of optical images affected by artificial influences in the focal plane (Physica **17**, 63-70, 1951).

Optical images can be influenced by an articifial modification of the diffraction field in the focal plane in the image space. The effects of this modification are discussed in this paper, starting from an arbitrary distribution of the corresponding attenuation and phase retardation imposed at each point of the focal plane. There distributions may be chosen such as to accentuate very special features of the object in its paraxial image. The variety of possibilities in this respect is illustrated by a number of examples. For instance, it appears to be possible to reproduce theoretically the modulus of the gradient of the transparency for absorbing objects, and the modulus of the gradient of the phase-rotation distribution for transparant objects.

**1981:** P. C. van der Willigen: Grepen uit de ontwikkeling van het booglassen (Electrotechniek **29**, 143-147, 1951). (Some remarks on the development of arc welding; in Dutch).

Two important questions in the field of electric arc-welding by hand are discussed in this article. The first part deals with the choice between direct current and alternating current. Their advantages and disadvantages are described, including the question of the open-circuit voltage of the welding transformer. In the second part the different methods for welding of steel are examined from the point of view of shielding of the deposited metal against nitrogen and oxygen from the air. In this connection the use of hydrogen and its influence on the weld are also briefly discussed.

VOL. 7 NR 4 AUGUST 1952

# Philips Research Reports

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

*R 193* *Philips Res. Rep.* 7, 241-250, 1952

## A SINGLE COMPONENT WHITE LUMINESCENT SCREEN FOR TELEVISION TUBES

by F. A. KRÖGER, A. BRIL and J. A. M. DIKHOFF

621.386.832:535.373.1

**Summary**

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-viewing television tubes. Efficiency and current saturation are similar to those of the sulphide mixtures normally used.

**Résumé**

On a développé un nouveau luminophore (Zn,Cd)S-Ag-Au-Al qui présente une forte luminescence blanche lorsqu'il est excité par les rayons ultra-violets, les rayons cathodiques ou les rayons X, et qui peut être utilisé pour les tubes de télévision à vision directe. Le rendement et le saturation sont analogues à ceux des mélanges de sulfures utilisés ordinairement.

**Zusammenfassung**

Es wurde ein neuer Lumineszenzstoff (Zn,Cd)S-Ag-Au-Al entwickelt, der bei Erregung durch ultraviolette Strahlen, Kathodenstrahlen oder Röntgenstrahlen eine starke weiße Lumineszenz zeigt und sich zur Verwendung in Fernsehempfangsröhren mit unmittelbarer Bildwiedergabe eignet. Wirkungsgrad und Stromsättiging entsprechen ungefähr demselben Ausmaß wie bei den normal gebräuchlichen Sulfidgemischen.

### 1. Introduction

In television receivers the picture is made visible on the luminescent screen of a cathode-ray tube. In one system the viewers see the picture directly on the tube face (direct viewing receiver); in another system the light from the tube is projected onto a larger screen (projection receiver). The latter system has even been adapted for use in cinemas.

In all cases the colour of the light is more or less white. For the United States of America the colour of the light a television tube should emit has been provisionally standardized by the Joint Electron Tube Engineering Council of the Institute of Radio Engineers [1]): it must be such that the corresponding colour point in the I.C.I. colour triangle must lie within a

polygon around the point of equi-energy white W (fig. 1). In Europe there is a tendency to prefer a slightly more bluish hue, but a definite choice has not yet been made.

With the tubes now in use the white luminescence is obtained with the aid of screens consisting of two luminescent components, one of which has a blue and the other a yellow fluorescence. Although the number of luminophors showing cathodo-luminescence is extremely great, only very few have properties making them suitable for practical use. For blue fluores-

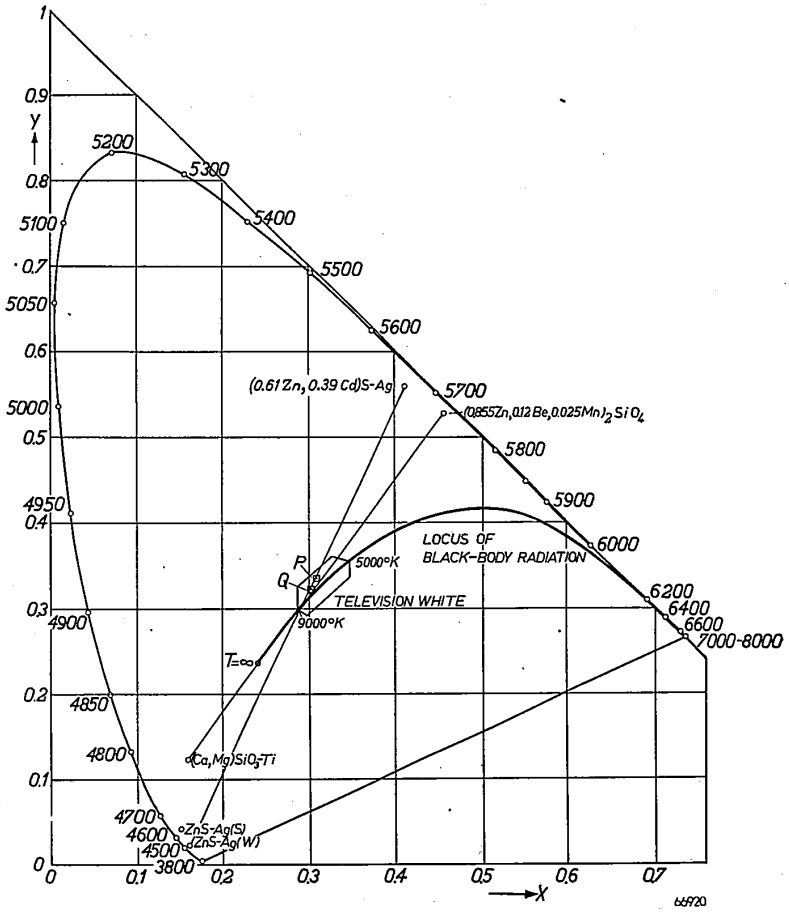



Fig. 1. Colour points of the luminescence of various luminophors used as components in white luminescent mixtures. *P* and *Q* denote such mixtures, the spectral distributions (and compositions) of which are shown in fig. 2. The symbols *S* and *W* denote the crystalline modifications of ZnS: sphalerite or wurtzite.

cence zinc sulphide activated by silver, and calcium-magnesium silicate activated by titanium have proved to be the best; for the yellow, zinc-cadmium sulphide-selenide activated by silver, and zinc-beryllium silicate activated by manganese offer the best solution. The sulphides and sulphide-selenides have the highest efficiency, amounting to 13-25% at 10-20 kV. Owing to the fact that they contain only a relatively small amount of activator ($\sim 10^{-4}$) they are easily saturated and therefore cannot be applied for tubes in which the cathode rays reach a high current density. The silicates on the other hand have a lower efficiency ($\approx 7\%$ at 20 kV) but they have a higher concentration of centres and therefore remain linear up to much higher current densities [2]). For these reasons most direct viewing tubes contain a sulphide screen, while the silicates are used mainly in projection tubes. The spectral distribution of a sulphide mixture ($P$) and a silicate mixture ($Q$) is shown in fig. 2 *). The corresponding colour points are shown in fig. 1, together with those of the fluorescence for the separate components. It is seen that the colour points of the mixtures lie fairly well on the line connecting the points of the components; this indicates that the influence of the absorption of light generated in one component by the crystals of the other component is negligible.

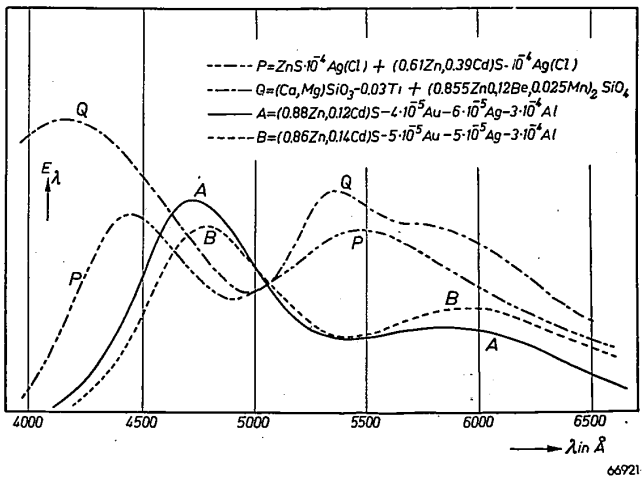



Fig. 2. Spectral distribution of the luminescence from various luminescent systems, having colour points within the tolerated white area (see figs 1 and 4). $P$ corresponds to a mixture of blue luminescent ZnS-Ag(Cl) and yellow luminescent (Zn,Cd)S-Ag(Cl); $Q$ corresponds to a mixture of blue luminescent $(Ca,Mg)SiO_3$-Ti and yellow luminescent $(Zn, Be, Mn)_2SiO_4$, while $A$ and $B$ correspond to a double activated sulphide.

*) Throughout this paper concentrations are given either in molar fractions, or in gram atoms per mole.

In the development of new screen materials various efforts have been made to obtain white luminescence from one luminophor rather than from a mixture [3]) [4]). Most of these attempts have failed, however, either because the products obtained were not sufficiently efficient, or because the colour of the fluorescence did not fall within the tolerated area. It is only recently that more satisfactory results seem to have been obtained [5]).

As we have seen above, white can be obtained by the superposition of fluorescence bands originating from two luminophors. Therefore the simplest way of obtaining a white radiation from one component will be to develop a luminescent system showing fluorescence in two sub-bands. These sub-bands must fulfil the same conditions required of the bands of a mixture: the line connecting the colour points of the two bands must pass through the white area; the proportion in which the bands are emitted must be such that the colour point of the two bands together falls within this same area, and finally the efficiency with which the luminescence is generated must be high.

In principle a two-band fluorescence can be obtained by the incorporation of two activators in one base material. Since, as we have seen, the zinc-cadmium sulphides have an exceptionnally high efficiency, while on the other hand various activators are known for these systems, these substances have been chosen for a renewed attack upon the problem of preparing a white luminescent luminophor.

## 2. (Zn,Cd)S-Ag,Cu

Zinc sulphide activated with silver shows fluorescence in a blue band, while copper as activator causes a green and a blue band; under appropriate conditions of preparation the green copper band predominates [6]),[7]). The various bands have a slightly different position for the two modifications of ZnS; for the wurtzite modification the maxima are Ag: 4350 Å, Cu: 5160 Å; for the sphalerite modification the maxima lie 150 Å towards longer wavelengths [7]). A similar difference exists between samples made under sulphurizing conditions and those made in the presence of oxygen [6]),[8]). Colour points of the blue silver band and the green copper band for oxygen-free ZnS in the sphalerite (*S*) and wurtzite modification (*W*) are given in fig. 3. Colour points representing the fluorescence attainable with zinc sulphide containing copper as well as silver in varying proportions are situated on straight lines connecting the points of the silver and copper bands. These lines pass far from the white area, indicating that a tolerably white luminescence cannot be obtained in this way. Introduction of cadmium sulphide into the system, causes the bands to shift towards longer wavelengths [9]); fig. 3 also shows the colour points of the silver and copper band in solid solutions (Zn, Cd) S of varying cadmium content. The points shift regularly

through the colour triangle, forming loci of Ag and Cu bands. Lines connecting corresponding points of these loci represent the colours attainable with double-activated (Zn, Cd) S luminophors. It is seen that although the lines of the solid solutions containing 10-16 mole % CdS approach the white area rather closely, this region is not crossed. For preparations made in the presence of oxygen all colour points shift along the loci over a distance corresponding to an increase of the CdS content by a few percent, but this

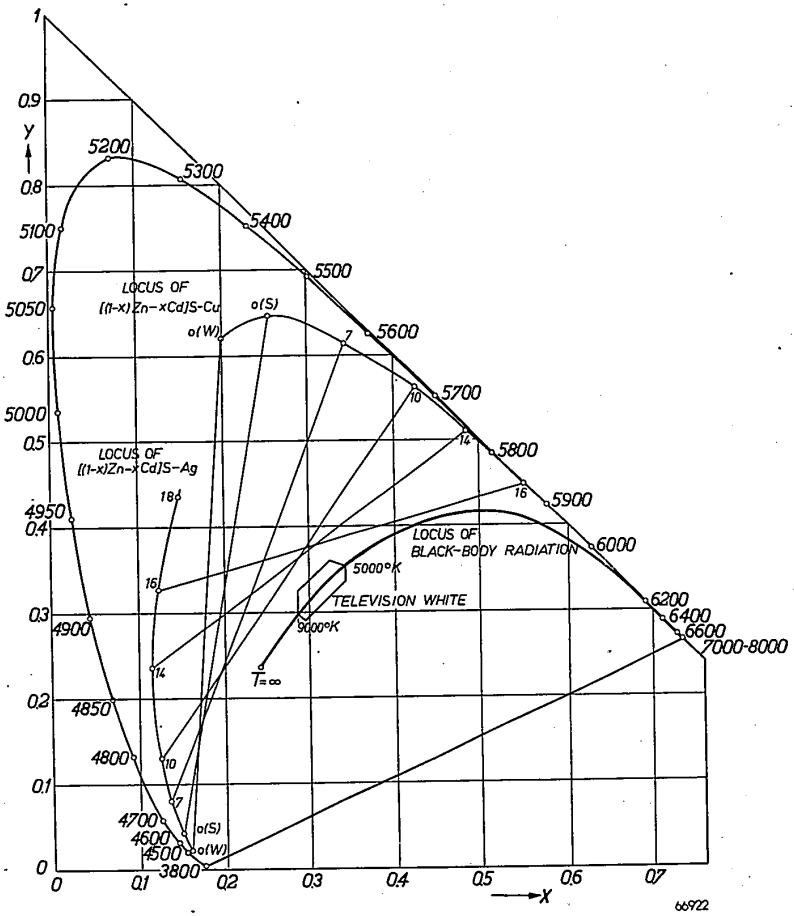



Fig. 3. Loci of silver and copper bands for (Zn,Cd)S having the wurtzite structure. The figures along the loci indicate the cadmium content ($x$). Colours attainable in double-activated systems are indicated by lines joining corresponding points of the two loci. Points for the sphalerite modification ($S$) of ZnS are also given.

obviously does not diminish the shortest distance between the white area and the colour line most closely approaching it. Therefore a white luminescence cannot be obtained with a luminophor of this composition.

### 3. (Zn,Cd)S-Ag,Au

From fig. 3 it is seen that a white luminescence could be obtained with the copper band in combination with a blue band with a maximum lying at a wavelength approximately 150 Å shorter than that of the silver band, or with the silver band in combination with a long-wave band lying over a similar distance at the long-wave side of the copper band. The only activator known to give a short-wave fluorescence in ZnS is Na(Cl) [10]), but the band of this activator is only emitted at low temperatures and therefore cannot be used. On the other hand several activators are known to give long-wave bands, viz Mn [11]), Pb [12]), Sn [13]), As [17]), P [5]), and Au [14]). The intensity of the fluorescence of the first four activators is not sufficiently great to make these bands of practical value. Phosphorus, however, gives rise to a yellow band of high intensity, and a satisfactory white luminescent system containing phosphorus has actually been made [5]). Yet the difficulties involved in preparing such luminophors have prevented their application in industry up till now.

Gold, finally, gives rise to a strong luminescence and the spectral distribution of the main band fulfils the requirements, as is shown in fig. 4; the lines connecting corresponding points of the Ag and Au loci pass through the white area for cadmium sulphide contents between 6 and 15 mole %.

The preparation of white luminescent (Zn, Cd) S-Ag, Au is complicated by the fact that gold, just like copper, gives rise to two emission bands. For ZnS (wurtzite) these bands have maxima at 5300 and 4900 Å. The long-wave band is the one required to obtain a white luminescent luminophor; the locus of the Au band in fig. 4 applies to this band. The short-wave band, on the other hand, must be suppressed; if this band is also present it shifts the colour points towards the green, and the white area cannot be reached.

Under normal conditions of preparation, using halides as fluxes, it is possible to prepare ZnS-Au showing fluorescence predominantly in the long-wave band. With products containing silver as well as gold, however, the appearance of the short-wave band cannot be prevented: it is even so that the silver band and the long-wave gold band disappear together, leaving only the short-wave gold band *).

This complication can be avoided by preparing the phosphors with the addition of trivalent ions instead of halides. The best results are obtained

*) This effect suggests an interaction between silver and gold centres which needs further investigation.

with aluminium [16]). Now the short-wave gold band does not appear and all colours determined by the loci in fig. 4 can be obtained. The spectral distribution of the fluorescence of two samples of slightly different composition (A and *B*) is given in fig. 2, and the corresponding colour points are given in fig. 4.

Products prepared in this way show a moderate phosphorescence in the

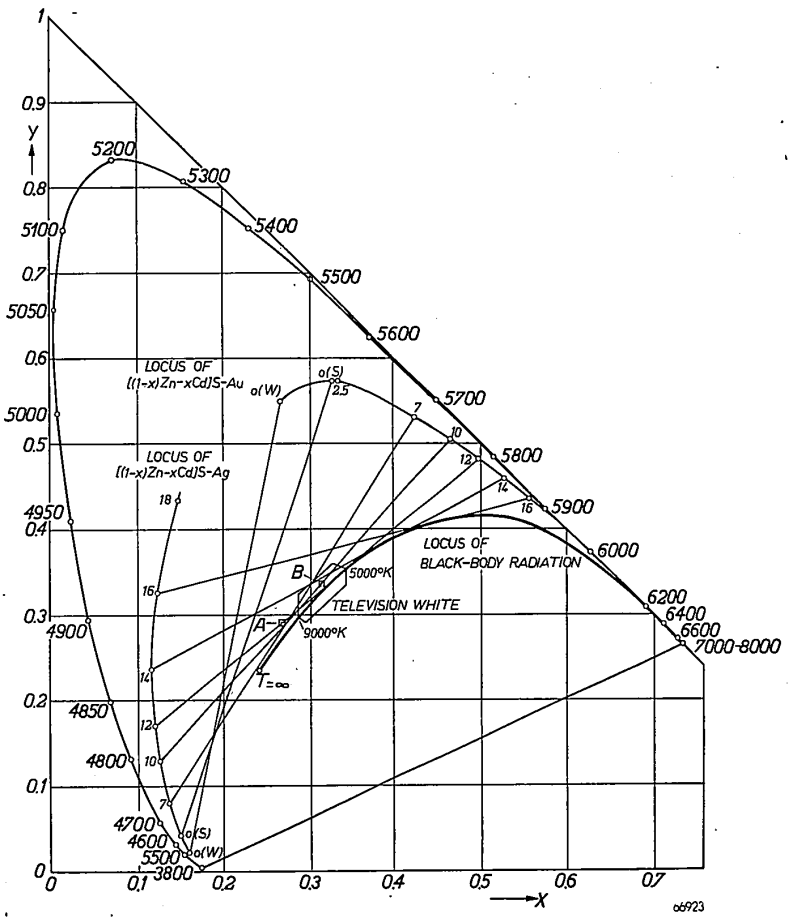



Fig. 4. Loci of silver and gold bands in (Zn, Cd)S,having the wurtzite structure. Points for the sphalerite modification (*S*) of ZnS are also given. The figures along the loci indicate the cadmium content ($x$). Colours attainable in double-activated systems are indicated by lines joining corresponding points of the two loci. *A* and *B* are the colour points of the two double-activated products; the spectral distributions (and the compositions) are given in fig. 2.

IPR2025-00217 EXHIBIT 2010 Page 251 of 485

long-wave band. This afterglow can be suppressed in the usual way by adding a killer. Satisfactory results are obtained with incorporation of $10^{-6}$ Ni.

## 4. Intensity and saturation properties

The properties of these double-activated luminophors are similar to those of the sulphide mixtures normally used. For excitation by cathode rays of 20 kV and low current densities under normal scanning conditions the energy efficiency is 15%, corresponding to 43 lumens per Watt. For 10 kV these values are 14 % and 40 lumens per Watt. These efficiencies were measured in a demountable tube with a relatively thick phosphor layer, so that practically all the fluorescent light was emitted to the same side (viz the gun side) of the layer. To prevent charging up of the phosphor, it was applied on a metal plate connected to the anode of the tube.

At higher currents the luminence tends towards saturation and the efficiency diminishes. Fig. 5 shows the saturation for a double-activated

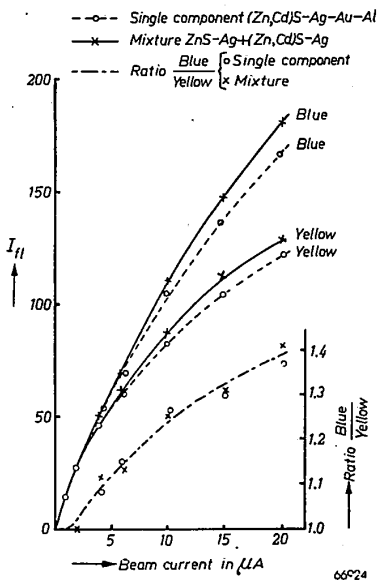



Fig. 5. Intensity of the cathodo-fluorescence in the blue and yellow sub-bands as a function of the beam current for a mixture of ZnS-Ag(Cl) and (Zn,Cd)S-Ag(Cl), and for a double-activated luminophor of the composition (0·88 Zn, 0·12 Cd)-$2{\cdot}5.10^{-5}$ Ag-$2{\cdot}5.10^{-5}$ Au-$10^{-4}$Al. The units are arbitrary and have been chosen in such a way that the curves start with an equal slope in the origin. The lower curve shows the ratio between the blue and yellow sub-bands.

luminophor together with that for a normal mixture of blue luminescent ZnS-Ag and yellow luminescent (Zn, Cd)S-Ag, measured separately for the blue and yellow bands. To show the relative amount of saturation more clearly, the slopes at low current densities are made equal for all the curves.

These curves are also the result of measurements in a demountable tube in which the two luminophors could be brought under the bombarding electron beam in succession without opening the tube. Thus all conditions and especially the focus diameter, which is very critical in such measurements, were exactly the same. A current of 20 μA in these experiments corresponds to peak intensities in direct viewing tubes (400-500 μA for a raster of 20 cm line amplitude and 10 cm image amplitude). The lower curve in fig. 5 gives the ratio between the blue and the yellow sub-bands of these luminophors; it shows that the yellow saturates more than the blue in both cases, this resulting in a change of colour with increasing current densities. In a direct-viewing tube this variation corresponds to a difference in colour point between the dark parts and the high lights of $\Delta x = 0{\cdot}03$ and $\Delta y = 0{\cdot}025$, i.e. a shift over one half to two thirds of the length of the white polygon, which is not unacceptable.

As is seen from the curve the colour change does not increase indefinitely but tends towards saturation at higher currents. This is due to two facts. In the first place the intensity of the fluorescence in each band has an asymptotic value, determined by the concentration of the centres present; therefore also the ratio between the two bands must approach a definite value. In the second place the focus area becomes larger with increasing beam current, so that the current density does not increase linearly with the current measured.

Experiments with double activated luminophors containing the activator in a fixed proportion but with a varying sum-concentration showed that the saturatoin of each sub-band decreases with increasing concentration of centres, but the difference in saturation increases with the concentration. Therefore in practice a medium concentration has to be applied for which neither the saturation nor the difference in saturation is too great. The data of fig. 5 refer to a product of such a composition.

*Eindhoven, January 1952*

## REFERENCES


[1]) R. M. Bowie, A. E. Martin, Proc. Inst. Radio Engrs **36**, 1023-1029 (1948).
[2]) A. Bril, Physica, 's-Grav., **15**, 361-379 (1949).
[3]) H. W. Leverenz and F. Seitz, J. appl. Phys. **10**, 479-493 (1939);
H. W. Leverenz, Proc. Inst. Radio Engrs **32**, 256-263 (1944).
[4]) S. Rothschild, Brit. Pat. Spec. 560,484.
[5]) A. H. McKeag and P. W. Ranby, J. electrochem. Soc. **96**, 85-90 (1949).
[6]) S. Rothschild, Trans. Faraday Soc. **42**, 635-642 (1946).
[7]) F. A. Kröger and J. E. Hellingman, J. electrochem. Soc. **93**, 156-171 (1948);
F. A. Kröger, J. E. Hellingman and N. W. Smit, Physica, 's-Grav., **15**, 990-1018 (1949).
[8]) A. L. Smith, J. electrochem. Soc. **93**, 324-333 (1949).
[9]) A. A. Guntz, Ann. chim. **5**, 157 (1926);
S. Rothschild, Z. Phys. **108**, 24-39 (1938).
[10]) F. A. Kröger, J. opt. Soc. Am. **39**, 670-672 (1949).
[11]) F. A. Kröger, Thesis, Amsterdam, 1940.
[12]) P. Lenard, Handb. Exp. Physik **23**, 397 (1928);
G. R. Fonda, J. opt. Soc. Am. **36**, 382-389 (1946);
N. W. Smit and F. A. Kröger, J. opt. Soc. Am. **39**, 661-663 (1949).
[13]) A. A. Cherepnev, Doklady Akademii Nauk S.S.S.R. **62**, 767-769 (1948);
G. F. J. Garlick and D. E. Mason, Proc. phys. Soc. **62A**, 817-822 (1949).
[14]) S. T. Henderson, Proc. roy. Soc. **A 173**, 323-338 (1939).
[15]) C. J. Milner, Trans. Faraday Soc. **35**, 101-111 (1939);
H. W. Leverenz, U. S. A. Pat. Spec. 2,402,757.
[16]) F. A. Kröger and J. A. M. Dikhoff, Physica, 's-Grav., **16**, 297-316 (1950).
[17]) F. A. Kröger and N. W. Smit, Belgian Patent 508,694;
J. S. Prener, J. electrochem. Soc. **98**, 406 (1951).

# HIGH-FREQUENCY DIODE ADMITTANCE WITH RETARDING DIRECT-CURRENT FIELD

by K. S. KNOL and G. DIEMER 621.385.2.011.21

**Summary**

A linear-field theory is given of the susceptance of a plan-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.

**Résumé**

Une théorie du champ linéaire est proposée pour la susceptance d'une diode à plans parallèles avec tensions d'anode négatives. Les contributions des électrons qui reviennent en arrière (susceptance d'émission totale) et des électrons qui traversent (susceptance exponentielle) sont considérées séparément. La théorie s'accorde qualitativement avec des résultats expérimentaux connus.

**Zusammenfassung**

Es wird eine Theorie der linearen Feldverteilung über die Suszeptanz einer planparallelen Diode mit negativen Anodenspannungen gegeben. Der Artikel behandelt die Beiträge, welche durch umkehrende Elektronen zustande kommen (Gesamtemissions-Suszeptanz), und sodann die Beiträge, die durchgehenden Elektronen zuzuschreiben sind (exponentielle Suszeptanz). Die Theorie stimmt qualitativ mit bekannten experimentellen Ergebnissen überein.

## 1. Introduction

The electronic conductance in diodes with negative anode voltages has been dealt with by several authors, some of them [1–5]) giving experimental results, whereas others [6–10]) calculated the conductance from transit-time considerations. Also a change in capacitance has been shown to be present [2]),[4]),[5]), but a theory of the electronic susceptance is lacking.

For a plan-parallel diode we will here derive formulae:
(a) for the total-emission susceptance, i.e., the electronic susceptance for such strongly negative anode voltages that the influence of the very few electrons that reach the anode can be ignored,
(b) for the electronic susceptance in the exponential region of the characteristic, due to crossing electrons ("exponential susceptance"). In both cases the retarding field is assumed to be linear.

## 2. Total-emission susceptance

We shall follow the same method of calculation as is usually applied for the total-emission (t.e.) conductance, i.e., the conductance due to electrons

that return to the cathode. The first-order calculation [6],[7]) assumes that the electrons are emitted from the cathode with a uniform velocity and enter a retarding linear d.c. field on which a small a.c. ripple is superimposed. These assumptions are too simple to give a satisfactory explanation of the experimental results, especially with large transit angles of the electrons. Begovich [8]) and Van der Ziel [9]) have dropped the assumption of uniform initial velocity and have carried out an integration over the thermal velocity distribution of the electrons emerging from the cathode. With small transit angles this refinement of the theory results only in an increase by a factor of about 1·5, but with large transit angles they found a dependence of the conductance on the transit angle that agreed much better with general physical considerations, especially concerning the fact that one must the t.e. conductance expect to become nowhere negative. However, the theory does not give until now a quantitative agreement with the experimental material.

For the calculation of the susceptance we assume, as Begovich and Van der Ziel did, that the static electric field strength $F$ in the diode is given by

$$F = -V_a/d\,,$$

where $-V_a$ is the anode voltage with respect to the cathode, and $d$ the cathode-anode distance. It can be shown that with normal cathodes and applied field strengths of at least some tenths of a volt per $\mu$ this approximation is sufficiently accurate [11]). (See, however, section 5.) We use nearly the same notation as Van der Ziel. The total-emission admittance $Y_T$ is given by

$$Y_T = \int_0^{V_a} \frac{4V_0}{V_a^2} \left[\tfrac{1}{8}\, j\omega\tau_{cc}\varphi_4(j\omega\tau_{cc})\right] I_s \exp\left[-V_0/V_T\right] \mathrm{d}V_0/V_T, \qquad (1)$$

where $V_0$ is the initial electron velocity in electron volts, $\tau_{cc}$ the transit time of an electron with initial velocity $V_0$ (in electron volts) from the cathode back to the cathode, $I_s$ the saturation current and $V_T = kT_c/e$ ($T_c$ is the cathode temperature in degrees Kelvin, $k$ Boltzmann's constant, $e$ electronic charge), $\varphi_4$ is a transit-time function given by Bakker and De Vries [6]):

$$\varphi_4 = \frac{6}{a^3}\left[(a-2) + (a+2)e^{-a}\right] =$$

$$= \frac{6}{(\omega\tau)^3}\left\{-(\omega\tau\cos\omega\tau - 2\sin\omega\tau + \omega\tau) + j(2\cos\omega\tau + \omega\tau\sin\omega\tau - 2)\right\}, \qquad (2)$$

with $a = j\omega\tau$.

Now we write

$$Y_T = G_T + jB_T, \tag{3}$$

$G_T$ being the t.e. conductance and $B_T$ the t.e. susceptance.

Eq. (1) is also given by Van der Ziel, who only makes use of the real part $G_T$. In substituting eqs (2) and (3) in (1) for the imaginary part $B_T$ we find:

$$B_T = \int_0^{V_a} \frac{4V_0}{V_a^2} \frac{2}{(\omega\tau_{cc})^2} (-\omega\tau_{cc}\cos\omega\tau_{cc} + 2\sin\omega\tau_{cc} - \omega\tau_{cc}) \times \\ \times I_s \exp[-V_0/V_T]\mathrm{d}V_0/V_T. \tag{4}$$

The following constants are introduced: $\tau_0 = 2d(2e/m)^{-1/2}V_a^{-1/2}$, the transit time of the slowest electron that in the static case can reach the anode; and $x_0 = \omega\tau_0(V_T/V_a)^{1/2}$.

With the conditions mentioned above we may substitute for the upper limit of the integral $\infty$ instead of $V_a$. Thus we obtain after some partial integrations

$$B_T = \frac{4I_sV_T}{V_a^2} \cdot \frac{\sqrt{\pi}}{2x_0} \{(1+2x_0^2)e^{-x_0^2} - 1\}.$$

Denoting the part that depends upon transit time by $b_T(x_0)$ we have

$$B_T = \frac{4I_sV_T}{V_a^2} b_T(x_0), \tag{5}$$

where

$$b_T(x_0) = \frac{\sqrt{\pi}}{2x_0} \{(1+2x_0^2)e^{-x_0^2} - 1\}, \tag{6}$$

which is plotted in fig. 1.

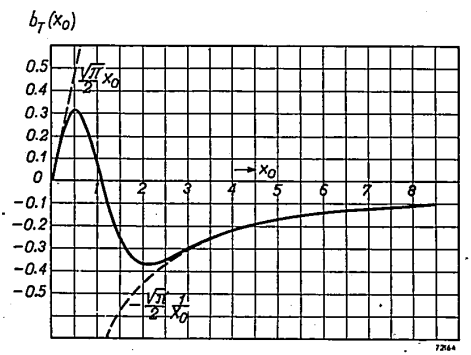



Fig. 1. Transit-time function of the total-emission susceptance; $x_0$ is the characteristic transit angle belonging to thermal initial velocity.

For small values of $x_0$ we have

$$b_T(x_0) \approx \frac{\sqrt{\pi}}{2} x_0; \tag{6a}$$

for large values of $x_0$

$$b_T(x_0) \approx -\frac{\sqrt{\pi}}{2} \frac{1}{x_0}. \tag{6b}$$

We see that, contrary to the t.e. conductance, the t.e. susceptance can be expressed in terms of elementary functions.

It is well known that under very general conditions there exist certain relations between the real and the imaginary part of a frequency dependent complex impedance [12]), [13]), [14]). Dr Bouwkamp *) of this laboratory showed that in this way the total-emission conductance $G_T(x_0)$ can be derived from $B_T(x_0)$ according to eq. (5). In a similar way in section 3 we shall derive the exponential susceptance from the exponential conductance.

### 3. Exponential susceptance

For the conductance due to those electrons that with negative anode voltages are still able to reach the anode Diemer [10]) has derived the following formula (we use a somewhat different notation to get in accordance with that of section 2):

$$G_e = \frac{I_a}{V_T} \psi_e(x_e), \tag{7}$$

with

$$\psi_e(x_e) = \frac{8}{x_e^4} (1 - \cos x_e + \sin^2 x_e - x_e \sin x_e - x_e \sin 2x_e + \tfrac{1}{2} x_e^2 + x_e^2 \cos^2 x_e), \tag{8}$$

where $I_a$ = static anode current; $x_e = \omega \tau_0 = (V_a/V_T)^{1/2} x_0$ = transit angle of the slowest electron that in the static case can reach the anode. We now regard $G_e$ as the real part of an as yet unknown admittance function

$$Y_e = Y_e(jx_e) = G_e(jx_e) + jB_e(jx_e), \tag{9}$$

and we want to calculate the electronic susceptance $B_e(jx_e)$. We remark that the function $G_e(z)$ according to eqs (7) and (8) is bounded on the imaginary axis ($\pm j\infty$ included). From physical considerations we know that the same must be the case for the susceptance $B_e(jx_e)$; so the function $Y_e$ is bounded on the imaginary axis. Furthermore we know that in the closed right-half plane, infinity included, $G_e(z)$ is bounded. Under these conditions we may write

*) Unpublished. Dr Bouwkamp kindly informed us concerning these methods giving a survey of four different types of calculation especially suitable for electronic admittances.

$$Y_e(z) = \int_{-\infty}^{+\infty} \frac{G_e(x)}{z^2 + x^2}\,dx =$$

$$= \frac{I_a}{\pi V_T} \int_{-\infty}^{\infty} \frac{8}{(z^2+x^2)x^4} (1 - \cos x + \sin^2 x - x\sin x - x\sin 2x + \tfrac{1}{2}x^2 + x^2\cos^2 x)\,dx \quad (10)$$

$$(\text{with } \mathrm{Im}(z) > 0).$$

The integral in eq. (10) can be evaluated with Cauchy's theorem, resulting in

$$Y_e(jx) = \frac{16 I_a}{V_T}\left[\frac{1}{4x_e^4}\left\{2 + x_e^2 + x\sin^2 x_e - 2\cos x_e - 2x_e\sin x_e - 2x_e\sin^2 x_e + 2x_e^2\cos^2 x_e\right\} + \right.$$

$$\left. + j\left\{-\frac{1}{3x_e} - \frac{1}{2x_e^4}(-\sin x_e - \tfrac{1}{2}\sin 2x_e + x_e\cos x_e + x_e\cos 2x_e + \tfrac{1}{2}x_e^2\sin 2x_e)\right\}\right], \quad (11)$$

in which $z = jx_e (x_e > 0)$.

We see that for $x_e \geqslant 0$ the real part of $Y_e$ according to eq. (11) equals $G_e$ according to eqs (7) and (8) as was to be expected.

It can be shown that for $x_e \geqslant 0$ the imaginary part is bounded; (see eqs 12a, b) so we can drop the assumption $x_e > 0$ and we may conclude that the imaginary part of eq. (11) represents the susceptance $B(jx_e)$ we are looking for:

$$B_e(jx_e) = \frac{16 I_a}{V_T}\frac{1}{x_e^4}\left\{-\frac{x_e^3}{3} + \tfrac{1}{2}\sin x_e + \tfrac{1}{4}\sin 2x_e - \tfrac{1}{2}x_e\cos x_e - \tfrac{1}{2}x_e\cos 2x_e - \tfrac{1}{4}x_e^2\sin 2x_e\right\} =$$

$$= \frac{I_a}{V_T} b_e(x_e), \quad (12)$$

with $x_e \geqslant 0$.

The function $b_e(x_e)$ is plotted in fig. 2.

For small values of $x_e$ we have

$$b_e(x_e) \approx \tfrac{4}{5} x_e; \quad (12a)$$

for large values of $x_e$

$$b_e(x_e) \approx -\frac{16}{3x_e}\left(1 + \frac{3}{4x_e}\sin 2x_e\right). \quad (12b)$$

**4. Discussion of the theoretical results**

From figs 1 and 2 we see that as a function of the transit angle the total-emission susceptance and the exponential susceptance behave quite similarly. With small transit angles they increase linearly, go through a maximum, become zero at a certain transit angle, have a negative minimum and approach zero from below in inverse proportion to the transit angle.

With small transit angles ($0 \leqslant x_e \leqslant 0{\cdot}5$) from (6) and (12) we can estimate the ratio of the total-emission susceptance to the exponential susceptance:

$$\frac{B_T}{B_e} \approx \tfrac{5}{2}\sqrt{\pi}\,\eta^{-5/2} \exp \eta\,, \tag{13}$$

with $\eta = V_a/V_T$.

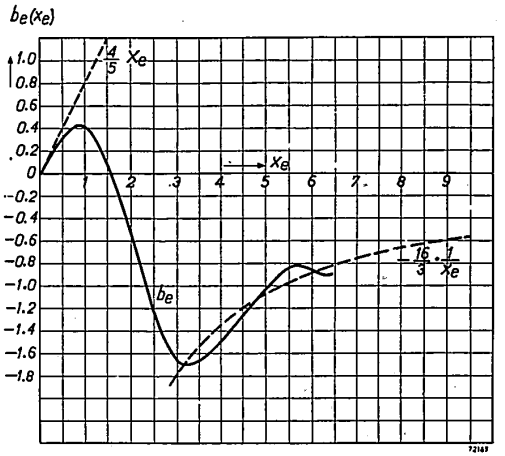



Fig. 2. Transit-time function of the exponential susceptance; $x_e$ is the transit angle of the slowest electrons that are able to reach the anode.

In fig. 3 the right-hand side of eq (13) is given as a function of $\eta$. At $\eta = 5/2$ a minimum occurs having a value of $(5/2)\sqrt{\pi}\,(e/2{\cdot}5)^{2{\cdot}5} \approx 5{\cdot}4$. So with $0 \leqslant x_e \leqslant 0{\cdot}5$ we may conclude that the total-emission susceptance is much more important than the exponential susceptance. With large transit angles ($x_0 \geqslant 2{\cdot}5$) we have analogously

$$\frac{B_T}{B_e} \approx \tfrac{3}{8}\sqrt{\pi}\,\eta^{-5/2} \exp \eta\,. \tag{14}$$

This is plotted in fig. 4, from which we see that also for large values of $x_0$ the total-emission susceptance is predominant (a minimum having a value of $\frac{3}{8}\sqrt{\pi}\,(\frac{2}{3}\,e)^{5/2} \approx 1{\cdot}64$ now occurs at $\eta = 1{\cdot}5$).

Since the theory developed here is only valid for such values of $\eta$ that the d.c. field may supposed to be linear, we may say that in general for the conditions we are investigating the total-emission conductance is practically the only term to be taken into account ($x_e \leqslant 0{\cdot}5$).

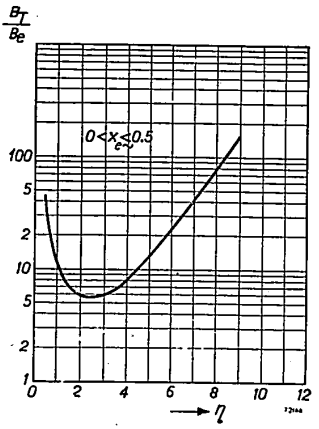



Fig. 3. Ratio of total-emission susceptance to exponential susceptance as a function of anode voltage ($\eta = V_a/V_T$) in the case of small transit angles.

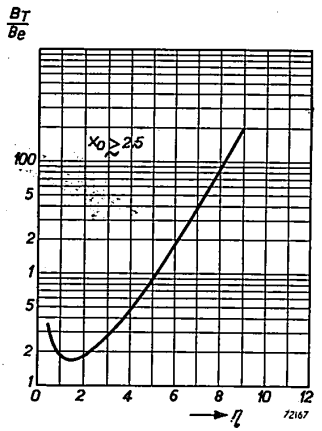



Fig. 4. The same as fig. 3, but with large transit angles.

## 5. Comparison with experiments

Several experiments have been performed on diode susceptance with negative anode voltages. In these experiments the transit angle is varied by varying the anode d.c. voltage. Qualitatively all the curves show the same general aspect. (See, e.g., figs 8 and 9 of reference [5]) and figs 10, 11 and 12 of reference [4]).) The trend of the transit-time dependence agrees well with the theory outlined above. At moderate negative anode volt-

ages the order of magnitude of $B$ is also what may be expected. A discrepancy exists, however, between the experimental and the theoretical values of the transit angle at which the susceptance goes through zero and where the minima and maxima occur. The experimental equivalent transit angles are much larger than those according to the linear-field theory. Since with not too small values of the saturation current the exponential point (i.e. the anode voltage at which the potential minimum is beginning to show up) lies at about 1 volt negative, strong deviations from the linear field will occur in this voltage region, giving rise to much larger transit angles than those occurring in the absence of space charge. We hope to come back upon this subject.

*Eindhoven, November 1951*

REFERENCES


1) C. N. Smyth, Nature **157**, 841, 1946.
2) A. van der Ziel, Nature **159**, 677, 1947; A. van der Ziel and A. Versnel, Nature, **159**, 640 1947; Philips Res. Rep. **3**, 13-23, 1948.
3) N. T. Lavoo, Proc. Inst. Radio Engrs **35**, 1248-1251, 1946.
4) S. D. Robertson. Bell Syst. tech. J. **28**, 647-655, 1949.
5) G. Diemer and K. S. Knol, Philips Res. Rep. **4**, 321-333, 1949.
6) C. J. Bakker and G. de Vries, Physica **2**, 683-697, 1935.
7) K. S. Knol, M. J. O. Strutt and A. van der Ziel, Physica **5**, 325-334, 1938.
8) N. A. Begovich, J. appl. Phys. **20**, 457-462, 1949.
9) A. van der Ziel, Proc. Inst. Radio Engrs **38**, 562, 1950.
10) G. Diemer, Philips Res. Rep. **6**, 211-223, 1951.
11) G. Diemer and H. Dijkgraaf, Philips Res. Rep. **7**, 45-53, 1952.
12) H. A. Kramers, Atti Congr. Fis. Como 545, 1927.
13) O. Brune, J. Math. Phys. **10**, 191-236, 1931.
14) W. Nijenhuis and F. L. Stumpers, Physica **8**, 289-307, 1941.

R 195 *Philips Res. Rep.* 7, 259-269, 1952

# A GENERAL NETWORK THEOREM, WITH APPLICATIONS

by B. D. H. TELLEGEN

621.392.1

**Summary**

It is proved that in a network configuration, for branch currents $i$ satisfying the node equations and branch voltages $v$ satisfying the mesh equations, $\Sigma iv$ 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 given, arbitrarily varying voltages are applied on a $2n$-pole at rest 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.

**Résumé**

Il a été démontré que dans une configuration de réseau pour courants de branche $i$ répondant aux équations nodales et pour tensions de branche $v$ répondant aux équations de maille, $\Sigma iv$ sommé sur tous les branches est égal à zéro. Ce théorème permet de prouver le théorème de l'énergie et la relation de réciprocité des réseaux et de démontrer que si des tensions données, variant arbitrairement, sont appliquées à un $2n$-pôle au repos, la différence entre les énergies électrique et magnétique dépendra à tout instant seulement de la matrice d'admittance du $2n$-pôle et non pas du réseau particulier, utilisé pour sa réalisation.

**Zusammenfassung**

Es wird gezeigt, daß in einer Netzwerkkonfiguration bezüglich Zweigströme $i$, die den Knotengleichungen genügen, und Zweigspannungen $v$, die den Maschengleichungen genügen, $\Sigma iv$, summiert über alle Zweige, gleich Null ist. Dieses Theorem ermöglicht es, das Energietheorem und die Reziprozitätsbeziehung von Netzwerken nachzuweisen, sowie zu zeigen, daß bei Anwendung gegebener, willkürlich veränderbarer Spannungen auf einen $2n$-Pol in Ruhe, die Differenz zwischen der elektrischen und der magnetischen Energie in jedem Zeitpunkt ausschließlich von der Admittanzmatrix des $2n$-Poles abhängt und nicht von dem besonderen Netzwerk, das zu ihrer Verwirklichung benutzt wird.

## 1. Introduction

Many properties of electrical networks can be derived from their impedance functions alone, without a detailed knowledge of their circuit diagrams. Certain properties, however, cannot be established in this way; for example, in order to prove the energy theorem and the reciprocity relation of networks and to investigate the distribution of the supplied energy in electric, magnetic, and dissipated energy, the inner structure of the network has to be taken into account. In this paper a general theorem is given which can be used in such cases. Considerations leading to the

theorem are inherently present in various network investigations; the theorem itself was, so far as the author is aware, never explicitly stated.

## 2. The general theorem

We prove the following

*Theorem:*

*In a network configuration, imagine branch currents $i$ such that for every node $\Sigma i = 0$, imagine branch voltages $v$ such that for every mesh $\Sigma v = 0$, and for every branch let the positive direction of the current be from the + to the — denoting the positive polarity of the voltage (fig. 1). Then $\Sigma iv = 0$, where the summation is over all branches.*

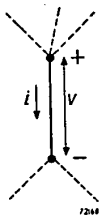



Fig. 1. Positive senses of current and voltage of a branch.

Since for every mesh $\Sigma v = 0$, there exist node potentials $V$ such that the voltage on each branch is equal to the difference between the potentials of its end points. We denote the potentials of the nodes $k$ and $l$ by $V_k$ and $V_l$, and the voltage on the branch connecting $k$ and $l$ by $v_{kl}$, where $v_{kl} = V_k - V_l$. We denote the current in the branch flowing from $k$ to $l$ by $i_{kl}$. Then

$$i_{kl}v_{kl} = i_{kl}(V_k - V_l) = i_{kl}V_k + i_{lk}V_l .$$

After applying this to all terms of $\Sigma iv$, we collect the terms containing $V_k$, $\Sigma i_{kl}V_k = V_k \Sigma i_{kl}$, where the summation is over all nodes that are connected to $k$ by a branch. Since the $i_{kl}$ now considered are the currents flowing away from $k$, their sum is zero and thus the sum of the considered terms of $\Sigma iv$ is zero. The same holds for the sum of the terms containing any other node potential, so that $\Sigma iv = 0$.

The theorem is the network equivalent of the well-known theorem that the volume integral of the scalar product of a solenoidal vector (comparable with $i$) and an irrotational vector (comparable with $v$) is zero. The theorem holds for all types of network, linear and nonlinear, constant and variable, passive and active. The $i$'s and $v$'s in the theorem may be complex quantities satisfying the node equations and the mesh equations, respectively. The

theorem can be applied to networks provided with terminal pairs by considering the terminal pairs as constituting branches of the network. If we take the positive senses of the current and the voltage of a terminal pair as indicated in fig. 2, we can write the theorem as $\Sigma i_t v_t = \Sigma i_b v_b$, where the first summation is over the terminal pairs and the second over the internal branches.

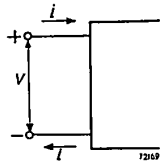



Fig. 2. Positive senses of current and voltage of a terminal pair.

### 3. The energy theorem of networks

If $i$ and $v$ denote branch currents and voltages simultaneously present in a network, $\Sigma iv = 0$ means that at any instant the total power consumption is zero. This constitutes the energy theorem of networks. For networks with terminal pairs it implies that the power absorbed by the branches is equal to the power delivered to the network through the terminal pairs.

If $I$ and $V$ are complex quantities representing sinusoidal branch currents and voltages of the same frequency in a constant, linear network we may write

$$\Sigma I_t V_t^* = \Sigma I_b V_b^* , \tag{1}$$

where the asterisk denotes the conjugate complex quantity, and the first summation is again over the terminal pairs and the second over the internal branches. The left-hand side of (1) equals the active power $+j$ reactive power delivered to the network. The right-hand side of (1) may be written as

$$\sum_k R_k I_k I_k^* + j\omega \sum_p C_p V_p V_p^* - j\omega \sum_{m,n} L_{mn} I_m I_n^* = W_{\mathrm{av}} + 2j\omega(U_{\mathrm{av}} - T_{\mathrm{av}}),$$

where the $R_k$ are the resistances, the $C_p$ are the capacitances, and the $L_{mn}$ are the self and mutual inductances of the network; $W_{\mathrm{av}}$, $U_{\mathrm{av}}$, and $T_{\mathrm{av}}$ are the average dissipated power, the average electric energy, and the average magnetic energy, respectively. Thus we find

$$\text{active power} = W_{\mathrm{av}}, \quad \text{reactive power} = 2\omega\,(U_{\mathrm{av}} - T_{\mathrm{av}}) .$$

These results are related to results obtained by Bode [1]).

## 4. The reciprocity relation

Let us consider two states of a constant, linear four-pole, the first state being denoted by unprimed quantities, the second by primed quantities. We can then write

$$I_1V_1' + I_2V_2' = \Sigma I_bV_b',$$
$$I_1'V_1 + I_2'V_2 = \Sigma I_b'V_b,$$

the subscripts 1 and 2 denoting the terminal pairs, the subscript $b$ the internal branches. Considering a branch with impedance $Z$, for that branch we have

$$I_bV_b' = I_bZI_b' = I_b'ZI_b = I_b'V_b.$$

Considering a set of $n$ coupled coils for which

$$V_{bi} = \sum_{k=1}^{n} Z_{ik}\, I_{bk}, \qquad (i = 1, \ldots, n;\ Z_{ik} = Z_{ki}),$$

we have

$$\sum_{i=1}^{n} I_{bi}V_{bi}' = \sum_{i=1}^{n} I_{bi} \sum_{k=1}^{n} Z_{ik}\, I_{bk}' = \sum_{k=1}^{n} I_{bk}' \sum_{i=1}^{n} Z_{ki}I_{bi} = \sum_{k=1}^{n} I_{bk}'V_{bk}.$$

Considering an ideal transformer for which

$$I_{b_1} = -T\, I_{b_2},$$
$$V_{b_2} = \quad T\, V_{b_1},$$

we have

$$I_{b_1}V_{b_1}' + I_{b_2}V_{b_2}' = -TI_{b_2}V_{b_1}' + I_{b_2}TV_{b_1}' = 0,$$

and also

$$I_{b_1}'V_{b_1} + I_{b_2}'V_{b_2} = 0.$$

Thus we see that for a four-pole composed of the elements under consideration

$$\Sigma I_bV_b' = \Sigma I_b'V_b,$$

and hence

$$I_1V_1' + I_2V_2' = I_1'V_1 + I_2'V_2. \qquad (2)$$

This is the reciprocity relation of four-poles in its general form. The proof shows that for this relation to hold, the four-pole need not be passive. It may contain negative resistances.

There are four important special cases of (2).

With $I_2 = 0$ and $I_1' = 0$, we get $V_2/I_1 = V_1'/I_2'$.
With $V_2 = 0$ and $V_1' = 0$, we get $I_2/V_1 = I_1'/V_2'$.
With $I_2 = 0$ and $V_1' = 0$, we get $V_2/V_1 = -I_1'/I_2'$.
With $V_2 = 0$ and $I_1' = 0$, we get $I_2/I_1 = -V_1'/V_2'$.

## 5. The distribution of energy

Let us consider a linear, constant, passive $2n$-pole the admittance matrix

of which is given as a function of the frequency parameter $\lambda$, and let it contain no energy: no voltages on the capacitors, no currents through the coils. At $t = 0$ we apply arbitrarily varying voltages to the $2n$-pole, which causes a certain power to be consumed by it. This power is partly dissipated in the resistors, partly stored as electric energy in the capacitors and as magnetic energy in the coils. If we know the $2n$-pole network we can at any instant calculate this distribution of energy. We know that a given admittance matrix can be realized by different networks. We now ask whether anything can be said about the distribution of energy that depends only on the admittance matrix and not on the particular network used for realizing it *).

To answer the question, imagine a $2n$-pole network composed of resistances, capacitances, inductances, and ideal transformers **). At $t = 0$ constant voltages $v_1, \ldots, v_n$ are applied to the $2n$-pole. We suppose the admittance matrix to have no pole at $\lambda = \infty$ (as would be the case if the network contained capacitances in parallel to its terminal pairs). Then at $t = 0$ the $2n$-pole currents $i_1, \ldots, i_n$ remain finite.

As a consequence, immediately after applying the voltages the $2n$-pole will contain no energy, i.e. the currents through the inductances and the voltages on the capacitances will still be zero. We further suppose the admittance matrix to have no pole at $\lambda = 0$ (as would be the case if the network contained inductances in parallel to its terminal pairs). Then at $t = \infty$ the $2n$-pole currents remain finite. Admittance matrices with poles at $\lambda = \infty$ or $\lambda = 0$ can be considered as limiting cases of those without such poles, and therefore these will not be investigated separately.

Let us begin by considering a $2n$-pole the admittance matrix of which has only two poles, $\lambda_a$ and $\lambda_b$. The poles may be real or conjugate complex. We suppose them to be unequal (equal poles can be considered as a limiting case of unequal ones). The poles determine the free oscillations of the $2n$-pole with short-circuited terminal pairs. For positive $t$ every $2n$-pole current is the sum of a term with $e^{\lambda_a t}$, a term with $e^{\lambda_b t}$, and a constant term. As the $2n$-pole voltages are constant, we can write the power delivered to the $2n$-poles as

$$i_1 v_1 + \ldots + i_n v_n = G_a e^{\lambda_a t} + G_b e^{\lambda_b t} + G_0. \qquad (3)$$

*) It was this question, for a two-pole composed of resistors and capacitors only, that induced me to undertake the investigations reported in this paper. This question was put to me by Dr A. J. Staverman and Dr F. Schwarzl of "Kunststoffeninstituut T.N.O." Delft, Netherlands, who were led to it by their study of visco-elastic matter, the electric equivalent of which is a system composed of resistors and capacitors.

**) Networks containing gyrators [2]) have to be excluded from the following considerations. Since a gyrator of which one terminal pair is connected to a capacitance behaves as an inductance, in such networks the distinction between electric and magnetic energy loses its interest.

IPR2025-00217 EXHIBIT 2010 Page 267 of 485

If the admittance matrix and $v_1, \ldots, v_n$ are known, $G_a$, $G_b$, $G_0$ can be calculated in several ways.
The current and the voltage of a resistance $R_k$ we write as

$$i_{Rk} = A_{ak}e^{\lambda_a t} + A_{bk}e^{\lambda_b t} + A_{0k}, \quad v_{Rk} = R_k(A_{ak}e^{\lambda_a t} + A_{bk}e^{\lambda_b t} + A_{0k}).$$

The current and the voltage of an inductance $L_m$ we write as

$$i_{Lm} = B_{am}(e^{\lambda_a t}-1) + B_{bm}(e^{\lambda_b t}-1), \; v_{Lm} = L_m(B_{am}\lambda_a e^{\lambda_a t} + B_{bm}\lambda_b e^{\lambda_b t}),$$

since $i_{Lm} = 0$ at $t = 0$.
The voltage and the current of a capacitance $C_p$ we write as

$$v_{Cp} = D_{ap}(e^{\lambda_a t}-1) + D_{bp}(e^{\lambda_b t}-1), \quad i_{Cp} = C_p(D_{ap}\lambda_a e^{\lambda_a t} + D_{bp}\lambda_b e^{\lambda_b t}),$$

since $v_{Cp} = 0$ at $t = 0$.

We now apply our general theorem and take the currents at $t = t'$ and the voltages at $t = t''$. In section 4 it has been seen that for the ideal transformers $\Sigma iv = 0$, and so we can disregard them here. Thus we can write

$$\begin{aligned} &G_a e^{\lambda_a t'} + G_b e^{\lambda_b t'} + G_0 = \\ &= \sum_k R_k(A_{ak}e^{\lambda_a t'} + A_{bk}e^{\lambda_b t'} + A_{0k})(A_{ak}e^{\lambda_a t''} + A_{bk}e^{\lambda_b t''} + A_{0k}) + \\ &+ \sum_m L_m\{B_{am}(e^{\lambda_a t'}-1) + B_{bm}(e^{\lambda_b t'}-1)\}\{(B_{am}\lambda_a e^{\lambda_a t''} + B_{bm}\lambda_b e^{\lambda_b t''}) + \\ &+ \sum_p C_p(D_{ap}\lambda_a e^{\lambda_a t'} + D_{bp}\lambda_b e^{\lambda_b t'})\{D_{ap}(e^{\lambda_a t''}-1) + D_{bp}(e^{\lambda_b t''}-1)\}. \end{aligned}$$

Since this must hold for all positive values of $t'$ and $t''$, we find

$$\begin{aligned} G_a &= \sum_k R_k A_{ak}A_{0k} - \lambda_a \sum_p C_p D_{ap}^2 \quad - \lambda_a \sum_p C_p D_{ap} D_{bp}, \\ G_b &= \sum_k R_k A_{bk}A_{0k} - \lambda_b \sum_p C_p D_{bp}^2 \quad - \lambda_b \sum_p C_p D_{ap} D_{bp}, \\ 0 &= \sum_k R_k A_{ak}A_{0k} - \lambda_a \sum_m L_m B_{am}^2 \quad - \lambda_a \sum_m L_m B_{am}B_{bm}, \\ 0 &= \sum_k R_k A_{bk}A_{0k} - \lambda_b \sum_m L_m B_{bm}^2 \quad - \lambda_b \sum_m L_m B_{am}B_{bm}, \\ 0 &= \sum_k R_k A_{ak}A_{bk} + \lambda_a \sum_m L_m B_{am}B_{bm} + \lambda_b \sum_p C_p D_{ap} D_{bp}, \\ 0 &= \sum_k R_k A_{ak}A_{bk} + \lambda_b \sum_m L_m B_{am}B_{bm} + \lambda_a \sum_p C_p D_{ap} D_{bp}, \end{aligned}$$

and three other relations which we do not need. From the last two equations it follows that

$$\sum_m L_m B_{am} B_{bm} = \sum_p C_p D_{ap} D_{bp}. \tag{4}$$

From the first four equations we then deduce

$$\lambda_a(\sum_m L_m B_{am}^2 - \sum_p C_p D_{ap}^2) = G_a, \; \lambda_b(\sum_m L_m B_{bm}^2 - \sum_p C_p D_{bp}^2) = G_b. \tag{5}$$

The magnetic energy $T$ is given by

$$\begin{aligned} T = \sum_m \tfrac{1}{2} L_m i_{Lm}^2 = \tfrac{1}{2}\sum_m L_m B_{am}^2(e^{\lambda_a t}-1)^2 + \tfrac{1}{2}\sum_m L_m B_{bm}^2(e^{\lambda_b t}-1)^2 + \\ + \sum_m L_m B_{am}B_{bm}(e^{\lambda_a t}-1)(e^{\lambda_b t}-1). \end{aligned}$$

The electric energy $U$ is given by

$$U = \sum_p \tfrac{1}{2} C_p v_{Cp}^2 = \tfrac{1}{2} \sum_p C_p D_{ap}^2 (e^{\lambda_a t} - 1)^2 + \tfrac{1}{2} \sum_p C_p D_{bp}^2 (e^{\lambda_b t} - 1)^2 + \\ + \sum_p C_p D_{ap} D_{bp} (e^{\lambda_a t} - 1)(e^{\lambda_b t} - 1) .$$

By subtraction, using (4) and (5), we arrive at

$$T - U = \frac{G_a}{2\lambda_a} (e^{\lambda_a t} - 1)^2 + \frac{G_b}{2\lambda_b} (e^{\lambda_b t} - 1)^2. \qquad (6)$$

The quantities $G_a$, $G_b$, $\lambda_a$, $\lambda_b$ occurring in the right-hand side of (6) depend, as seen above, only on $v_1, \ldots, v_n$ and the admittance matrix, but not on the particular network used to realize the latter, and so this also holds for $T-U$.

This result, derived for a $2n$-pole at rest on which at $t = 0$ constant voltages are applied and of which the admittance matrix has only two poles, remains valid in more general cases. To investigate $2n$-poles the admittance matrices of which have more than two poles, let us consider the expressions derived for the energy more closely. For an admittance matrix with only one pole, $\lambda_a$ or $\lambda_b$, $T$ only contains terms with $\sum_m L_m B_{am}^2$ or $\sum_m L_m B_{bm}^2$ respectively. If $\lambda_a$ and $\lambda_b$ are both poles, $T$ also contains terms with $\sum_m L_m B_{am} B_{bm}$. Similar remarks apply to $U$. Because of (4), in $T-U$ these latter terms cancel, and thus (6) consists of the sum of a term arising from $\lambda_a$ and a term arising from $\lambda_b$ *). Thus, if the admittance matrix has more than two poles, the corresponding expression for $T-U$ will be equal to the sum of the terms arising from each pole separately, and thus (6) can immediately be extended to any number of poles. It can also be extended to variable voltages, as will be shown in section 7. Thus we have the

*Theorem:*

*If on a linear, constant, passive, gyratorless 2n-pole at rest given, arbitrarily varying voltages are applied, the difference between the magnetic and the electric 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.*

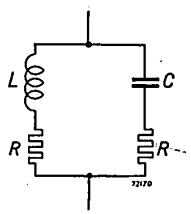



Fig. 3. Two-pole network.

*) This property is related to the conjugate property derived by Heaviside [3]).

Since for the energy $A$ delivered to the $2n$-pole we can write $A = W+T+U$, where $W$ is the dissipated energy, the above theorem not only holds for $T-U$ but also for $W+2T$ and for $W+2U$. If one of the energies is zero, as in resistanceless, inductanceless, or capacitanceless $2n$-poles, the theorem holds for any of the two remaining energies.

That in the general case $W$, $T$, and $U$ are not separately determinable from the admittance matrix alone, is shown by the two-pole network of fig. 3, where $L/C = R^2$. As is well known, the impedance of the two-pole equals $R$. Thus, when starting from rest, $T$ and $U$ of the two-pole network are not determinable from the impedance function; only $T-U$ is, and here it is zero.

## 6. A theorem of Heaviside

From (6) we can draw an interesting conclusion. If we disregard admittance matrices that have poles at imaginary values of $\lambda$, the real parts of $\lambda_a$ and $\lambda_b$ will both be negative. Therefore for large $t$ we get

$$T - U = \frac{G_a}{2\lambda_a} + \frac{G_b}{2\lambda_b}. \tag{7}$$

The energy $A$ delivered to the $2n$-pole is determined according to (3) by

$$\frac{\mathrm{d}A}{\mathrm{d}t} = G_a e^{\lambda_a t} + G_b e^{\lambda_b t} + G_0.$$

Integrating with respect to $t$ between the limits 0 and $t$, we find

$$A = \frac{G_a}{\lambda_a}(e^{\lambda_a t}-1) + \frac{G_b}{\lambda_b}(e^{\lambda_b t}-1) + G_0 t.$$

For large $t$ we get

$$A = G_0 t - \frac{G_a}{\lambda_a} - \frac{G_b}{\lambda_b}. \tag{8}$$

The quantity $G_0$ is the power delivered to the $2n$-pole at large $t$, which power is dissipated. Therefore $G_0 t$ is the amount of energy that would have been dissipated up to the time $t$ if from $t = 0$ the dissipation rate would have had the value $G_0$, which energy we denote by $W'$. From (7) and (8) it then follows that

$$A - W' = 2(U - T). \tag{9}$$

This relation remains valid if the admittance matrix has more than two poles. It expresses a theorem due to Heaviside and proved by Lorentz [2]), which may be stated as follows:

*Theorem:*

*If on a linear, constant, passive, gyratorless 2n-pole at rest suddenly constant voltages are applied, then, when the final state has been reached, the total energy delivered to the 2n-pole exceeds the energy representing the loss by dissipation at the final rate, supposed to start at once, by twice the excess of the electric over the magnetic energy.*

## 7. Variable voltages

To investigate the distribution of energy when variable voltages are applied to a $2n$-pole at rest, we first consider a four-pole the admittance matrix of which has two poles, on whose first terminal pair a constant voltage $v_1$ is applied at $t = t_1$, and on whose second terminal pair a constant voltage $v_2$ is applied at $t = t_2$. For $t > t_1$ and $t > t_2$ the currents and the voltages are the superposition of those due to $v_1$ and those due to $v_2$. The currents to the terminal pairs can be written as

$$i_1 = \{G_{a11}\, e^{\lambda_a(t-t_1)} + G_{b11}\, e^{\lambda_b(t-t_1)} + G_{011}\}\, v_1 + \\ + \{G_{a12}\, e^{\lambda_a(t-t_2)} + G_{b12}\, e^{\lambda_b(t-t_2)} + G_{012}\}\, v_2\,,$$

$$i_2 = \{G_{a12}\, e^{\lambda_a(t-t_1)} + G_{b12}\, e^{\lambda_b(t-t_1)} + G_{012}\}\, v_1 + \\ + \{G_{a22}\, e^{\lambda_a(t-t_2)} + G_{b22}\, e^{\lambda_b(t-t_2)} + G_{022}\}\, v_2\,.$$

If the admittance matrix is known, the $G$'s can be calculated. They do not depend on $v_1$, $v_2$, $t_1$, $t_2$. The power delivered to the four-pole amounts to

$$i_1 v_1 + i_2 v_2 = G_{a11}\, v_1^2\, e^{\lambda_a(t-t_1)} + G_{a12}\, v_1 v_2 \{e^{\lambda_a(t-t_1)} + e^{\lambda_a(t-t_2)}\} + G_{a22}\, v_2^2\, e^{\lambda_a(t-t_2)} + \\ + G_{b11}\, v_1^2\, e^{\lambda_b(t-t_1)} + G_{b12}\, v_1 v_2 \{e^{\lambda_b(t-t_1)} + e^{\lambda_b(t-t_2)}\} + G_{b22}\, v_2^2\, e^{\lambda_b(t-t_2)} + \\ + G_{011}\, v_1^2 + 2\, G_{012}\, v_1 v_2 + G_{022}\, v_2^2\,.$$

For the current through the resistance $R_k$ we take the same expression as in section 5.

The current through the inductance $L_m$ we now write as

$$i_{Lm} = B_{a1m} v_1 \{e^{\lambda_a(t-t_1)} - 1\} + B_{a2m}\, v_2 \{e^{\lambda_a(t-t_2)} - 1\}, \\ + B_{b1m} v_1 \{e^{\lambda_b(t-t_1)} - 1\} + B_{b2m}\, v_2 \{e^{\lambda_b(t-t_2)} - 1\}, \qquad (10)$$

since $i_{Lm} = 0$ at $t = t_1$ if $v_2 = 0$, and at $t = t_2$ if $v_1 = 0$.

The voltage on the capacitance $C_p$ we write accordingly as

$$v_{Cp} = D_{a1p}\, v_1 \{e^{\lambda_a(t-t_1)} - 1\} + D_{a2p}\, v_2 \{e^{\lambda_a(t-t_2)} - 1\} + \\ + D_{b1p}\, v_1 \{e^{\lambda_b(t-t_1)} - 1\} + D_{b2p}\, v_2 \{e^{\lambda_b(t-t_2)} - 1\}.$$

As in section 5, we apply our general theorem with the currents at $t = t'$ and the voltages at $t = t''$. Instead of (4) we now arrive at

$$\sum_m L_m (B_{a1m} v_1 e^{-\lambda_a t_1} + B_{a2m} v_2 e^{-\lambda_a t_2})(B_{b1m} v_1 e^{-\lambda_b t_1} + B_{b2m} v_2 e^{-\lambda_b t_2}) =$$
$$= \sum_p C_p (D_{a1p} v_1 e^{-\lambda_a t_1} + D_{a2p} v_2 e^{-\lambda_a t_2})(D_{b1p} v_1 e^{-\lambda_b t_1} + D_{b2p} v_2 e^{-\lambda_b t_2}).$$

Since this must hold for all values of $t_1$ and $t_2$, we find

$$\left.\begin{aligned} \sum_m L_m B_{a1m} B_{b1m} &= \sum_p C_p D_{a1p} D_{b1p}, \quad \sum_m L_m B_{a2m} B_{b2m} = \sum_p C_p D_{a2p} D_{b2p}, \\ \sum_m L_m B_{a1m} B_{b2m} &= \sum_p C_p D_{a1p} D_{b2p}, \quad \sum_m L_m B_{a2m} B_{b1m} = \sum_p C_p D_{a2p} D_{b1p}. \end{aligned}\right\} \quad (11)$$

By means of these relations, instead of (5) we now arrive at

$$\lambda_a \{\sum_m L_m (B_{a1m} v_1 e^{-\lambda_a t_1} + B_{a2m} v_2 e^{-\lambda_a t_2})(B_{a1m} v_1 + B_{a2m} v_2) -$$
$$- \sum_p C_p (D_{a1p} v_1 e^{-\lambda_a t_1} + D_{a2p} v_2 e^{-\lambda_a t_2})(D_{a1p} v_1 + D_{a2p} v_2)\} =$$
$$= G_{a11} v_1^2 e^{-\lambda_a t_1} + G_{a12} v_1 v_2 (e^{-\lambda_a t_1} + e^{-\lambda_a t_2}) + G_{a22} v_2^2 e^{-\lambda_a t_2},$$

and a similar equation derived therefrom by changing the index $a$ into $b$. Since this must hold for all values of $v_1$ and $v_2$, we find

$$\left.\begin{aligned} \lambda_a (\sum_m L_m B_{a1m}^2 - \sum_p C_p D_{a1p}^2) &= G_{a11}, \\ \lambda_a (\sum_m L_m B_{a1m} B_{a2m} - \sum_p C_p D_{a1p} D_{a2p}) &= G_{a12}, \\ \lambda_a (\sum_m L_m B_{a2m}^2 - \sum_p C_p D_{a2p}^2) &= G_{a22}, \end{aligned}\right\} \quad (12)$$

and three similar relations by changing the index $a$ into $b$. The magnetic energy and the electric energy are given by

$$T = \tfrac{1}{2} \sum_m L_m [B_{a1m} v_1 \{e^{\lambda_a (t-t_1)} - 1\} + B_{a2m} v_2 \{e^{\lambda_a (t-t_2)} - 1\} +$$
$$+ B_{b1m} v_1 \{e^{\lambda_b (t-t_1)} - 1\} + B_{b2m} v_2 \{e^{\lambda_b (t-t_2)} - 1\}]^2,$$
$$U = \tfrac{1}{2} \sum_p C_p [D_{a1p} v_1 \{e^{\lambda_a (t-t_1)} - 1\} + D_{a2p} v_2 \{e^{\lambda_a (t-t_2)} - 1\} +$$
$$+ D_{b1p} v_1 \{e^{\lambda_b (t-t_1)} - 1\} + D_{b2p} v_2 \{e^{\lambda_b (t-t_2)} - 1\}]^2.$$

By subtraction, using (11) and (12), we arrive at

$$T - U = \frac{1}{2\lambda_a} [G_{a11} v_1^2 \{e^{\lambda_a (t-t_1)} - 1\}^2 + 2 G_{a12} v_1 v_2 \{e^{\lambda_a (t-t_1)} - 1\} \{e^{\lambda_a (t-t_2)} - 1\} +$$
$$+ G_{a22} v_2^2 \{e^{\lambda_a (t-t_2)} - 1\}^2] +$$
$$+ \frac{1}{2\lambda_b} [G_{b11} v_1^2 \{e^{\lambda_b (t-t_1)} - 1\}^2 + 2 G_{b12} v_1 v_2 \{e^{\lambda_b (t-t_1)} - 1\} \{e^{\lambda_b (t-t_2)} - 1\} +$$
$$+ G_{b22} v_2^2 \{e^{\lambda_b (t-t_2)} - 1\}^2].$$

As is well known, the application of a variable voltage can be conceived as the successive application of incremental constant voltages. Therefore, if a constant voltage $v$, applied at $t = 0$ in some point $A$ of a network at rest, gives rise, for $t > 0$ in some point $B$ of the network, to a voltage or current $vf(t)$, then a variable voltage $v$, applied at $t = 0$ in the point

$A$ of the network at rest, will give rise, for $t > 0$ in the point $B$, to a voltage or current $\int_0^t v'(\tau)f(t-\tau)\mathrm{d}\tau$. If variable voltages $v_1$ and $v_2$ are applied at $t = 0$ to the four-pole at rest, we can thus for $t = 0$ write instead of (10)

$$i_{Lm} = B_{a1m}\int_0^t v_1'(\tau)\left\{e^{\lambda_a(t-\tau)}-1\right\}\mathrm{d}\tau + B_{a2m}\int_0^t v_2'(\tau)\left\{e^{\lambda_a(t-\tau)}-1\right\}\mathrm{d}\tau + \\ + B_{b1m}\int_0^t v_1'(\tau)\left\{e^{\lambda_b(t-\tau)}-1\right\}\mathrm{d}\tau + B_{b2m}\int_0^t v_2'(\tau)\left\{e^{\lambda_b(t-\tau)}-1\right\}\mathrm{d}\tau.$$

The $B$'s are the same as in (10), but the time functions have changed. Similarly, in $vC_p$, $T$, $U$, and $T-U$ the coefficients remain the same, and only the time functions are altered. This proves our statement in section 5 that also in the case of variable voltages $T-U$ does not depend on the particular network used for realizing a given admittance matrix. This result can be extended to networks with any number of terminal pairs and with admittance matrices having any number of poles.

*Eindhoven, February 1952*

## REFERENCES


1) H. W. Bode, Network analysis and feedback amplifier design, 1945, p. 130.
2) B. D. H. Tellegen, Philips Res. Rep. 3, 81, 1948.
3) O. Heaviside, Electrical Papers, Vol. II, p. 202.
4) O. Heaviside, Electrical Papers, vol. II, p. 412;
H. A. Lorentz, Proc. nat. Acad. Sciences 8, 333, 1922 = Collected Papers, Vol. III, p. 331. See also Balth. van der Pol, Physica, 's-Grav. 4, 585, 1937.

# ON THE SCATTERING MATRIX OF SYMMETRICAL WAVEGUIDE JUNCTIONS *)

by A. E. PANNENBORG 621.392.26:621.317.3

## CHAPTER V. DIRECTIONAL COUPLERS

### V. 1. Condition for perfect directivity

A junction of four waveguides constitutes a directional coupler if the four output leads can be divided into two pairs with the property that no coupling exists between the leads of each pair. Junctions having a scattering matrix $S$ of the form

$$S = \begin{pmatrix} \alpha & \beta & \gamma & \delta \\ \beta & \alpha & \delta & \gamma \\ \gamma & \delta & \alpha & \beta \\ \delta & \gamma & \beta & \alpha \end{pmatrix} \qquad (1)$$

satisfy this definition if one of the transmission coefficients $\beta$, $\gamma$ or $\delta$ vanishes. For lossless junctions it can be proved[1]) that this condition necessarily implies

$$\alpha = 0\,. \qquad (2)$$

The converse is true also: if $\alpha$ vanishes, i.e. if the junction is completely matched, then it is a directional coupler. Hence eq. (2) is a necessary and sufficient condition for a lossless junction to be a directional coupler. This theorem is not limited to junctions described by eq. (1): any completely matched, lossless junction of four waveguides is a directional coupler [18]).

Let us now consider a lossless junction with two-fold planar symmetry consisting of two parallel sections of rectangular waveguide which have one of their side walls in common. If only discontinuities lying in the common wall, assumed to be infinitely thin, are permitted, the elements of the seattering matrix are given by eqs (IV. 11), viz.

$$\left.\begin{aligned} \alpha &= \tfrac{1}{4}(s_1 + s_2) \quad , \\ \beta &= \tfrac{1}{4}(s_1 - s_2 + 2), \\ \gamma &= \tfrac{1}{4}(s_1 + s_2) \quad , \\ \delta &= \tfrac{1}{4}(s_1 - s_2 - 2)\,. \end{aligned}\right\} \qquad (3)$$

To recall the physical meaning of the various elements, the waves scattered by the junction are shown schematically in fig. 5.1 for an incident wave

*) Continued from Philips Res. Rep. 7, 131-157, 169-188, 1952.

in lead (1) only. From now on we confine the discussion to junctions of the type described by eqs (1) and (3).

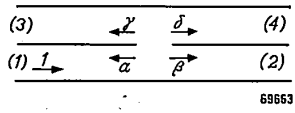



Fig. 5.1. The amplitudes of the scattered waves.

Because

$$\alpha = \gamma, \tag{4}$$

the only way in which the junction can act as a directional coupler ($\alpha = 0$) is the decoupling of lead (3) from lead (1) and of lead (4) from lead (2).

From eqs (2) and (3) it follows that the condition for perfect directivity is equivalent to

$$s_1 + s_2 = 0\,. \tag{5}$$

With the aid of this equation we can deduce the relevant condition for the wave matrix of a directional coupler. From eq. (IV. 21) it is found that in the two-mode basis the wave matrix of a directional coupler is diagonal. It is evident that a junction with a diagonal wave matrix is matched. In that case the +mode, which in general is partially reflected, only suffers a change of phase in passing the discontinuities.

The two quantities commonly used to specify the behaviour of a directional coupler are the directivity $D$ and the coupling $C$. The former indicates quantitatively how effectively the desired decoupling has been realized. It is defined as the ratio of the power coupled in the wanted direction to the power coupled in the unwanted direction. Expressed in decibels it is for the junction of fig. 5.1 given by

$$D = 20 \log \frac{|\delta|}{|\gamma|}. \tag{6}$$

The coupling $C$ is defined as the ratio of the incident power to the power split off in the wanted direction. Thus

$$C = -20 \log |\delta|\,. \tag{7}$$

## V. 2. Bethe-hole coupler

A directional coupler of which the properties seem to be in direct contradiction to the conclusions reached in the preceding section is the well-known Bethe-hole coupler [15]), [19]), [20]). It consists of two rectangular waveguides which are coupled by a circular hole in the centre of their common wide side. The small-hole theory indicates that the coefficient $\delta$ of

this junction vanishes if the guide wavelength equals the cut-off wavelength. From eq. (3), however, it follows that for $\delta = 0$ the quantities $s_1$ and $s_2$ necessarily have their unperturbed values $+1$ and $-1$ respectively, which situation corresponds to the trivial case of no coupling at all between the two sections of waveguide. The solution of this apparent paradox lies in the fact that Bethe's theory is a first-order approximation; accordingly the coefficient $\delta$ of a Bethe-hole coupler is not zero but small of the second order as will now be shown.

The junction is assumed to be lossless. Then we may write (cf. eq. (IV.13))

$$s_1 = e^{i\vartheta_1}, \qquad s_2 = -e^{i\vartheta_2}, \tag{8}$$

where $\vartheta_1$ and $\vartheta_2$ are real. The signs in eqs (8) have been chosen in conformity with eqs (IV. 14) so that for weak coupling the absolute magnitude of both $\vartheta_1$ and $\vartheta_2$ is small compared with unity. With the aid of eqs (8) we obtain from eqs (3)

$$|\gamma| = \tfrac{1}{2}|\sin \tfrac{1}{2}(\vartheta_1 - \vartheta_2)|, \tag{9}$$

$$|\delta|^2 = \tfrac{1}{4}[1 + \cos^2 \tfrac{1}{2}(\vartheta_1 - \vartheta_2) - 2 \cos \tfrac{1}{2}(\vartheta_1 - \vartheta_2) \cos \tfrac{1}{2}(\vartheta_1 + \vartheta_2)]. \tag{10}$$

Apparently $|\gamma|$ is determined by $(\vartheta_1 - \vartheta_2)$ only, whilst $|\delta|$ is a function of $(\vartheta_1 + \vartheta_2)$ as well. Minimizing $|\delta|$ with respect to the latter quantity we have

$$|\delta|_{\min} = \tfrac{1}{2}[1 - \cos \tfrac{1}{2}(\vartheta_1 - \vartheta_2)]. \tag{11}$$

As for small holes

$$|\vartheta_1 - \vartheta_2| \ll 1, \tag{12}$$

the first term of the series expansions for $|\gamma|$ and $|\delta|$ are given by

$$|\gamma| \approx \tfrac{1}{4}|\vartheta_1 - \vartheta_2|, \tag{13}$$

$$|\delta|_{\min} \approx \tfrac{1}{16}(\vartheta_1 - \vartheta_2)^2. \tag{14}$$

Thus it is seen that $|\delta|_{\min}$ is small of the second order; in fact $|\delta|_{\min}$ is just equal to the square of $|\gamma|$. In terms of coupling and directivity, which for the Bethe-hole coupler are *not* given by eqs (6) and (7), but rather by $-20 \log|\gamma|$ and $20 \log|\gamma/\delta|$ respectively, we can conclude that for this type of coupler the directivity is never greater than the coupling.

### V. 3. Systems of identical junctions connected in cascade

In this section we shall treat waveguide systems obtained by cascading a number of identical junctions. In particular the consequences of the condition for perfect directivity will be investigated.

Our first aim is to express the quantities describing the behaviour of the complete system in terms of the elements of the scattering matrix of the individual junctions. The way to achieve this has been prepared in Sec. IV.4 by the introduction of the wave matrix. It was pointed out there

that the wave matrix of a number of junctions in cascade is equal to the product of the wave matrices belonging to each junction individually. If we confine ourselves to systems of $n$ identical junctions in cascade as shown schematically in fig. 5.2, we have

$$\mathfrak{W}'_n = \mathfrak{W}'^n. \tag{15}$$

Here $\mathfrak{W}'_n$ represents the wave matrix of the system and $\mathfrak{W}'$ the wave matrix of a single junction. The primes in eq. (15) indicate that the wave matrices do not refer to the symmetry planes, but rather to the boundary planes between the individual junctions shown in fig. 5.2 by dashed lines. For $\mathfrak{W}'_n$ the reference planes are given by the two outermost planes in fig. 5.2.

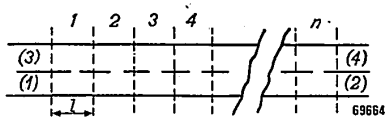



Fig. 5.2. System of $n$ junctions connected in cascade.

The wave matrix $\mathfrak{W}'$ is given by eq. (IV. 28), viz.

$$\mathfrak{W}' = \begin{pmatrix} -\dfrac{2s'_1 s'_2}{s'_1 - s'_2} & \dfrac{s'_1 + s'_2}{s'_1 - s'_2} & 0 & 0 \\ -\dfrac{s'_1 + s'_2}{s'_1 - s'_2} & \dfrac{2}{s'_1 - s'_2} & 0 & 0 \\ 0 & 0 & e^{2i\varphi} & 0 \\ 0 & 0 & 0 & e^{-2i\varphi} \end{pmatrix}, \tag{16}$$

where $2\varphi$ is the phase angle corresponding to the length $l$ of each junction. Introducing the phase angles of $s_1$ and $s_2$ according to eq. (8) we have by eq. (IV. 29)

$$s'_1 = e^{2i\varphi} s_1 = e^{i(2\varphi+\vartheta_1)}, \quad s'_2 = e^{2i\varphi} s_2 = -e^{i(2\varphi+\vartheta_2)}, \tag{17}$$

where $s_1$ and $s_2$ are as usual defined with respect to the symmetry plane of an individual junction.

Insertion of eq. (16) in eq. (15) and subsequent evaluation of the $n$th power of $\mathfrak{W}'$ yields

$$\mathfrak{W}'_n = \begin{pmatrix} \cos n\zeta \pm \dfrac{1 + s'_1 s'_2}{s'_1 - s'_2} \dfrac{\sin n\zeta}{\sin \zeta} & \dfrac{s'_1 + s'_2}{s'_1 - s'_2} \dfrac{\sin n\zeta}{\sin \zeta} & 0 & 0 \\ -\dfrac{s'_1 + s'_2}{s'_1 - s'_2} \dfrac{\sin n\zeta}{\sin \zeta} & \cos n\zeta \mp \dfrac{1 + s'_1 s'_2}{s'_1 - s'_2} \dfrac{\sin n\zeta}{\sin \zeta} & 0 & 0 \\ 0 & 0 & e^{2in\varphi} & 0 \\ 0 & 0 & 0 & e^{-2in\varphi} \end{pmatrix}. \tag{18}$$

The auxiliary quantity $\zeta$ is determined by

$$\cos \zeta = \frac{1 - s_1' s_2'}{s_1' - s_2'} . \tag{19}$$

The ambiguity of sign in two elements of $W_n'$ arises from the definition of $\zeta$, eq. (19), which leaves the sign of $\zeta$ undetermined. Which sign is the correct one will be apparent at a later stage of the discussion.

In Sec. 1 it has been shown that the wave matrix of a coupler with perfect directivity is diagonal. It follows that the condition for perfect directivity for the system described by $W_n'$, eq. (18), is given by

$$\frac{s_1' + s_2'}{s_1' - s_2'} \frac{\sin n\zeta}{\sin \zeta} = 0 . \tag{20}$$

Equation (20) has two solutions, viz.

$$s_1' + s_2' = 0 , \tag{21}$$

and

$$\frac{\sin n\zeta}{\sin \zeta} = 0 . \tag{22}$$

We shall first deal with systems obeying eq. (21). By this condition it is required that each individual junction is by itself a directional coupler; the complete system then is a directional coupler irrespective of the number $n$ of junctions contained in it.

Throughout this chapter we assume that the phase angles $\vartheta_1$ and $\vartheta_2$ of the individual junctions, as introduced in eqs (8), are considerably smaller than $\pi$. Then the condition for directivity, eq. (21), has only one solution, viz.

$$\vartheta_1 = \vartheta_2 \equiv \vartheta . \tag{23}$$

Let us assume that this condition is fulfilled. Equation (19) now becomes

$$\cos \zeta = \cos (2\varphi + \vartheta) , \tag{24}$$

while further

$$\frac{1 + s_1' s_2'}{s_1' - s_2'} = - i \sin (2\varphi + \vartheta) . \tag{25}$$

By eqs (21), (24) and (25) the wave matrix of the system, eq. (18), takes the simple form

$$W_n' = \begin{pmatrix} e^{\pm in(2\varphi+\vartheta)} & 0 & 0 & 0 \\ 0 & e^{\mp in(2\varphi+\vartheta)} & 0 & \\ 0 & 0 & e^{2in\varphi} & 0 \\ 0 & 0 & 0 & e^{-2in\varphi} \end{pmatrix} , \tag{26}$$

It is now clear that the upper sign in the exponentials should be taken

as for vanishing $\vartheta$ the matrix $W_n'$ should represent a simple phase shift.

Equation (26) can be written as

$$W_n' = \begin{pmatrix} e^{in\varphi} & 0 & 0 & 0 \\ 0 & e^{-in\varphi} & 0 & 0 \\ 0 & 0 & e^{in\varphi} & 0 \\ 0 & 0 & 0 & e^{-in\varphi} \end{pmatrix} \begin{pmatrix} e^{in\vartheta} & 0 & 0 & 0 \\ 0 & e^{-in\vartheta} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e^{in\varphi} & 0 & 0 & 0 \\ 0 & e^{-in\varphi} & 0 & 0 \\ 0 & 0 & e^{in\varphi} & 0 \\ 0 & 0 & 0 & 0^{-in\varphi} \end{pmatrix}. \quad (27)$$

Hence the wave matrix $W_n$ referred to the symmetry plane of the system is given by the central matrix in the above product.

The scattering matrix in the two-mode basis $S_{tm}$ can be derived from $W_n$ with the aid of eqs (IV. 20) and (IV. 21). We find

$$S_{tm} = \begin{pmatrix} 0 & e^{in\vartheta} & 0 & 0 \\ e^{in\vartheta} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad (28)$$

whence by eq. (IV. 17)

$$S = \begin{pmatrix} 0 & \beta & 0 & \delta \\ \beta & 0 & \delta & 0 \\ 0 & \delta & 0 & \beta \\ \delta & 0 & \beta & 0 \end{pmatrix}, \quad (29)$$

with

$$\beta = \tfrac{1}{2}(e^{in\vartheta} + 1), \qquad \delta = \tfrac{1}{2}(e^{in\vartheta} - 1). \quad (30)$$

A graphical representation of $\beta$ and $\delta$ is shown in fig. 5.3. The circle is the locus of the end points of the vectors regarded as functions of $n$.

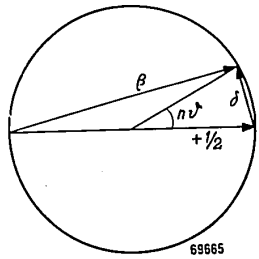



Fig. 5.3. Locus of the transmission coefficients $\beta$ and $\delta$ in the complex plane, both as a function of $n$.

With increasing $n$ the radius of the circle rotates counter-clockwise so that alternately $|\beta|$ and $|\delta|$ go through zero and a maximum respectively. In a system with $n\vartheta \gg \pi$, $\vartheta \ll \pi$, power incident e.g. in lead (1) will be transferred many times to the upper guide and back again to the lower

IPR2025-00217 EXHIBIT 2010 Page 279 of 485

guide before leaving the system by leads (3) and (4). This phenomenon has sometimes been called spatial beating and is known to happen also in parallel waveguides with a long slot in the common wall [20]), [21]).

The conditions corresponding to special values of the coupling can be read directly from fig. 5.3. For equal power partition (e.g.) we must have

$$|\beta| = |\delta| \quad \text{for } n\vartheta = \pi/2, 3\pi/2, 5\pi/2 \ldots ; \tag{31}$$

for complete transfer of power from one guide to the other

$$\beta = 0, \quad |\delta| = 1 \quad \text{for } n\vartheta = \pi, 3\pi, 5\pi , \ldots . \tag{32}$$

The general expression for the coupling $C$ in terms of $n\vartheta$ is by eqs (7) and (30)

$$C = -20 \log |\delta| = -20 \log |\sin \tfrac{1}{2} n\vartheta| . \tag{33}$$

A very effective design for a directional coupler of the type just discussed has been described in a paper by Riblet and Saad [22]). The discontinuities in the common wide side between the waveguides consist here of two slots: one transverse and one longitudinal with respect to the axes of the waveguides, as shown in fig. 5.4. Because the slots are narrow, coupling by the electric field vector is negligible. With regard to the magnetic coupling we recall that the fields corresponding to $s_1$ and $s_2$ are obtained by inserting in the symmetry plane $F_1$ a magnetic and an electric wall respectively.

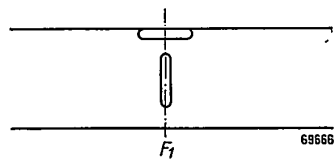



Fig. 5.4. Coupling slots used in directional couplers.

The magnetic wall annuls the coupling through the transverse slot; $s_1$ therefore is determined only by the size of the longitudinal slot and unaffected by the transverse slot. An electric wall in $F_1$ places a short circuit across the longitudinal slot, thereby greatly reducing its effect. Hence $s_2$ is determined almost exclusively by the transverse slot.

For slots short compared with the wavelength the value of $\vartheta_1$ belonging to the longitudinal slot and the value of $\vartheta_2$ belonging to the transverse slot can be calculated with the aid of Bethe's theory. For larger slots measurements taken on each slot separately will yield the dependence of $\vartheta_1$ and $\vartheta_2$ on slot size. By combination of slots with equal values of $\vartheta$ a directional coupler is obtained.

Once one pair of slots has been designed so as to give high directivity and the value of $\vartheta$ is known, the number of junctions needed to give a specified coupling can be calculated with the aid of eq. (33).

As a specific example we quote some results obtained with systems constructed from standard size 3-cm waveguide having 0·400″ × 0·900″ ID. The dimensions of the slots cut in the common wall of thickness 0·050″ are given in fig. 5.5. The phase angle of both $s_1$ and $s_2$ is ten degrees. The dependence of the coupling on the number $n$ of cascaded junctions as calculated from eq. (33) is given for this specific case in table 5.1.

TABLE 5.1

| $n$ | $\lvert\delta\rvert$ | $C$ db. |
|---|---|---|
| 1 | 0·087 | 21·2 |
| 2 | 0·17 | 15·2 |
| 3 | 0·26 | 11·8 |
| 4 | 0·34 | 9·3 |
| 5 | 0·42 | 7·5 |
| 6 | 0·50 | 6.0 |
| 9 | 0·71 | 3·0 |
| 18 | 1·00 | 0·0 |

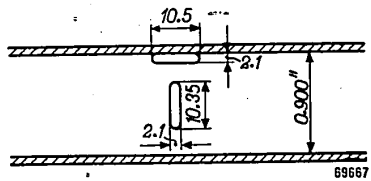



Fig. 5.5. Slot dimensions in millimetres so as to make $\vartheta = 10°$ for a frequency of 9400 Mc/s.

We shall limit the investigation of the consequences of the second directivity condition, viz.

$$\frac{\sin n\zeta}{\sin \zeta} = 0\,, \tag{22}$$

to the practically important case that the two sections of waveguide are coupled by only two identical holes in their common wall. For $n = 2$ eq. (22) becomes

$$\cos \zeta = 0\,. \tag{34}$$

Insertion of eqs (22) and (34) in the general expression for the wave matrix eq. (18), yields

$$W_2' = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & e^{4i\varphi} & 0 \\ 0 & 0 & 0 & e^{-4i\varphi} \end{pmatrix}. \tag{35}$$

The corresponding scattering matrix is found to be

$$S' = \begin{pmatrix} 0 & \beta' & 0 & \delta' \\ \beta' & 0 & \delta' & 0 \\ 0 & \delta' & 0 & \beta' \\ \delta' & 0 & \beta' & 0 \end{pmatrix}, \tag{36}$$

with

$$\beta' = -\tfrac{1}{2}(1 - e^{4i\varphi}), \quad \delta' = -\tfrac{1}{2}(1 + e^{4i\varphi}). \tag{37}$$

The individual junctions used in this type of coupler are usually completely non-directional. This is the case if either $s_1$ or $s_2$ retains its unperturbed value in spite of the presence of the discontinuity. We have seen above that this happens for each of the slots of the junction shown in fig. 5.4. As the field of the lowest mode in a rectangular waveguide is independent of its height, a hole in the narrow side will behave in exactly the same manner as a longitudinal slot cut in the wide side adjacent to the narrow side. Hence for a junction of two rectangular waveguides having a small hole in the common narrow side the value of $s_2$ is unperturbed so that

$$s_2 = -1, \quad \vartheta_2 = 0. \tag{38}$$

Let us restrict the discussion to two-hole systems obeying eq. (38). By substitution of this equation in the expression for cos $\zeta$, eq.(19), the directivity condition eq. (34) becomes

$$e^{i(4\varphi + \vartheta_1)} + 1 = 0. \tag{39}$$

The smallest root of eq. (39) is given by

$$2\varphi = \tfrac{1}{2}(\pi - \vartheta_1). \tag{40}$$

Now $2\varphi$ is the phase angle corresponding to the length of each junction as is also shown in fig. 5.6. Hence it follows from eq. (40) that the distance between the symmetry planes of the individual non-directional junctions

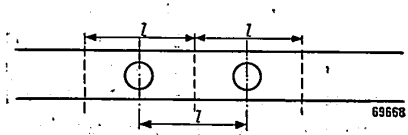



Fig. 5.6. A two-hole coupler.

should be smaller than a quarter wavelength in a two-hole coupler. The amount of the correction depends on the value of $\vartheta_1$, which quantity is decisive for the coupling as will be shown below. With the aid of eq. (39) $\varphi$ can be eliminated from the expressions for $\beta'$ and $\delta'$, eqs (37). Thus

$$\beta' = -\tfrac{1}{2}(1 + e^{-i\vartheta_1}) , \quad \delta' = -\tfrac{1}{2}(1 - e^{-i\vartheta_1}) . \tag{41}$$

Referred to the symmetry plane of the complete coupler the relevant quantities are

$$\beta = e^{-2i\varphi}\, \beta' = -e^{i\vartheta_1}\, \beta' = \tfrac{1}{2}(e^{i\vartheta_1} + 1) , \tag{42}$$

and analogously

$$\delta = \tfrac{1}{2}(e^{i\vartheta_1} - 1) . \tag{43}$$

These equations are seen to be identical with eqs (30) for $n = 1$. This, indeed, was to be expected, because transducer theory does not discriminate between structures, provided their scattering properties are the same, even if internally they differ radically. Hence if a number of two-hole couplers are connected in cascade, the discussion leading up to eqs (30) is fully applicable; in this fashion the theory for binomial couplers e.g. can be derived.

The coupling $C$ of a two-hole coupler in terms of $\vartheta_1$ is given by

$$C = -20 \log |\delta| = -20 \log |\sin \tfrac{1}{2} \vartheta_1| . \tag{44}$$

Once the transmission through one hole and hence $\vartheta_1$ has been determined, the correct spacing for a two-hole directional coupler can be calculated from eq. (40) and the coupling to be expected from eq. (44). In fact by substituting eq. (38) in eqs (3) it can be proved that $|\delta|$ of a two-hole coupler is just twice the absolute value of the transmission coefficient $\delta$ of a single hole.

In a paper by Rosen and Bangert [23]) the conclusion is reached that with a two-hole coupler perfect directivity can never be achieved. As, however, their theory is founded on Bethe's first-order approximation for the coupling through small holes, their results regarding possible second order effects are not conclusive.

### V. 4. Frequency dependence of directional couplers

The fundamental theorem derived in Sec. I.7. is the starting-point for the investigation of the frequency dependence of directional couplers. It states that the phase angle of the reflection coefficient of a lossless termination always increases with frequency. As the quantities $s_1$ and $s_2$, which specify the behaviour of couplers of the type discussed, are essentially reflection coefficients, the above theorem applies to them.

The coefficients involved in the definitions of coupling and directivity are given by

$$\gamma = \tfrac{1}{4}(s_1 + s_2)\,; \qquad \delta = \tfrac{1}{4}(s_1 - s_2 - 2)\,. \tag{45}$$

Note that here $s_1$ and $s_2$ refer to the complete system, which may consist of several cascaded junctions.

Let again

$$s_1 = e^{i\vartheta_1}, \qquad s_2 = -e^{i\vartheta_2}. \tag{8}$$

Substituting eqs (8) in eqs (45) and differentiating the latter with respect to the angular frequency $\omega$ we obtain

$$\frac{\mathrm{d}\gamma}{\mathrm{d}\varphi} = \frac{i}{4}\left(s_1\frac{\mathrm{d}\vartheta_1}{\mathrm{d}\omega} + s_2\frac{\mathrm{d}\vartheta_2}{\mathrm{d}\omega}\right), \qquad \frac{\mathrm{d}\delta}{\mathrm{d}\omega} = \frac{i}{4}\left(s_1\frac{\mathrm{d}\vartheta_1}{\mathrm{d}\omega} - s_2\frac{\mathrm{d}\vartheta_2}{\mathrm{d}\omega}\right). \tag{46}$$

If for $\omega = \omega_0$ the junction is perfectly directional, $s_2$ can be eliminated from eqs (46) with the aid of eq. (5). Thus

$$\left(\frac{\mathrm{d}\gamma}{\mathrm{d}\omega}\right)_{\omega_0} = \frac{is_1}{4}\left(\frac{\mathrm{d}\vartheta_1}{\mathrm{d}\omega} - \frac{\mathrm{d}\vartheta_2}{\mathrm{d}\omega}\right)_{\omega_0}, \qquad \left(\frac{\mathrm{d}\delta}{\mathrm{d}\omega}\right)_{\omega_0} = \frac{is_1}{4}\left(\frac{\mathrm{d}\vartheta_1}{\mathrm{d}\omega} + \frac{\mathrm{d}\vartheta_2}{\mathrm{d}\omega}\right)_{\omega_0}. \tag{47}$$

By the general theorem both

$$\frac{\mathrm{d}\vartheta_1}{\mathrm{d}\omega} > 0 \quad \text{and} \quad \frac{\mathrm{d}\vartheta_2}{\mathrm{d}\omega} > 0\,. \tag{49}$$

Hence it is possible that $(\mathrm{d}\gamma/\mathrm{d}\omega)_{\omega_0}$ vanishes, in which case the directivity is broadband. For broadband coupling it is required that $(\mathrm{d}\delta/\mathrm{d}\omega)_{\omega_0}$ be perpendicular to $\delta$. This can only be achieved in special cases, as can be seen from fig. 5.7. In this diagram the circle is the locus of $\frac{1}{2}s_1$ and $\frac{1}{2}s_2$ regarded as a function of frequency. By the general theorem both $s_1$ and $s_2$ rotate counter-clockwise with increasing frequency. It is immediately clear that perfect directivity is maintained as long as $s_1$ and $s_2$ are exactly opposite. On the other hand it is seen that, as $(\mathrm{d}\delta/\mathrm{d}\omega)_{\omega_0}$ is tangent to the

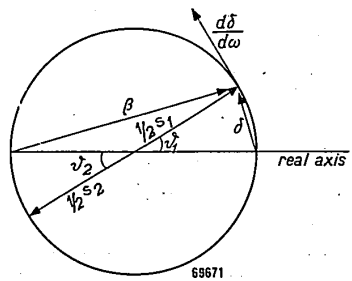



Fig. 5.7. Locus of $\frac{1}{2}s_1$ and $\frac{1}{2}s_2$ as a function of frequency.

circle, $|\delta|$ necessarily increases with frequency for small values of $\vartheta$. However, under conditions for total power transfer, i.e. $|\delta| = 1$, $\vartheta = \pi$, the derivative of $\delta$ with respect to frequency is perpendicular to $\delta$ and the coupling is broadband.

## V. 5. Variable attenuator *)

A cascaded directional coupler can be used with advantage as the basic unit for a variable attenuator. In Sec. 3 it has been shown that if the number $n_0$ of directional couplers connected in cascade obeys the equation

$$n_0\vartheta = \pi\,, \tag{32}$$

power incident in a lead is completely transferred to the lead diametrically opposite. If now it were possible to vary the number of junctions from $n_0$ downwards, the transmission through the system could be controlled and the instrument would be a variable attenuator. This, of course, cannot be done, but in a slightly different way the same effect can be achieved.

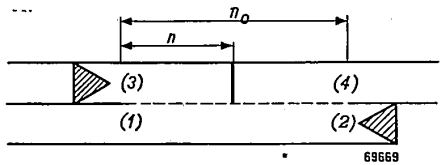



Fig. 5.8. Waveguide system used in a variable attenuator.

Let us consider the system schematically shown in fig. 5.8. A large number $n_0$ of directional couplers each having a small value of $\vartheta$ is connected in cascade. Leads (2) and (3) of the system are terminated by matched loads. A partition has been placed in the upper guide between the junctions numbered $n$ and $n + 1$ from lead (1). Both faces of the partition are completely absorbing, i.e. non-reflecting for waves incident on it.

The scattering matrix $S'_n$ for the system comprising the first $n$ junctions from the left is obtained from eq. (29) by deleting the third and fourth rows and columns because the corresponding leads are terminated by matched loads. Thus

$$S'_n = \begin{pmatrix} 0 & \beta'_n \\ \beta'_n & 0 \end{pmatrix}, \tag{49}$$

with

$$\beta'_n = \tfrac{1}{2}\, e^{2in\varphi}\,(e^{in\vartheta} + 1)\,. \tag{50}$$

Analogously the scattering matrix $S'_{n_0-n}$ for the system comprising the

*) The original ideas underlying the construction of the instruments described in this and the next section are due to Ir W. J. van de Lindt of Philips Research Laboratories.

remaining junctions is found from eq. (29) by omitting the second and third rows and columns. Hence

$$S'_{n_0-n} = \begin{pmatrix} 0 & \delta'_{n_0-n} \\ \delta'_{n_0-n} & 0 \end{pmatrix}, \tag{51}$$

where

$$\delta'_{n_0-n} = \tfrac{1}{2} e^{2i(n_0-n)\varphi} \left(e^{i(n_0-n)\vartheta} - 1\right). \tag{52}$$

Because the partial systems are connected in cascade and both are matched, the scattering matrix $S'$ of the complete system is matched also while its transmission coefficient is simply the product of the transmission coefficients of the partial junctions. Thus we have

$$S' = \begin{pmatrix} 0 & \beta'_n \delta'_{n_0-n} \\ \beta'_n \delta'_{n_0-n} & 0 \end{pmatrix}. \tag{53}$$

By inserting eqs (50) and (52) in eq. (53) we find for the absolute value of the transmission coefficient $t_n$ of the complete system

$$|t_n| = |\beta'_n \delta'_{n_0-n}| = \tfrac{1}{2} |\sin \tfrac{1}{2} n_0 \vartheta + \sin (\tfrac{1}{2} n_0 - n) \vartheta| . \tag{54}$$

Now the position of the partition *can* be varied so that this system forms, indeed, a variable attenuator. The attenuation, referred to the transmission through the instrument with the partition on the extreme left ($n = n_0$), is given by

$$\left|\frac{t_n}{t_0}\right| = \tfrac{1}{2} \left| 1 + \frac{\sin (\tfrac{1}{2} n_0 - n) \vartheta}{\sin \tfrac{1}{2} n_0 \vartheta} \right|. \tag{55}$$

It is, of course, desirable that the minimum attenuation should be zero, i.e. $|t_0| = 1$. If this condition for complete power transfer, eq. (32), is fulfilled, eq. (55) simplifies to

$$\left|\frac{t_n}{t_0}\right| = |t_n| = \cos^2 \tfrac{1}{2} n\vartheta , \qquad n_0 \vartheta = \pi . \tag{56}$$

It is evident that the curve relating attenuation to the position of the thin partition is of a discontinuous nature. In order to obviate this undesirable feature in the actual instruments, a long wedge of absorbing material extending over several junctions is used instead of the partition which was assumed to be short in the idealized theory. The curve is then smoothed out and $n$ in eqs (55) and (56) can be regarded as a continuous variable proportional to the coordinate specifying the position of the wedge.

In the instrument shown in fig. 5.9, which is intended for use in the 3-cm wavelength band, the slots of the directional junctions have dimensions as given in fig. 5.5. From table 5.1 it follows that the number $n$ to satisfy eq. (32) for total power transfer must be equal to 18.

Equations (55) and (56) indicate infinite attenuation for $n = n_0$. Now it is evident that the maximum attenuation obtainable in practice is determined by the properties of the wedge, which up till now was assumed to be completely absorbing. In the actual models built, the maximum attenuation, measured with the wedge right up in output lead (4), fig. 5.8, is over 40 db. On the other hand for settings of attenuation well under 40 db the properties of the wedge are relatively unimportant.

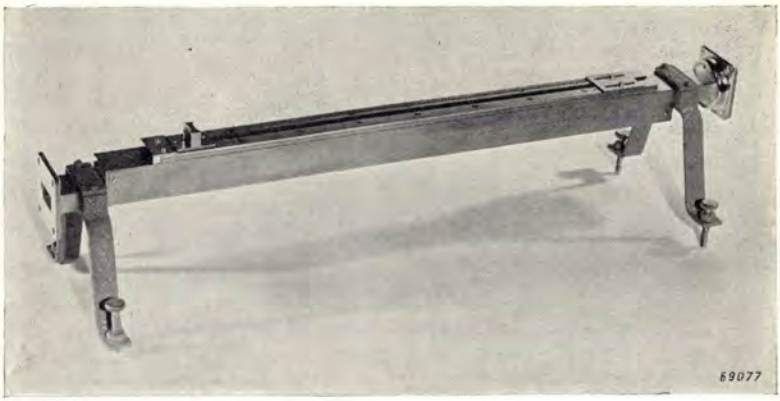



Fig. 5.9. Variable attenuator.

From eq. (53) it is clear that the instrument is matched irrespective of the value of $n$, i.e., the position of the wedge. This is only true if the directivity is perfect and the wedge matched. By careful design this ideal can be closely approximated and measurements indicate that the input standing-wave ratio can be kept below 1·1 in the frequency range between 8600 and 9600 Mc/s.

In the preceding section it was shown that the coupling of a system obeying the condition for total power transfer is broadband. This implies that the insertion loss $t_0$ of the attenuator is independent of frequency, which fact is confirmed by measurements. Broadband coupling does not, however, insure insensitivity to frequency of the calibration curve, as differentiation of eq. (55) with respect to $\vartheta$ demonstrates. To indicate the nature of this effect eq. (55) has been plotted in fig. 5.10 on a logarithmic scale for $n_0 = 18$ and three values of $\vartheta$. The centre curve corresponds to the ideal case of eq. (56) with $n_0\vartheta = 180°$. The point $n = \frac{1}{2}n_0$, $|t_n/t_0| = \frac{1}{2}$ is seen to be independent of $\vartheta$ and hence of frequency. The calibration curves taken on actual models fit the theoretical curves very closely. The cross-over point is found to lie slightly higher than indicated in fig. 5.10.

Possibly this is due to the wedge, the length of which was assumed to be very small in the theory.

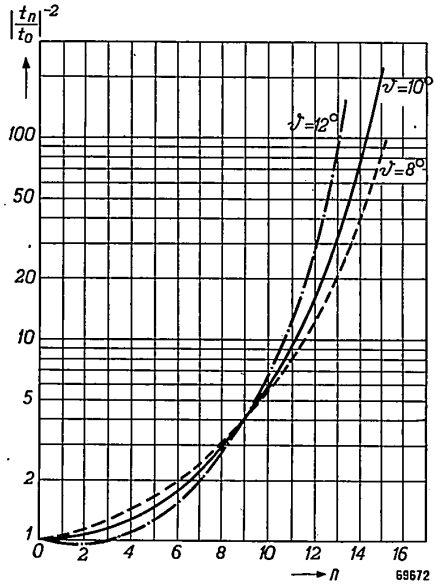



Fig. 5.10. Theoretical calibration curves for the variable attennuator. The spread in the values of $\vartheta$ has been exaggerated for the sake of clarity.

## V. 6. Standard matching transformer

A directional coupler with equal power division is equivalent to a Magic Tee [1]). Now it is known that by insertion of movable plungers in the side arms and a non-reflecting termination in the H-arm of a Magic Tee both phase and magnitude of the reflection suffered by a signal incident in the E-arm can be controlled at will through all possible values [24]). It must be possible to achieve the same effect with the aid of a directional coupler giving equal power division.

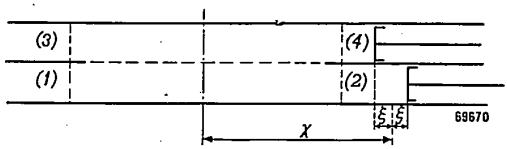



Fig. 5.11. Waveguide system used as a standard matching transformer.

Let us consider the system shown diagramatically in fig. 5.11. It comprises a directional coupler with movable plungers inserted in leads (2) and (4). Let their reflection coefficients referred to the symmetry plane of the coupler be denoted by $g_2$ and $g_4$ respectively. If directivity is assumed to be perfect the behaviour of the coupler is described in the usual notation by

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix} = \begin{pmatrix} 0 & \beta & 0 & \delta \\ \beta & 0 & \delta & 0 \\ 0 & \delta & 0 & \beta \\ \delta & 0 & \beta & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix}. \tag{57}$$

By substitution of the conditions

$$a_2 = g_2 b_2, \qquad a_4 = g_4 b_4, \tag{58}$$

in eq. (57) the scattering matrix for the complete structure having only two output leads, numbered (1) and (3) in fig. 5.11, is found to be

$$S = \begin{pmatrix} \beta^2 g_2 + \delta^2 g_4 & \beta\delta(g_2 + g_4) \\ \beta\delta(g_4 + g_4) & \delta^2 g_2 + \beta^2 g_4 \end{pmatrix}, \tag{59}$$

This matrix takes a particularly simple form if the directional coupler satisfies the condition for equal power partition, viz.

$$|\beta| = |\delta|, \qquad n\vartheta = \pi/2 . \tag{31}$$

By eq. (31) we have in this case

$$\beta = \tfrac{1}{2}(i + 1), \qquad \delta = \tfrac{1}{2}(i - 1). \tag{60}$$

Elimination of $\beta$ and $\delta$ from eq. (59) with the aid of eq. (60) yields

$$S = \tfrac{1}{2} \begin{pmatrix} i(g_2 - g_4) & -(g_2 + g_4) \\ -(g_2 + g_4) & -i(g_2 - g_4) \end{pmatrix}. \tag{61}$$

The form of the elements of $S$ now obtained suggests the introduction of the phase angles $\chi$ and $\xi$, as shown in fig. 5.11, to specify the location of the plungers. The angle $\chi$ corresponds to the distance between the reference plane of $S$ and the median plane between the two plungers. The angle $2\xi$ corresponds to the distance between the two plungers. For lossless plungers we have

$$g_2 = e^{2i(\chi+\xi)}, \qquad g_4 = e^{2i(\chi-\xi)}. \tag{62}$$

The scattering matrix $S$, eq. (61), can now be written as a function of $\chi$ and $\xi$ only. Thus finally

$$S = \begin{pmatrix} -\sin 2\xi\, e^{2i\chi} & -\cos 2\xi\, e^{2i\chi} \\ -\cos 2\xi\, e^{2i\chi} & \sin 2\xi\, e^{2i\chi} \end{pmatrix}. \tag{63}$$

Equation (63) clearly fulfils the general requirements imposed on the scattering matrix of a lossless junction. If $\xi$ and $\chi$ are considered as variables, it is clear that the elements of $S$ can attain all possible values. But the most important point to note is the fact that the magnitude of the elements is a function of $\xi$ only, whilst their phase angles are controlled independently by $\chi$. It is here that the superiority for this application of the directional coupler over the Magic Tee becomes apparent. The scattering matrix of a Magic Tee with plungers in its arms, it is true, may be brought in the form of eq. (63) by the introduction of suitable coordinates for the plunger positions. But the two plungers are physically spaced wide apart so that it is difficult to realize a mechanical drive for the separate control of $\xi$ and $\chi$. In our case the two plungers are adjacent and the required mechanism is quite simple as is shown in fig. 5.12.

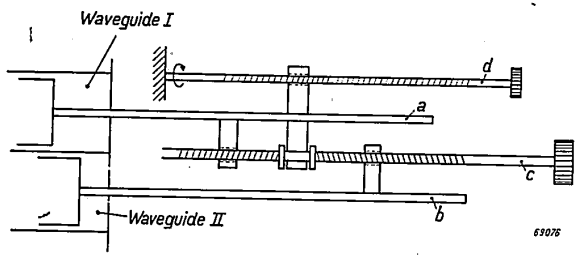



Fig. 5.12. Outline of the plunger drive needed for separate control of $\xi$ and $\chi$.

The advantages of an instrument which allows separate control of modulus and argument of its reflection coefficient are evident. In microwave measurement work the diagram most often used is a polar plot of reflection coefficients. The controls of the instrument hence correspond to the natural coordinates of this diagram. Especially in measurements involving the plotting of lines of constant value of certain quantities, such as power or frequency, on the reflection-coefficient diagram (Rieke-diagram) such an instrument is highly desirable.

There are two methods to measure unknown reflection coefficients with the instrument. The first consists in its use as a variable standard termination. A matched load is connected to lead (3). The reflection coefficient $\Gamma$ looking into lead (1) is then given by

$$\Gamma = -\sin 2\xi \, e^{2i\chi}. \tag{64}$$

Hence if the scale readings of the controls are calibrated in terms of $\xi$ and $\chi$, the reflection coefficient $\Gamma$ is known and it can be compared with the unknown reflection by known bridge methods.

In the other measuring method the instrument is used as a standard matching transformer. Then the termination of which the reflection coefficient is to be determined is connected to output lead (3). The condition for a match in lead (1) under these circumstances will now be derived. Let the unknown reflection coefficient be denoted by $g_3$. Then

$$a_3 = g_3 b_3 . \tag{65}$$

The reflection coefficient in lead (1), found by insertion of eq. (65) in eq. (63), is given by

$$\frac{b_1}{a_1} = \frac{g_3 e^{2i\chi} - \sin 2\xi}{e^{-2i\chi} - g_3 \sin 2\xi}. \tag{66}$$

Hence the condition for a match on the input side of the instrument is

$$g_3 = \sin 2\xi \; e^{-2i\chi}. \tag{67}$$

By comparison of eq. (67) with eq. (63) it can be concluded that the input lead is matched if the reflection coefficient looking into the output lead is complex conjugate to the reflection coefficient of the termination. This theorem applies to any lossless transformer, as by a suitable choice of reference planes the scattering matrix of such a structure can always be brought into the form of eq. (63). For the determination of $g_3$ the same calibration of the controls can be used as before with the only difference that the sign of the phase angle should be reversed.

An actual model of the instrument for use in the 3-cm wavelength band is shown in fig. 5.13. It is composed of nine directional junctions of the type shown in fig. 5.5, which number gives equal power division (cf. table 5.1).

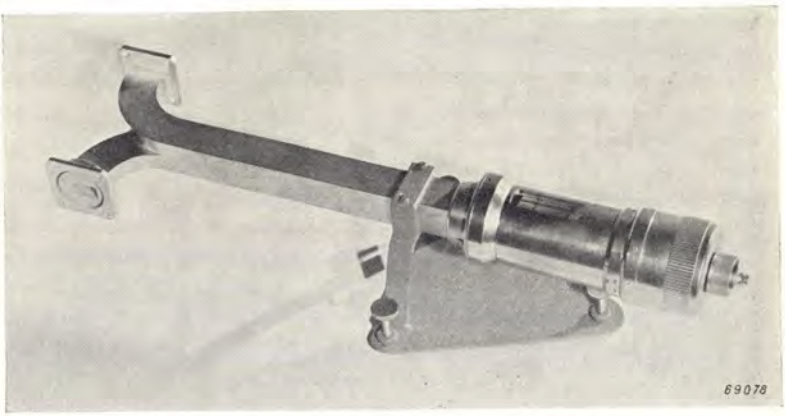



Fig. 5.13. Standard matching transformer.

The phase of the reflection coefficient is set by the large annular knob, while its modulus is controlled by the knob on the extreme right.

The application of the instrument as a standard termination and a standard matching transformer is illustrated by figs 5.14 and 5.15 respectively. In the first case the instrument is terminated by a matched load and connected with its input lead to one of the side arms of a double Tee. The circuit element of which the reflection coefficient is to be measured is connected to the other side arm. The indication for the proper setting of the controls of the standard termination is zero signal in the H-arm of the double Tee. This arm is connected to a measuring receiver. It is to be noted that it is not necessary here to use a *Magic* Tee as bridge element.

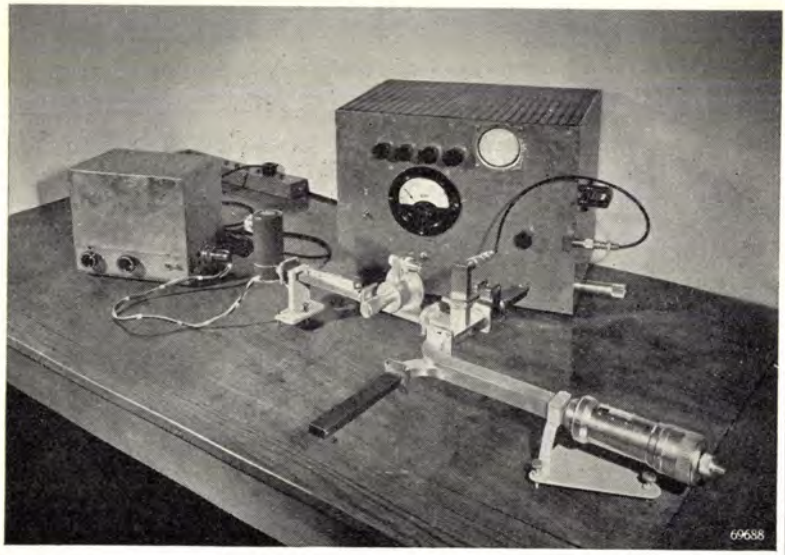



Fig. 5.14. Measurement of a reflection coefficient by comparison with the reflection of a standard termination.

In fig. 5.15 the circuit element to be measured forms the termination of the standard matching transformer. A directional coupler is used for the indication of the reflected wave, which must vanish for the proper setting.

The precision plunger drive with which the instrument is equipped enables us to measure accurately the directivity of the system of slots. To show this we shall derive the condition imposed on the plunger positions, i.e. $g_2$ and $g_4$, for total decoupling of the input lead (1) from the output lead (3). If in the equation

$$\boldsymbol{B} = \boldsymbol{S}\,\boldsymbol{A}, \tag{68}$$

where $S$ is given by eq. (1), the conditions

$$a_2 = g_2 b_2, \quad a_3 = 0, \quad b_3 = 0, \quad a_4 = g_4 b_4, \tag{69}$$

are inserted, it follows from the resulting secular equation that

$$\frac{g_2 + g_4}{1 + g_2 g_4} = \frac{\alpha}{\alpha^2 - \beta\delta} = \frac{s_1 + s_2}{1 + s_1 s_2}. \tag{70}$$

If directivity is perfect ($\alpha = 0$), eq. (70) requires $g_2 + g_4 = 0$, in agreement with eq. (61). We now assume that $\alpha$ is small compared with unity and that

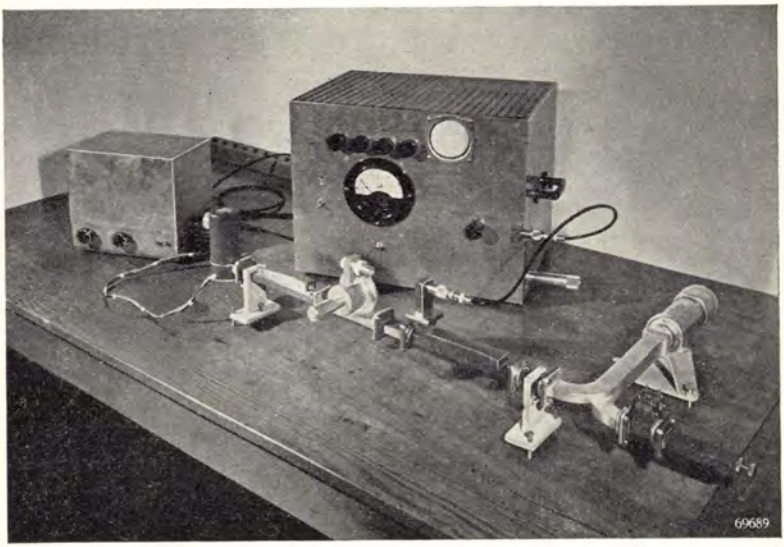



Fig. 5.15. Measurement of a reflection coefficient with the aid of a standard matching transformer.

$\beta$ and $\delta$ differ little from the ideal values given by eqs (60). The right-hand side of eq. (70) is real and to the first order equal to 2 $\alpha$. Expressing $g_2$ and $g_4$ in terms of $\chi$ and $\xi$ according to eqs (61) we obtain from eq. (70)

$$\cos 2\xi = 2\,\alpha \cos 2\chi\,. \tag{71}$$

In this equation $\chi$, controlled by the "phase knob" of the matching transformer, is to be regarded as the independent variable. The procedure for the determination of $\alpha$ then consists in taking, for several values of $\chi$, the reading of $\xi$ that decouples the output lead from the input. For small values of $\alpha$ a graph of $2\xi$ against $2\chi$ obtained in this way will show a

sinusoidal fluctuation of $2\xi$ about $\frac{1}{2}\pi$. The amplitude of this sine curve is equal to $\alpha$. This method for the measurement of $\alpha$ is closely related to the so-called S-curve method used for the determination of the characteristics of lossless transformers [2],[6],[25]. This is brought out even more closely by writing eq. (71) in the form

$$\tan(\chi-\xi) = \frac{1+2\alpha}{1-2\alpha}\operatorname{cotan}(\chi+\xi). \tag{73}$$

It can be shown by an analysis of the behaviour of the matching transformer for $\alpha$, $\beta$ and $\delta$ differing slightly from their ideal values that the possible error in the reading for the modulus of the reflection coefficient depends mainly on the directivity and not so much on the coupling. Deviation of the latter quantity from its ideal value causes, apart from phase errors, only a reduced accuracy for low values of $|\Gamma|$. The possible error due to finite directivity, on the other hand, attains for $|\Gamma| = \frac{1}{2}\sqrt{2}$ a maximum value which is equal to the directivity itself. As a directivity of 40 db has been realized in actual models, the corresponding absolute accuracy of $|\Gamma|$ is $\pm$ 0·01. Summarizing we may state that the standard matching transformer is best suited for measurements in the range $0{\cdot}1 < |\Gamma| < 1{\cdot}0$. For the accurate measurement of small reflections conventional apparatus such as a standing-wave detector or a squeeze section are to be preferred.

## CHAPTER VI. RESONANT SLOTS

In this chapter we shall extend the treatment of the type of junction discussed in the previous chapter so as to include resonant coupling. The literature on this subject [26] is scanty, whereas extensive theoretical and experimental material is available on both resonant slots connecting the interior of a waveguide with outside space [26], [27] and on resonant diaphragms across a section of rectangular waveguide [28].

### VI. 1. Single slots

In the preceding chapter it was shown that $s_2$ is unaffected by the presence of a longitudinal slot in the wide side common to two rectangular waveguides. The properties of the slot are completely specified by the value of $s_1$. A transverse slot, on the other hand, leaves $s_1$ undisturbed; its behaviour can be described fully by $s_2$.

A slot in the wall of a waveguide may be regarded as a magnetic dipole. If its length becomes of the order of half a wavelength it will accordingly display resonance effects.

We shall first treat the case of two sections of rectangular waveguide coupled by a single *longitudinal* slot in their common wide side, as shown in fig. 6.1. As $s_2$ is equal to its unperturbed value —1, the scattering coefficients are given by

$$\alpha = \gamma = \delta = \tfrac{1}{4}(s_1 - 1)\,, \qquad \beta = \tfrac{1}{4}(s_1 + 3)\,, \tag{1}$$

as follows from the general equations (V. 3).

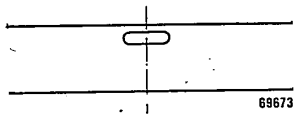



Fig. 6.1. Resonant longitudinal slot.

In Chap. II the expression for the reflection coefficient of a resonant termination has been derived. The quantity $s_1$ which describes the behaviour of a resonant longitudinal slot is essentially a reflection coefficient, so that eq. (II.36) applies. The off-resonance reflection is given in this equation as $i$. The unperturbed (i.e. off-resonance) value of $s_1$, referred to the symmetry plane of the junction, has been found to be +1, however. Hence, in order to represent $s_1$ correctly, equation (II.36) must be multiplied by —$i$, corresponding to an inward shift of the reference plane over a distance of one-eighth of a wavelength. Thus we have

$$s_1 = 1 - \frac{2Q_L}{Q_E(1 - 2i\,Q_L\,\Delta\omega/\omega_0)}\,. \tag{2}$$

To simplify the discussion we will, for the moment, assume the junction to be lossless. Then

$$Q_E = Q_L \equiv Q\,. \tag{3}$$

In general this assumption introduces considerable discrepancies between theory and experiment. For resonant slots, however, it is often a useful approximation because of the heavy loading due, in our case, to the four output leads of the junction. Substituting eq. (3) in eq. (2) and introducing the phase angle of $s_1$ we have

$$s_1 = e^{i\vartheta_1} = -\frac{1 + 2iQ\,\Delta\omega/\omega_0}{1 - 2iQ\,\Delta\omega/\omega_0}\,. \tag{4}$$

Hence

$$\tan\tfrac{1}{2}\vartheta_1 = -\frac{\omega_0}{2Q\Delta\omega}\,. \tag{5}$$

The introduction of $\vartheta_1$ in eqs (1) yields

$$|\alpha|^2 = |\gamma^2| = |\delta|^2 = \tfrac{1}{4}\sin^2\tfrac{1}{2}\vartheta_1, \qquad |\beta|^2 = 1 - \tfrac{3}{4}\sin^2\tfrac{1}{2}\vartheta_1 . \tag{6}$$

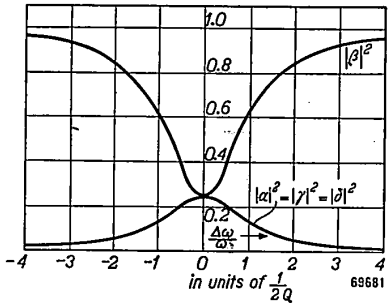



Fig. 6.2. Power-scattering coefficients of a single slot.

These power-scattering coefficients are plotted in fig. 6.2 against frequency according to eqs (5) and (6). At the resonance frequency we have by eqs (1) and (4)

$$\Delta\omega = 0, \quad s_1 = -1, \quad \alpha = \gamma = \delta = -\tfrac{1}{2}, \quad \beta = +\tfrac{1}{2} . \tag{7}$$

The resonance scattering by a longitudinal slot is illustrated in fig. 6.3.

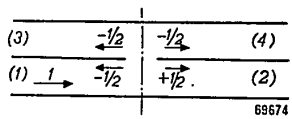



Fig. 6.3. Resonance scattering by a lossless longitudinal slot.

If the internal losses are accounted for by a finite value of $Q_U$, the scattering coefficients at the resonance frequency are found to be by eqs (1) and (2)

$$\Delta\omega = 0, \; s_1 = -1 + \frac{2Q_L}{Q_U}, \; \alpha = \gamma = \delta = -\tfrac{1}{2}\left(1 - \frac{Q_L}{Q_U}\right), \; \beta = \tfrac{1}{2}\left(1 + \frac{Q_L}{Q_U}\right) . \tag{8}$$

A *transverse* slot, as shown in fig. 6.4, does not affect $s_1$. Substitution of the unperturbed value $s_1 = +1$ in eqs (V.3) yields

$$\alpha = \gamma = -\delta = \tfrac{1}{4}(1 + s_2), \quad \beta = \tfrac{1}{4}(3 - s_2) . \tag{9}$$

In the vicinity of the resonance frequency of the slot, $s_2$ is given by the

expression for the reflection coefficient of a resonant termination, eq. (II.36). Because the off-resonance value of $s_2$ equals $-1$, the expression must be multiplied by $i$, corresponding to an inward shift of the reference plane over a distance of three-eighth of a wavelength. Thus

$$s_2 = -1 + \frac{2\,Q_L}{Q_E(1-2i\,Q_L\,\Delta\omega/\omega_0)}. \tag{10}$$

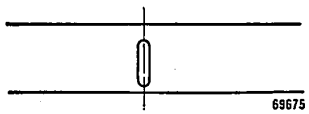



Fig. 6.4. Resonant transverse slot.

Assuming again for the moment the junction to be lossless we have

$$s_2 = -e^{i\vartheta_2} = \frac{1+2i\,Q\,\Delta\omega/\omega_0}{1-2i\,Q\,\Delta\omega/\omega_0}, \tag{11}$$

whence

$$\tan\tfrac{1}{2}\vartheta_2 = -\frac{\omega_0}{2Q\,\Delta\omega}. \tag{12}$$

The frequency dependence of the phase angles $\vartheta_1$ and $\vartheta_2$ of a longitudinal and a transverse slot respectively is given by the same function as is evident from eqs (5) and (12). The dependence of the absolute value of the scattering coefficients on the phase angle is also the same in both cases, so that fig. 6.2 applies to a transverse slot as well. The phase of the waves scattered by a transverse slot differs, however, from the phase of the waves scattered by a longitudinal slot. At resonance we have for a transverse slot

$$\Delta\omega = 0, \quad s_2 = +1, \quad \alpha = \beta = \gamma = +\tfrac{1}{2}, \quad \delta = -\tfrac{1}{2}, \tag{13}$$

which situation is illustrated in fig. 6.5.

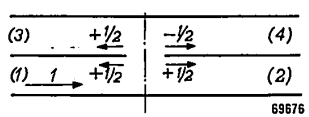



Fig. 6.5. Resonance scattering by a lossless transverse slot.

From eqs (7) and (13) it follows that in lossless junctions containing either a longitudinal or a transverse slot resonance is characterized by the scattering of one quarter of the incident power into each of the four output leads.

If the dissipation within the junction is included in the calculation, the scattering coefficients at resonance are given by

$$\Delta\omega = 0,\ s_2 = 1 - \frac{2Q_L}{Q_U},\ \delta = \gamma = -\delta = \tfrac{1}{2}(1 - \frac{Q_L}{Q_U}),\ \beta = \tfrac{1}{2}(1 + \frac{Q_L}{Q_U}). \quad (14)$$

The strength of the coupling between the slot and the fields in the waveguides, as expressed quantitatively by $Q_E$, is determined mainly by the position of the slot relative to the axes of the waveguides. Clearly the lowest values of $Q_E$ can be expected with slots cut in a position where the component of the undisturbed magnetic field parallel to the slot attains its maximum value. From fig. 4.1 it can thus be seen that for longitudinal slots in the wide side the minimum value of $Q_E$ is achieved by slots adjacent to the narrow side. Transverse slots, on the other hand, display low values of $Q_E$ if the slot is located in a central position. These effects are shown schematically in fig. 6.6.

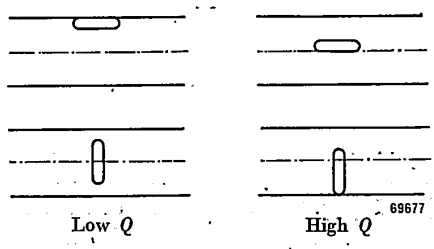



Fig. 6.6. Dependence of $Q_E$ on slot position.

The dependence of the quality factor on slot width is only slight. Both the coupling of the slot to the fields in the waveguides and the energy-storing capacity of the slot increase with slot width. The net effect is a slow decrease of $Q_E$ with increasing slot width.

Variation of the wall thickness affects only the energy stored in the slot, so that $Q_E$ increases with the wall thickness. It should be realized, however, that the theory holds strictly only for an infinitely thin wall between the two sections of waveguide.

The general relation between the resonant frequency of a slot and its size, position and the dimensions of the waveguides is not easily accessible to theoretical treatment. This is even more so for slots filled with a dielectric material which differs from the dielectric within the waveguides. For the time being one can do little else but collect enough experimental material in order to be able to predict the slot dimensions needed in a specific case.

## VI. 2. Measurements on single slots

The three quantities specifying the behaviour of a single slot, either longitudinal or transverse, are $\omega_0$, $Q_L$ and $Q_U$. They may be determined experimentally by measuring the absolute value of $\alpha$, $\beta$, $\gamma$ or $\delta$ as a function of frequency. If $Q_L/Q_U$ is small compared with unity, however, the determination of $Q_U$ by this method requires very high accuracy in the measurements, as follows from eqs (8) and (14). We have therefore developed an alternative procedure, which yields $Q_L$ and $Q_U$ with about the same accuracy.

The waveguide system used for the measurement is shown schematically in fig. 6.7. A generator $G$ is connected to lead (1) of the junction and a receiver $R$ to output lead (2). The frequency of the generator is variable;

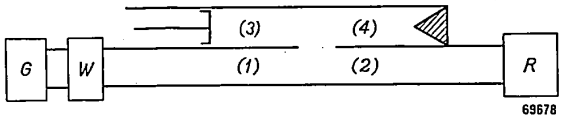



Fig. 6.7. Outline of the waveguide system used for the measurement of slot constants. $G$ = generator, $W$ = wavemeter, $R$ = matched receiver.

its value can be determined with the aid of the wavemeter $W$. In output lead (3) a movable plunger has been inserted, whilst lead (4) is terminated by a matched load. It is evident that the value of the transmission coefficient $b_2/a_1$ between lead (1) and lead (2) of this system depends on the position of the plunger in lead (3). By moving the plunger we can find a maximum and a minimum of the absolute value of this transmission coefficient. The measurement consists in the determination of the ratio of these two extreme values of $|b_2/a_1|$.

We shall now derive the equations needed for the interpretation of the measurements. The conditions to be inserted in the four equations contained in

$$\boldsymbol{B} = \boldsymbol{S}\boldsymbol{A}, \tag{15}$$

are given by

$$a_2 = 0, \qquad a_3 = g_3 b_3, \qquad a_4 = 0. \tag{16}$$

Here $g_3$ represents the reflection coefficient of the plunger ($|g_3| = 1$). Substituting eq. (16) in eq. (15) and solving for $b_2$ we obtain

$$\frac{b_2}{a_1} = 1 + \frac{\beta - 1}{1 - \alpha g_3}. \tag{17}$$

In the derivation of eq. (17) nothing special has been assumed about the

junction except its symmetry. The transmission coefficient $b_2/a_1$ is related to $g_3$ by a bilinear transformation. Because $g_3$, as a function of the plunger position, describes a circle (in fact the unit circle) in the complex plane, the locus of $b_2/a_1$ is a circle also. The centre of the latter is located at the point $(\beta - |\alpha|^2)/(1 - |\alpha|^2)$; its radius equals $|\alpha||\beta - 1|/(1 - |\alpha|^2)$. Hence the extreme values of $|b_2/a_1|$ are given by

$$\left|\frac{b_2}{a_1}\right|_{\max} = \left|\frac{\beta - |\alpha|^2}{1 - |\alpha|^2}\right| + \frac{|\alpha|\,|\beta - 1|}{1 - |\alpha|^2}, \tag{18}$$

$$\left|\frac{b_2}{a_1}\right|_{\min} = \left|\frac{\beta - |\alpha|^2}{1 - |\alpha|^2}\right| - \frac{|\alpha|\,|\beta - 1|}{1 - |\alpha|^2}. \tag{19}$$

The ratio $H$ between these quantities is thus found to be

$$\left(\frac{|b_2|_{\max}}{|b_2|_{\min}}\right)_{a_1=\text{const}} = \frac{|\beta - |\alpha|^2| + |\alpha|\,|\beta - 1|}{|\beta - |\alpha|^2| - |\alpha|\,|\beta - 1|} \equiv H\,. \tag{20}$$

We now suppose that the only discontinuity in the common wide side between the two sections of waveguide is a single slot. From eqs (1), (2), (9) and (10) it follows that the expressions for $|\alpha|$ and $\beta$ as functions of frequency and the slot constants are the same for both longitudinal and transverse slots, viz.

$$|\alpha| = \frac{Q_L}{2Q_E|1 - 2i\,Q_L\,\Delta\omega/\omega_0|}, \quad \beta = 1 - \frac{Q_L}{2Q_E(1 - 2iQ_L\,\Delta\omega/\omega_0)}. \tag{21}$$

There is no need to derive explicitly the rather intricate expression for $H$ as a function of frequency. It is sufficient to calculate $H$ for two special values of $\Delta\omega$. At resonance $H$ attains a maximum value, which is according to eqs (20) and (21) given by

$$\Delta\omega = 0\,, \quad H_{\text{res}} = \frac{1 + \dfrac{Q_L}{Q_U}}{\dfrac{Q_L}{Q_U}\left(3 - \dfrac{Q_L}{Q_U}\right)}. \tag{22}$$

As the other special value of $\Delta\omega$ we take the half-power points defined by

$$2\,Q_L\,\frac{\Delta\omega}{\omega_0} = \pm 1\,. \tag{23}$$

For the corresponding expression $H_{1/2}$ we find from eqs (20), (21) and (23)

$$H_{1/2} = \frac{\sqrt{\left(5-\frac{Q_L}{Q_U}\right)^2\left(1+\frac{Q_L}{Q_U}\right)^2+4\left(1-\frac{Q_L}{Q_U}\right)^2}+\left(1-\frac{Q_L}{Q_U}\right)^2}{\sqrt{\left(5-\frac{Q_L}{Q_U}\right)^2\left(1+\frac{Q_L}{Q_U}\right)^2+4\left(1-\frac{Q_L}{Q_U}\right)^2}-\left(1-\frac{Q_L}{Q_U}\right)^2}. \quad (24)$$

In figs 6.8 and 6.9 $H_{res}$ and $H_{1/2}$ respectively are plotted as a function of $Q_L/Q_U$. In fig. 6.8 a graph is given also of the reciprocal of $|b_2/a_1|_{min}$ at the resonance frequency. The corresponding equation is obtained from eqs (19) and (21). Thus

$$\left|\frac{b_2}{a_1}\right|^{-1}_{min} = \tfrac{1}{2}\left(1+\frac{Q_U}{Q_L}\right), \quad \Delta\omega = 0. \quad (25)$$

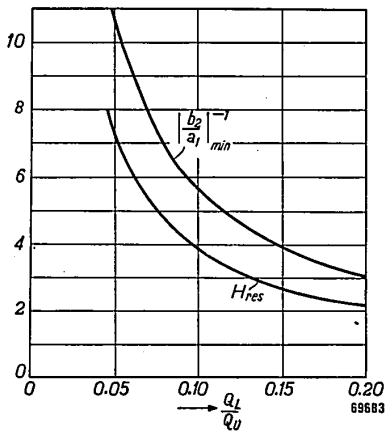



Fig. 6.8. The ratio $H$ and the maximum insertion loss at the resonance frequency as a function of $Q_L/Q_U$.

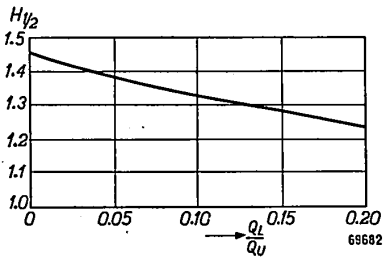



Fig. 6.9. The ratio $H$ at half-power points as a function of $Q_L/Q_U$

The measurement of this quantity, taken at the resonance frequency with a plunger setting so as to give maximum insertion loss, can be used conveniently as a check on the accuracy of the value of $Q_L/Q_U$ calculated from the measured value of $H_{res}$.

A typical example of an experimentally determined plot of $H$ against frequency is shown in fig. 6.10; the isolated point represents the maximum insertion loss. The measurements were taken on a junction having a single transverse glass-filled slot. The resonance frequency is seen to be 9070 Mc/s. The value of $Q_L/Q_U$ corresponding to $H_{res} = 3{\cdot}45$ is read from fig. 6.8 to be 0·110. The value derived from a maximum insertion loss of 5·2

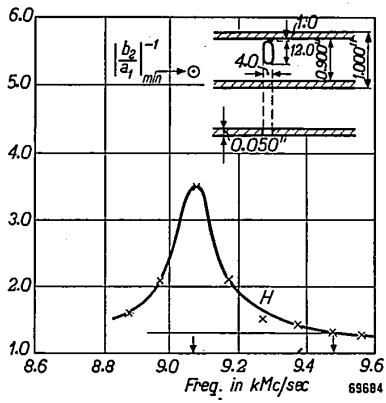



Fig. 6.10. Experimental plot for a transverse slot of $H$ versus frequency. Slot dimensions in millimeters. The isolated point represents the maximum insertion loss at the resonance frequency.

is 0·106. Putting $Q_L/Q_U = 0{\cdot}108$ we derive from fig. 6.9 that $H_{1/4}$ equals 1·32. Returning now to fig. 6.10 we deduce $Q_L = 11{\cdot}1$ and hence finally $Q_U = 103$.

A check on the symmetry of the junction may readily be achieved by an interchange of the positions of the plunger and the matched load in leads (3) and (4). The transmission coefficient $b_2/a_1$ becomes under these conditions

$$\frac{b_2}{a_1} = 1 + \frac{\beta - 1}{1 - \alpha g_4}, \tag{26}$$

where $g_4$ denotes the reflection coefficient of the plunger now in lead (4). As the dependence of $b_2/a_1$ on $g_4$ is identical with its dependence on $g_3$, eq. (17), the measuring results should be the same in both cases.

## VI. 3. Resonant directional couplers

In Sec. V. 3 it has been proved that a lossless junction of rectangular waveguides having both a longitudinal and a transverse slot in the common wide side, as shown in fig. 5.4, constitutes a directional coupler. For perfect directivity the slot dimensions must be so as to make

$$s_1 + s_2 = 0. \tag{27}$$

This equation is satisfied if the phase angles $\vartheta_1$ and $\vartheta_2$ are equal. In the case of a lossless junction with resonant slots the phase angles are determined by eqs (5) and (12). Hence perfect directivity requires that

$$\left(\frac{\omega_0}{2Q\Delta\omega}\right)_1 = \left(\frac{\omega_0}{2Q\Delta\omega}\right)_2. \tag{28}$$

Subscripts 1 and 2 are used to distinguish between quantities referring to the longitudinal and the transverse slot respectively. It follows from eq. (28) that perfect directivity is maintained irrespective of frequency if both the quality factor and the resonant frequency are the same for the two slots. Let us assume this to be the case. Then we have (cf. eq. (V. 30))

$$\alpha = \gamma = 0, \quad |\beta|^2 = \cos^2 \tfrac{1}{2}\vartheta, \quad |\delta|^2 = \sin^2 \tfrac{1}{2}\vartheta, \tag{29}$$

where $\vartheta$ denotes, as before, the common value of $\vartheta_1$ and $\vartheta_2$. The appearance of the graphs of $|\beta|^2$ and $|\delta|^2$ as a function of frequency is very much like the curves in fig. 6.2. The vertical scale must be taken differently, however, for $|\beta|^2$ is zero and $|\delta|^2$ equals 1 at the resonance frequency. This means that complete power transfer takes place in a resonant directional coupler at the resonance frequency. The power transfer, as a function of frequency, drops on both sides of the resonance frequency. A resonant directional coupler thus constitutes a branching filter.

An interesting application of the structure just described is its use as a radar duplexer. Let us consider the system shown in fig. 6.11. Let a signal of frequency equal to the resonance frequency of the directional coupler

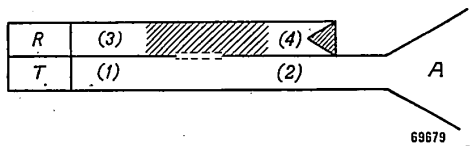



Fig. 6.11. Application of a resonant directional coupler as a radar duplexer.

enter the aerial $A$. It will travel down lead (2) towards the junction. There it is transferred completely to lead (3), thus reaching the receiver $R$ irrespective of the terminations of leads (1) and (4).

A high-power pulse leaving the transmitter $T$ will, in passing the junction, build up a high voltage across the slots. If now the upper section of waveguide contains a suitable gas filling, which is confined to that space by gas-tight dielectric layers in the slots, the high voltage will fire the gas. The ionized gas short-circuits the slots, thereby impeding the passage of the pulse towards lead (4). Hence the pulse leaves the system by lead (2) and the aerial. After the pulse the gas deionizes, so that the path from the aerial towards the receiver is open again for reflected pulses.

The main advantage of the system just described over duplexers of conventional design lies in the way it handles the spike. By the spike is meant that part of the transmitter pulse which passes through the gas-filled space before ionization starts. It can be seen from fig. 6.11 that the spike travels down lead (4) and is absorbed harmlessly by the matched load, whereas in conventional systems the spike reaches the receiver with the inherent danger of damage to it.

In actual models the slots are fitted with glass windows in order to ensure a gas-tight partition between the two sections of waveguide. For an accurate description of the behaviour of the junction it is then no longer justified to neglect dissipation.

In general the condition for a matched junction, $\alpha = 0$, ensures perfect directivity only if the junction is lossless. But in our case, because symmetry guarantees $\alpha = \gamma$, the same condition is still sufficient if dissipation takes place. If, accordingly, the general expressions for $s_1$ and $s_2$, eqs (2) and (10), are inserted in the directivity condition, eq. (27), it follows that for broadband directivity all three relevant quantities, viz. $\omega_0$, $Q_L$ and $Q_U$, must have the same values for the two slots. Let us assume this to be the case. Then we have

$$\alpha = \gamma = 0, \quad \beta = 1 - \frac{Q_L}{Q_E\,(1 - 2i\,Q_L\,\Delta\omega/\omega_0)}, \quad \delta = -\frac{Q_L}{Q_E\,(1 - 2i\,Q_L\,\Delta\omega/\omega_0)}. \tag{30}$$

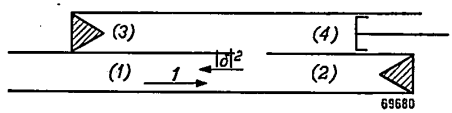



Fig. 6.12. Measurements of the insertion loss.

The quantity specifying the behaviour of the "cold" duplexer, i.e. as far as received signals are concerned, is the insertion loss. This can be measured in a very direct way as shown in fig. 6.12. It is evident, that with leads (2) and (3) terminated by matched loads the absolute value of the reflection coefficient in lead (1) caused by a plunger in lead (4) equals

$|\delta|^2$. But this is just the fraction of the power entering the aerial (fig. 6.11) that reaches the receiver. In an analogous way we can measure the other power-scattering coefficient, $|\beta|^2$, by interchanging the terminations of leads (2) and (4). In the ideal case of slots with identical properties the expressions for $|\beta|^2$ and $|\delta|^2$ as a function of frequency are according to eq. (30) given by

$$|\beta|^2 = \frac{\left(\frac{Q_L}{Q_U}\right)^2 + \left(2\,Q_L \frac{\Delta\omega}{\omega_0}\right)^2}{1 + \left(2\,Q_L \frac{\Delta\omega}{\omega_0}\right)^2}, \qquad |\delta|^2 = \frac{Q_L^2}{Q_E^2\left\{1 + \left(2\,Q_L \frac{\Delta\omega}{\omega_0}\right)^2\right\}}. \tag{31}$$

In fig. 6.13 measured values of $|\beta|^2$ and $|\delta|^2$ are shown for a junction with two resonant slots. The dimensions of the transverse slot are the same as of the single slot shown in fig. 6.10. By measurements on single longitudinal

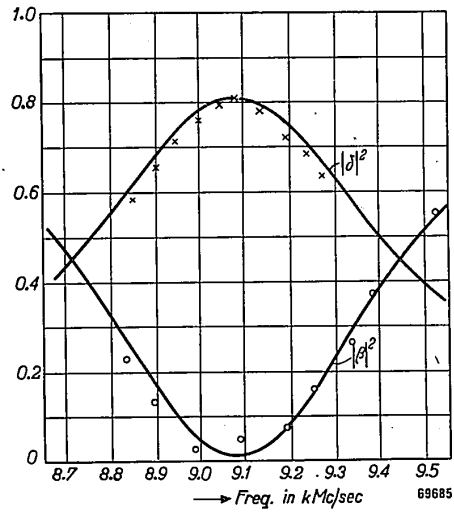



Fig. 6.13. Power-scattering coefficients of a radar duplexer versus frequency. The crosses and circles are experimental, the curves are theoretical.

slots the size was determined of the slot with as nearly as possible the same value of the slot constants as the transverse slot. The curves in fig. 6.13 have been calculated from eqs (31) after insertion of the constants of the transverse slot, as recorded in the preceding section.

**Acknowledgement**

The author wishes to express his thanks to Prof. R. Kronig for many stimulating discussions. Also he is grateful to Prof. H. B. G. Casimir for his interest in the subject.

*Eindhoven, January 1952*

REFERENCES

1) C. G. Montgomery, R. H. Dicke and E. M. Purcell, Principles of microwave circuits, McGraw-Hill, New York, 1948, Chap. 12.
2) J. C. Slater, Rev. mod. Phys. **18**, 441-512, 1946.
3) S. Tomonaga, J. phys. Soc. Japan **2**, 158-171, 1947; **3**, 93-104, 1948.
4) D. M. Kerns, J. Res. Bur. Stand. **42**, 515-540, 1949.
5) S. A. Schelkunoff, Electromagnetic waves, van Nostrand, New York, 1943, p. 35
6) G. B. Collins, Microwave magnetrons, McGraw-Hill, New York, 1948, Chap. 5.
7) W. Altar, Proc. Inst. Radio Engrs **35**, 355-361; 478-484, 1947.
8) D. M. Kerns, J. Res. Bur. Stand. **46**, 267-278, 1951.
9) J. Rosenthal and G. M. Murphy, Rev. mod. Phys. **8**, 317-346, 1936.
10) E. Bauer, Introduction à la théorie des groupes et à des applications à la physique quantique, Presses universitaires, Paris, 1933.
11) H. Eyring, J. Walter and G. E. Kimball, Quantum chemistry, Wiley & Sons, New York, 1944.
12) J. A. Schouten, Tensor analysis for physicists, Clarendon Press, Oxford, 1951.
13) H. A. Bethe, Phys. Rev. **66**, 163-182, 1944.
14) H. A. Bethe, Radiation Laboratory Report 43-22.
15) H. A. Bethe, Radiation Laboratory Report 43-27.
16) H. J. Riblet, Proc. Inst. Radio Engrs **35**, 1307-1313, 1947.
17) J. Hessel, G. Goubau and L. R. Battersby, Proc. Inst. Radio Engrs **37**, 990-1000, 1949.
18) C. G. Montgomery, R. H. Dicke and E. M. Purcell, Principles of microwave circuits, McGraw-Hill, New York, 1948, Chap. 9.
19) M. Surdin, J. Inst. elect. Engrs **93**, III A, 725-736, 1946.
20) C. G. Montgomery, Technique of microwave measurements, McGraw-Hill, New York, 1947, Chap. 14.
21) P. E. Krasnushkin and R. B. Khokhlov, J. Tech. Phys. (Russian) **19**, 931-942, 1949.
22) H. J. Riblet and T. S. Saad, Proc. Inst. Radio Engrs **36**, 61-64, 1948.
23) S. Rosen and J. T. Bangert, Proc. Inst. Radio Engrs **27**, 393-401, 1949.
24) H. G. Beljers and W. J. van de Lindt, Philips Res. Rep. **6**, 96-104, 1951.
25) N. Marcuvitz, Waveguide handbook, McGraw-Hill, New York, 1951, p. 132 et seq.
26) W. H. Watson, The physical principles of waveguide transmission and antenna systems, Clarendon Press, Oxford, 1947.
27) A. F. Stevenson, J. appl. Phys. **19**, 24-38, 1948.
28) L. D. Smullin and C. G. Montgomery, Microwave duplexers, McGraw-Hill, New York, 1948, Sec. 3.10.

*Philips Res. Rep.* **7**, *303-313, 1952*

# SOME REMARKS CONCERNING CLOSE-PACKING OF EQUAL SPHERES

by A. H. BOERDIJK

513.468

**Summary**

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 maximal 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 falsc.

**Resumé**

Définition de trois critères pour estimer la densité moyenne dans des régions locales des configurations des sphères identiques. Quelques configurations sont décrites ayant une densité locale dans certain régions dépassant celle de l'empillage plus dense des sphères identiques. Ces régions peuvent avoir une volume infinie. Ensuite on prouve que le nombre maximum des sphèrcs tangentes à une sphère est douze. Une supposition de Fejes concernant une quatorzième sphère adjointe à cette configuration paraît incorrecte.

**Zusammenfassung**

Drei Kriterien werden aufgestellt, die dazu dienen die Dichte von Kugelpackungen (Kugeln gleichen Durchmessers) in endlich großen Bereichen miteinander zu vergleichen. Einige Anordnungen werden angegeben, die gemäß dieser Kriterien in gewissen Bereichen eine größere Dichte aufweisen als die bekannte dichteste gitterförmige Kugelpackung. Diese Bereiche können selbst einen unendlich großen Rauminhalt haben. Weiter wird bewiesen, daß die größte Anzahl der Kugeln, die gleichzeitig eine Kugel berühren können, zwölf ist. Es wird bewiesen, daß eine Vermutung von Fejes über eine vierzehnte Kugel, zu diesen dreizehn Kugeln hinzugefügt, falsch ist.

## 1. Criteria for estimating the mean density in local regions

Let us consider the configuration of equal spheres (of diameter 1) known as "close-packing" or "normal piling". If the density inside the spheres is unity, and zero outside the spheres, the mean density $D$ of the configuration will be

$$D = (\pi/6)\ \sqrt{2} \approx 0{\cdot}7405\,. \qquad (1)$$

Is this the maximal mean density obtainable in any possible configuration of equal spheres, or may this density be surpassed, either in all space or locally? It may well be surpassed locally. For in close-packing we can extend the configuration uniformly in any direction over any distance and we shall see that there are denser local configurations that cannot be so extended.

For the investigation of this problem we shall need means for estimating the densities of configurations in local regions. We shall give three criteria.

1. One measure for the local mean density is the number $N$ of centres

of spheres of the configuration located inside or on a sphere with radius $R$ and centre $M$.

II. Another measure for the density is the mean density $d$ in such a sphere.

These criteria are correlated inter alia in the following way. If we have two configurations $C$ and $C'$ for which

$$N \geqslant N', \quad (0 \leqslant R < R_1)$$

and

$$N > N', (R_1 \leqslant R < R_2),$$

then it follows easily that

$$d > d', (R_1 - \tfrac{1}{2} < R < R_2 - \tfrac{1}{2}). \tag{2}$$

The third criterion is based on the following argument. In close-packing the planes through the centres of adjacent spheres form a quasi-regular honeycomb. Each centre is surrounded by six regular octahedra and eight regular tetrahedra, both (in our case) having an edge-length 1. Calculating the mean densities $d_4$ and $d_8$ for the space enclosed by a tetrahedron and an octahedron we find respectively

$$d_4 = (6\varphi - \pi)\sqrt{2} \approx 0{\cdot}7797, \tag{3}$$

$$d_8 = (\tfrac{3}{4}\pi - 3\varphi)\sqrt{2} \approx 0{\cdot}7209, \tag{4}$$

where

$$\varphi = \arcsin \tfrac{1}{3}\sqrt{3} \approx 35°16', \tag{5}$$

$2\varphi$ being the dihedral angle of a regular tetrahedron.

It is rather surprising to find that $d_4$ and $d_8$ differ by 8%. On the other hand it is a known fact from crystallography (see also [1])) that the diameter of the largest possible sphere inside such an octahedron (which sphere touches 6 spheres of the configuration) is almost twice that for the largest sphere inside a tetrahedron (touching 4 spheres of the configuration).

In close-packing the number of tetrahedra turns out to be twice the number of octahedra, whereas the volume of each tetrahedron is one fourth of that of each octahedron [2]). As a check we have indeed from (1,3,4)

$$D = \tfrac{1}{3}(d_4 + 2d_8). \tag{6}$$

The proportion of the numbers of tetrahedra and octahedra in a configuration is a good measure for the density. At least this should be so in the inside of the configuration, for there all this space is made up of tetrahedra and octahedra. This measure will be accurate, even if the tetrahedra and octahedra are not regular, but do not deviate "too much" from regularity as turns out to be the case in all configurations to be considered later. For practical reasons we divide the octahedron into two four-sided pyramids. This leads to our third criterion.

III. If the numbers of three- and four-sided pyramids in a configuration

are $t$ and $f$ respectively, $t/f$ can be regarded as a measure for the local mean density of that configuration.

For close-packing we have $t/f = 1$.

## 2. The density of close-packing and locally denser configurations

We will now consider six different configurations 1, 2, ..., 6, all with a central sphere. For $M$ we choose the centre of that sphere. We build up the configurations by addition of other spheres, in the order of increasing distances from $M$. All $n$ spheres of which the centres have the same distance from $M$ form together a "shell". So we have

$$N = \sum_0^R n. \tag{7}$$

The configurations 1, 2 and 3 occur in close-packing. Configuration 1 is conformable hexagonal close-packing, 3 is conformable face-centred cubic close-packing and 2 is a mixture of the two types. In other words, in 1 the succession of planes of largest density is of the type *a-b-a-b-a*, in 2 of type *b-c-a-b-a*, in 3 of type *b-c-a-b-c* ($M$ being in the middle plane). In the region $R \leqslant \sqrt{3}$ which will be considered here, these three turn out to be the only possible configurations with a central sphere occurring in close-packing.

Configuration 4 consists of 12 spheres touching the central sphere, with their centres on the vertices of a cuboctahedron (as in configurations 2 and 3). Six spheres are added with their centres on the symmetry axes of the square faces of the cuboctahedron and 8 spheres are added in the same way to the triangular faces.

Configuration 5 consists of 12 spheres, touching the central sphere, with their centres on the vertices of a regular icosahedron. These 12 spheres do not touch each other [3], [4]) as the edge-length of the icosahedron will be

$$l_{20} = \frac{4}{\sqrt{10 + 2\sqrt{5}}} \approx 1{\cdot}05. \tag{8}$$

Twenty spheres are added with their centres on the symmetry axes of the faces of the icosahedron. For this shell $R$ will be

$$a = \tfrac{2}{3}\,\frac{3\sqrt{3} + \sqrt{15}}{\sqrt{10 + 2\sqrt{5}}} \approx 1{\cdot}59. \tag{9}$$

In configuration 6 we have one sphere touching the central sphere in its "north pole". Five spheres are added one by one, all touching the north-pole sphere, the second touching the first one, the third one touching the second one, and so on. Between the first one and the last one of these 5 spheres there will be a gap, as the dihedral angle $2\varphi$ of a regular tetrahedron

TABLE I

| Shell nr | $R$ | Num. value | | Configuration 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0·00 | $n,N$ | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1·00 | $n$ | 12 | 12 | 12 | 12 | 12 | 12 |
| | | | $N$ | 13 | 13 | 13 | 13 | 13 | 13 |
| | | | $t$ | 8 | 8 | 8 | 8 | 20 | 10 |
| | | | $f$ | 6 | 6 | 6 | 6 | 0 | 5 |
| | | | $t/f$ | 1·33 | 1·33 | 1·33 | 1·33 | $\infty$ | 2·00 |
| 2 | $b$ | 1·35 | $n$ | 0 | 0 | 0 | 0 | 0 | 1 |
| | | | $N$ | 13 | 13 | 13 | 13 | 13 | 14 |
| | | | $t$ | 8 | 8 | 8 | 8 | 20 | 10 |
| | | | $f$ | 6 | 6 | 6 | 6 | 0 | 6 |
| | | | $t/f$ | 1·33 | 1·33 | 1·33 | 1·33 | $\infty$ | 1·67 |
| 3 | $\sqrt{2}$ | 1·41 | $n$ | 6 | 6 | 6 | 6 | 0 | 4 |
| | | | $N$ | 19 | 19 | 19 | 19 | 13 | 18 |
| | | | $t$ | 8 | 8 | 8 | 8 | 20 | 10 |
| | | | $f$ | 12 | 12 | 12 | 12 | 0 | 10 |
| | | | $t/f$ | 0·67 | 0·67 | 0·67 | 0·67 | $\infty$ | 1·00 |
| 4 | $a$ | 1·59 | $n$ | 0 | 0 | 0 | 0 | 20 | 0 |
| | | | $N$ | 19 | 19 | 19 | 19 | 33 | 18 |
| | | | $t$ | 8 | 8 | 8 | 8 | 40 | 10 |
| | | | $f$ | 12 | 12 | 12 | 12 | 0 | 10 |
| | | | $t/f$ | 0·67 | 0·67 | 0·67 | 0·67 | $\infty$ | 1·00 |
| 5 | $\frac{2}{3}\sqrt{6}$ | 1·63 | $n$ | 2 | 1 | 0 | 8 | 0 | 10 |
| | | | $N$ | 21 | 20 | 19 | 27 | 33 | 28 |
| | | | $t$ | 10 | 9 | 8 | 16 | 40 | 20 |
| | | | $f$ | 12 | 12 | 12 | 12 | 0 | 10 |
| | | | $t/f$ | 0·83 | 0·75 | 0·67 | 1·33 | $\infty$ | 2·00 |
| 6 | $\sqrt{3}$ | 1·73 | $n$ | 18 | 21 | 24 | 0 | 0 | 0 |
| | | | $N$ | 39 | 41 | 43 | 27 | 33 | 28 |
| | | | $t$ | 28 | 30 | 32 | 16 | 40 | 20 |
| | | | $f$ | 24 | 26 | 28 | 12 | 0 | 10 |
| | | | $t/f$ | 1·17 | 1·15 | 1·14 | 1·33 | $\infty$ | 2·00 |

Table of the values of $n$, $N$, $t$ and $f$ as functions of $R$ (from 0 to $\sqrt{3}$) for the configurations 1, 2, ..., 6 described in the text.
$R$ = distance of the centres of the spheres from the centre $M$ of the central sphere
$n$ = number of spheres with the same value of $R$
$N$ = total number of spheres of which the centres are at a distance not larger than $R$ from $M$
$t$ = number of three-sided pyramids contained in the configuration
$f$ = number of four-sided pyramids contained in the configuration $a$ and $b$ are defined by eq. (9) and (10) respectively.

(equation (5)) is somewhat smaller than $2\pi/5$. Six more spheres are added causing the configuration to be symmetrical with respect to the equatorial plane of the central sphere. The centres of the 12 outer spheres are arranged in the vertices of an irregular polyhedron with 15 faces, of which 10 are triangular, 4 square and 1 rectangular. Fifteen more spheres are added with their centres on or close to the symmetry axes of these faces. Five of these spheres have their centres in the equatorial plane of the central sphere. The centre of the one facing the gap mentioned above will have a distance $R$ from $M$ equal to

$$b = \sqrt{3} \cos(\pi - 4\varphi) \approx 1{\cdot}347. \tag{10}$$

For all other shells the values of $R$ and $n$ can be found easily, so that we will not give them separately. Table I contains for the 6 configurations $n$, $N$, $t$, $f$ and $t/f$ for $R$ from 0 to $\sqrt{3}$.

We shall indicate a part of each of these configurations by the number of the configuration followed by the number of the outer shell. So 3-1 will be a sphere touched by 12 others, with their centres in the vertices of a cuboctahedron.

From the table it follows that the configurations 4-5, 5-4, 5-5, 6-2 and 6-5 have a larger local mean density than configurations of the same type occurring in close-packing, according to I and III; moreover the configurations 4-5 and 6-2 have for a certain region of $R$ a larger density according to our criterion II, as follows from (2). It is remarkable that in the region $0 \leqslant R < \sqrt{3}$ the local mean density according to all criteria of configuration 4 never falls below that of close-packing.

It turns out to be impossible to extend any of the configurations 4, 5 and 6 over a larger part of space, without causing the density to decrease locally for some shells. This effect is so strong that it seems unlikely that, if extended, the density of such configurations will be larger than that of close-packing.

Comparison of configurations with other types of centre, e.g., with a regular tetrahedron as a centre, leads to similar conclusions. In these cases too the density of close-packing can be surpassed locally.

## 3. Configurations denser than close-packing in regions with infinite volume

In the configuration of face-centered cubic close-packing the common element of two adjacent tetrahedra can be a vertex, or at the most an edge, but never a face. In hexagonal close-packing this element may be a face. However, two such tetrahedra having a face in common never have other faces in common with other tetrahedra. Of course it is possible to build configurations of equal spheres, of which the centres form the ver-

tices of tetrahedra in such a way that we could visit each tetrahedron by perforating common faces of tetrahedra only. Such configurations are of interest, since in a connected part of space with a volume exceeding that of two tetrahedra they will have a local mean density exceeding that of close-packing according to criterion III.

As a typical example we take four spheres touching each other. Their centres are the vertices of a regular tetrahedron. We add four spheres with their centres on the symmetry axes of the faces. So we obtain five tetrahedra, forming a polyhedron with $4 \times 3 = 12$ faces. Six more spheres can be added with their centres on the symmetry axes of 6 of these 12 faces. Now $t = 17, f = 0$. Of course such a configuration will not occur in close-packing. Alternatively, it can be regarded as built up from "five-rings", consisting of five spheres of which the centres are arranged on a circle of radius of $\frac{1}{2}\sqrt{3} \approx 0.865$. These five-rings play here the role of the configurations of six spheres of which the centres are arranged on a circle of radius 1, in the planes of maximal density in close-packing ("six-rings"). Here our configuration contains 6 of such five-rings.

Let us compare the six-ring and the five-ring. In the former we can place one sphere, touching all 6 spheres. In the latter two spheres can be placed, each touching the plane of the centres of the 5 spheres, and also the 5 spheres themselves. Both configurations so obtained contain 7 spheres. In the six-ring there are 12 contact points, whereas this number may be 15 ("almost 16") for the five-ring. The six-ring is the base of six regular tetrahedra in close-packing, whereas the five-ring may contain four regular tetrahedra and a slightly irregular one in itself.

The configuration considered above has been built around a point (the centre of the first tetrahedron). It turns out to be possible to build up close configurations along a straight line also. We will consider three types of such "needles".

There is a needle that contains regular tetrahedra only. This is possible because a regular tetrahedron can be screwed around a certain axis in such a way that one of its faces in the original position coincides with another of its faces in the second position. In fact a regular tetrahedron has 12 of such screw-axes, 6 right-handed and 6 left-handed. The lines joining the mid-points of opposite edges in a regular tetrahedron are mutually perpendicular and all pass through the centre of the tetrahedron. Each of the screw-axes intersects one of these lines perpendicularly in a point located at a distance of $\frac{1}{10}\sqrt{2}$ from the centre of the tetrahedron and encloses an acute angle $\arctan 3 \approx 71°34'$ with the direction of the nearest of the two opposite edges. In fig. 1 we have drawn a regular tetrahedron projected onto a plane at right angles to one of its screw-axes $Z$. Along this axis identical tetrahedra can be piled so that they fit together exactly (fig. 2$a$).

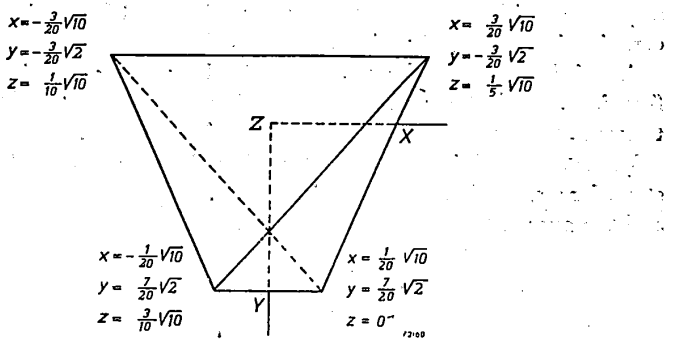



Fig. 1. Projection of a regular tetrahedron of edge-length 1 onto a plane at right angles to one of its 12 screw-axes.

Each tetrahedron is rotated relative to its neighbours through an angle $\alpha = 2 \arctan \sqrt{5} \approx 131°49'$ around $Z$. The relative displacement in the direction of the screw-axis is $\frac{1}{10}\sqrt{10}$. It appears that $\alpha$ is incommensurable with $\pi$. So the orientations of no two tetrahedra can be the same. Placing equal spheres of diameter 1 with their centres on the vertices of the tetrahedra we obtain the first needle configuration (fig. 2*b*). In this type of needle every sphere has 6 contact points. Since the 12 screw-axes occur in 6 pairs of members crossing at right angles two such needles can cross each other at right angles, having in common four spheres with their centres on the vertices of one tetrahedron.

A second more complicated needle may be obtained by adding to each of the two free faces of each tetrahedron in the above-mentioned needle

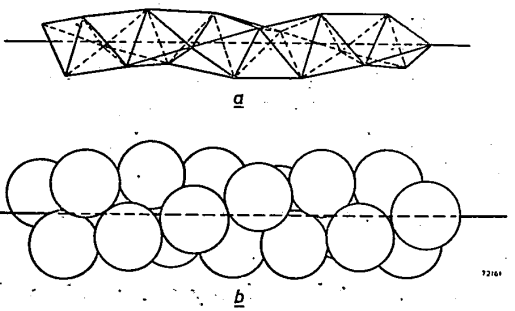



Fig. 2. *a*. Regular tetrahedra piled along a straight line so that they fit together exactly. This line is a screw-axis common to all these tetrahedra. *b*. First needle configuration formed by equal spheres of radius 1, of which the centres are on the vertices of the tetrahedra of fig. 2*a*.

a sphere with its centre on the symmetry axis of that face. This configuration contains interwoven five-rings.

A third type of needle can be obtained by piling identically oriented five-rings, filled up by spheres with their centres on a straight line at distances 1. Our configuration 6-1 is a part of this needle.

The latter needle is especially interesting, since it can be used as a building stone for a close plane configuration. In fig. 3 the plane of this configuration and the axes of the needles are perpendicular to the plane of the drawing.

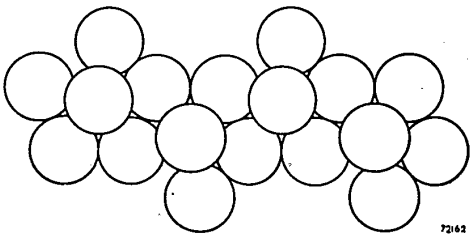



Fig. 3. Close plane configuration obtained by piling needle configurations of the third type with alternate orientations of the five-rings. The plane and the axes of the needles are perpendicular to the plane of the drawing.

To investigate the density of the above-mentioned configurations our criteria I and II can be extended in an obvious way. Instead of considering the volume of a sphere of radius $R$ we have to consider the volumes of a cylinder of unit length and radius $R$, or that of a rectangular parallelopiped with edge-lengths 1, 1 and $2R$. In this way we find that in the needles and in the plane configurations the local mean density surpasses that of any configurations of the same type occurring in close-packing, according to one or more of our three criteria.

Piling of the plane configuration, or parts of it, leads to configurations filling all space. The mean density however does not exceed that of close-packing. Further investigation shows that it is likely that 3/4 of all space could be filled by parallel needles of the first type. In that 3/4 part the mean density is about 0·73. This is rather high, considering the absence of periodicity in the direction of the axes of the needles.

## 4. Proof that the maximal number of spheres simultaneously touching a sphere is twelve

Fejes [3]) has shown that for arbitrary configurations of equal spheres we must have for the mean density

$$D < \frac{4\pi}{30\sqrt{2}(65-29\sqrt{5})} < 0{\cdot}7547 \,. \qquad (11)$$

This inequality of course only holds for the mean density over all space. So (3) is not in contradiction with it. His proof is based on the assumption that no more than 12 spheres can touch a central sphere (all spheres of equal diameters). He considers this an experimental fact. We shall give here a brief account of a proof of this assumption.

Using central projection, with the centre of the central sphere as a centre, we can transform the problem to a spherical one. This can be stated as follows. How many circles with a geodetical radius of arcsin $\frac{1}{2}$ can be drawn on the surface of a sphere with radius 1? Of course the circles may not overlap.

The area of each of the circles will be

$$A_1 = 2\pi\,(1 - \tfrac{1}{2}\sqrt{3}) > 0{\cdot}8417. \tag{12}$$

The number $C$ of circles therefore cannot exceed

$$\frac{4\pi}{A_1} = \frac{2}{1 - \frac{1}{2}\sqrt{3}} < 15, \tag{13}$$

i.e. $C$ will be 14 or less.

With each of the circles however a certain empty area will be associated. As we have for the dihedral angle $2\varphi$ of a regular tetrahedron

$$\frac{2\pi}{6} < 2\varphi < \frac{2\pi}{5}, \tag{14}$$

the number of circles that can simultaneously touch a certain circle cannot exceed 5. The closest configuration in the immediate neighbourhood of a certain circle will be like fig. 4. This is a part of the surface of the sphere drawn in stereographic projection, in which the plane of the drawing touches the sphere at the centre of the central circle in the drawing. Each of the areas $A_2$ is associated in the same way with each of the three circles sur-

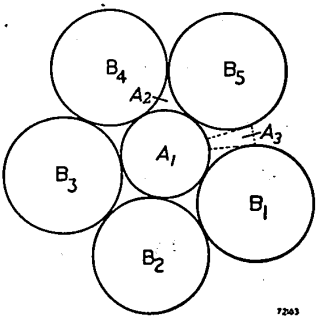



Fig. 4. Stereographic projection of a part of the surface of a sphere of radius 1, on which 6 equal circles of radius arcsin $\frac{1}{2}$ have been drawn, five of which touch a central one.

rounding it. As there are 5 of such areas, the empty area associated with the central circle will be at least

$$\tfrac{5}{3} A_2 = \tfrac{5}{3}\left(6\varphi - \pi - \frac{6\varphi}{2\pi} A_1\right) > 0{\cdot}0942\,. \tag{15}$$

It is obvious that this empty area will be associated with each of the circles. So $C$ cannot exceed

$$\frac{4\pi}{A_1 + \tfrac{5}{3} A_2} < 14\,, \tag{16}$$

i.e. $C$ will be 13 or less.

An investigation shows that with each of the circles also an empty area $A_3$ will be associated, which area has a radial shape. Around the centre of the central circle it will have an extension

$$\gamma = 2\pi - 10\varphi > 7°21'. \tag{17}$$

Of course it may be divided into several sectors by spacing the 5 circles in an arbitrary way along the circumference of the central circle. This will not influence the total angular extension of the sum of the radial areas. The radial length however will have a minimal value if this empty area is not divided between the circles. In that case it will extend over a distance of

$$\delta = \arctan\left\{\sqrt{3}\cos(\pi - 4\varphi)\right\} > 53°24' \tag{18}$$

from the centre of the central circle. So we have

$$A_3 \geqslant \gamma\,(\tfrac{1}{2}\sqrt{3} - \cos\delta) > 0{\cdot}0346\,. \tag{19}$$

It turns out that the area $A_3$ can never be effectively common to two or more circles, nor can any part of it. This can be seen as follows. Building the configuration further we can consider the circle $B_1$ as a new centre. We have already three circles ($A_1$, $B_2$ and $B_5$) around $B_1$. Of the other two one (say $B_6$) could touch $B_1$ and $B_2$, the other one ($B_7$) could touch $B_1$ and $B_5$. Then there must be a gap between $B_6$ and $B_7$. Calculation shows that this gap gives a bigger empty space than $A_3$. Of course we may displace the circles, but this only displaces the gaps and does not make them smaller. Therefore $C$ cannot exceed

$$\frac{4\pi}{A_1 + \tfrac{5}{3} A_2 + A_3} < 13\,, \tag{20}$$

i.e. $C$ will be 12 or less.

Returning to our original problem we find that the maximal number $C$ of spheres that may simultaneously touch a sphere (all spheres of equal diameters) is 12. This result is in agreement with the suggestion of Schütte

and van der Waerden [5]) (i.e. that for 13 equal outer spheres the diameter of the central sphere has to be 1·045 at least), but in contradiction with the assertion of Gregory (in an unpublished note mentioned in [6])) that 13 equal nonoverlapping spheres can touch another sphere of the same diameter. Since each of the contact points is associated with two spheres, the number of contact points for any configuration of $K$ spheres cannot exceed $6K$. So in close-packing it is already maximal [2]).

## 5. Proof that the conjecture of Fejes concerning a fourteenth sphere added to a configuration of twelve spheres touching a central sphere is false

Fejes states as a conjecture that, if a sphere is touched by 12 spheres (all spheres of diameter 1), a 14th sphere added to this configuration will have a distance from the centre of the central sphere in excess of 1·38. He mentions "a crude experiment" in support of his conjecture.

Apparently he has considered a configuration similar to our configuration 5-1, to which the 14th sphere is added with its centre on the symmetry axis of one of the faces of the icosahedron and pressed towards the central sphere until the other 12 spheres are in their most close position on the surface of the central sphere. This configuration indeed supports his conjecture.

From our configuration 6-2 and from (10) however it is clear that this is false, as the 14th sphere is here at a distance of 1·347 from the centre of the central sphere. We have not succeeded in proving that shorter distances are impossible. With the aid of (20) an estimation may be obtained of a lower limit of this distance. This is much lower than 1·347.

We express our gratitude to Dr Wise for informing us in detail of the inspiring contents of his paper on close-packing of unequal spheres.

*Eindhoven, January 1952*

## REFERENCES

) D. R. Hudson. Density and packing in an aggregate of mixed spheres, J. Appl. Phys. **20**, 154-162, 1949.
) H. S. M. Coxeter. Regular Polytopes, London, 1948, pp. 69-71.
) L. Fejes. Über die dichteste Kugellagerung, Math. Z. **48**, 676-684, 1943.
) A. H. Boerdijk, Thesis Delft 1951, stelling 25.
) K. Schütte and B. L. van der Waerden, Auf welcher Kugel haben 5, 6, 7, 8 oder 9 Punkte mit Mindestabstand Eins Platz? Math. Ann. **123**, 96-124, 1951.
) H. S. M. Coxeter, reviewing [5]) in Math. Rev. **13**, 61, 1952.

IPR2025-00217 EXHIBIT 2010 Page 317 of 485

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

Reprints of the majority of these papers can be obtained on application to the Administration of the Research Laboratory, Kastanjelaan, Eindhoven, Netherlands. Those papers of which no reprints are available are marked with an asterisk (*).

**1982:** N. Warmoltz: On the application of a Philips ionization Gauge type of ion source in a mass spectrometer leak detector (Appl. sci. Res. **B2**, 61-65, 1951).

A system of electrodes as in the Philips ionization Gauge can function as ion source giving a diverging ion beam in which the ions have different velocities. When placing this source behind a slit, it is possible to focus all the ions of the same mass in a second slit by combining a magnetic field parallel to the slits with an eletric field perpendicular to the slits and to the line joining the slits. Between the slits the ions describe a kind of epicycloidal trajectories. In this way it is possible to construct a simple leak detector.

**1983:** C. J. Bouwkamp and H. Bremmer: A note on Kline's Bessel-function expansion (Proc. Kon. Ned. Akad. Wetensch. Amsterdam **A 54**, 130-134, 1951).

A certain combination of Bessel functions, viz. $J_\nu(\nu \sec\vartheta)\, Y_\nu(z + \nu\sec\vartheta) - Y_\nu(\nu\sec\vartheta)\, J_\nu(z + \nu\sec\vartheta)$, in which $\nu$, $\vartheta$, $z$ are complex numbers, is expanded in a power series in the variable $1/\nu$. This series converges if $|\nu| > |z\cos\vartheta|$. The first three coefficients are given explicitly.

**1984*:** J. H. van Santen and G. H. Jonker: Combinaisons ferromagnétiques du manganèse à structure pérovskite (J. Phys. Radium **12**, 202-204, 1951). (Ferromagnetic compounds of manganese with perovskite structure; in French).

From investigations on the magnetic behaviour of ferromagnetic mixed crystals with perovskite structure of the type $(1 - x)$ La $Mn^{III}O_3 - x$ Me $Mn^{IV}O_3$ (Me = Ca, Sr or Ba) it seems probable that the (indirect) exchange interaction between $Mn^{IV}$ and $Mn^{IV}$ is negative and that between $Mn^{III}$ and $Mn^{III}$ is positive. This is the first known example of a positive indirect exchange interaction in oxidic compounds.

**1985:** J. J. Went: Linear magnetostriction of homogeneous nickel alloys (Physica **17**, 98-116, 1951).

The magnetostriction of homogeneous Ni-alloys is investigated. The

magnetostriction is measured as a function of the composition of the alloy, of the induction caused by an external magnetic field and of the temperature. A comprehensive table of data is given. Several general relationships for binary alloys are found. The saturation magnetostriction $\lambda_s$ at 0°K can be predicted if the saturation induction $I_s$ is known; $\lambda_s$ at a higher temperature follows directely from the $\lambda_s$ value at 0 °K and the relation $I_s$ vs $T$ (except for Ni-Mn-alloys). A single relationship is found between $\lambda$ and $I$ for alloys except those containing Co or Mn. From all these data the magnetostriction for ternary alloys may be calculated.

**1986:** G. W. Rathenau and G. Baas: Grain growth in a texture, studied by means of electron-emission microscopy (Physica **17**, 117-128, 1951).

Grain growth and secondary recrystallization in rolled face-centred Ni-Fe-alloys is studied by electron-optical means, an image of the activated hot metal surface being formed. Grain growth in an imperfect cubic texture proved to be a discontinuous process. Almost all neighbouring grains invade quickly one grain or group of grains. Grain-boundary movement at low-energy boundaries occurs on a small scale. The high-energy parts of the boundary between a cubic crystal and its (near) twin proved to move in a direction parallel to the common (111) plane. The surface tension of the boundary between a secondary crystal and a cubic crystal which is to be absorbed is about twice the value corresponding to the surface between two cubic crystals.

**1987:** W. de Groot: A graphical method for solving problems of low-level photometry (Appl. sci. Res. **B 2**, 131-148, 1951).

A certain level of subjective brightness may be characterized by the minimum power per unit area and solid angle, $E_m$, with which the corresponding brightness impression can be evoked monochromatically, the corresponding wavelength being $\lambda_m$. If the visibility function $V(\lambda_m, \lambda)$ for that level of subjective brigthness is known, it is possible to find the luminance $L_E = K \int E(\lambda)\ V_p(\lambda) \mathrm{d}\lambda$ of a radiation $E(\lambda)$ for which $\int E(\lambda) V(\lambda_m, \lambda)\ \mathrm{d}\lambda = E_m$, giving the same brightness impression as $E_m$. This is achieved bij a graphical method which can be incorporated in a slide rule for low-level photometry. Luminance scales for different relative energy distributions are easily obtained and, after choosing one of them as a standard of subjective brightness, other luminances may be corrected according to the standard scale chosen. Luminances for monochromatic radiations of wavelength $\lambda$ may be obtained in the same way. When plotting these versus $\lambda$ for a number of values of $E_m$ one obtains Bouma's diagram in which the data for composite radiations may be also drawn.

IPR2025-00217 EXHIBIT 2010 Page 319 of 485

An analysis is given of Weaver's data regarding $V(\lambda_m, \lambda)$, and the various functions discussed in the preceding sections are calculated using these data. Finally the function $E_m$ $(\lambda_m)$, as obtained from these calculations, is checked with other data from literature.

**1988:** J. J. Went and H. P. Wijn: The magnetisation process in ferrites (Phys. Rev. **82**, 169-270, 1951).

Data are obtained relating to the magnetization curves of certain ferrites at low field strengths ($\mu_0 H = 10^{-4}$ Wb/m$^2$) as a function of frequency (0·3 - 1000 kc/s). From these measurements it may be deduced that in ferrites with high permeability at field strengths of the order of $H_c$ an additional magnetization process occurs which (a) is frequency-dependent at frequencies below the gyromagnetic resonance frequency, (b) takes place at a lower field strength as the density of the ferrite is greater, (c) is irreversible. Contrary to conclusions arrived at by Rado e.a. the additional process is ascribed to irreversible Bloch-wall displacements whereas the initial permeability is thought to be related to a pure rotational process.

**1989:** F. A. Kröger and P. Zalm: The origin of multiple bands in luminophors activated by divalent manganese (J. electrochem. Soc. **98**, 177-182, 1951).

The two principal theories which have been advanced for the explanation of the appearance of multiple bands with magnanese-activated luminophors, particularly the silicates and germanates, are the "cluster theory" and the "coordination theory". The former accounts quite naturally for the dependence of the red-to-green ratio on the Mn concentration, whereas the coordination theory requires both the effect of coordination and concentration quenching (or possible some other factor) for a self-consistent picture. On the other hand, the effect of Be in promoting the red emission is quite naturally explained by the coordination theory, whereas the cluster theory requires an additional postulate concerning the effect of Be in promoting clusters. Taking these and other considerations into account, the cluster theory seems to present the simpler and more satisfactory picture. In the cluster theory, the variation of luminescence observed is attributed to the interaction between manganese ions belonging to clusters, and it seems that this influence must affect mainly the lifetime of the various excited states. It seems probable that the fluorescence bands are due to electronic transitions from various excited states to one single ground state and that several different bands may originate in one centre. The variation of fluorescence with temperature can be explained on the basis of this picture.

**1990*:** P. Cornelius: Kurze Zusammenfassung der Elektrizitätslehre, eine Einführung des rationalisierten Giorgischen Maßsystems. (Wien, Springer Verlag, 1951, VI + 89 pp., 11 figs, 7 tables) (A short survey of the theory of electricity, an introduction of the rationalized Giorgi system of electromagnetic units; in German).

For the contents of his booklet the reader is referred to abstract nr **1803**. The German edition contains an addendum, in which the author deals in short with possible objections of physicists regarding his way of exposition. An English edition is in preparation.

**1991:** J. L. Snoek: Conférence sur les propriétés magnétiques des ferrites (J. Phys. Radium **12**, 228-238, 1951). (Conference on magnetic properties of ferrites; in French).

Survey of magnetic properties of ferrites, in which the following points are discussed: $Fe_3O_4$, ferrites $XY_2O_4$, ferrite technology, inverted spinels, Néel's theory of antiferromagnetism, variation of saturation magnetization with temperature, magnetic resonance, magnetic crystal energy, magnetostriction, dispersion and absorption at high frequencies. Follows a report of a discussion, in which E. W. Gorter deals with experimental evidence for the influence of the angle Fe-O-Fe on the interaction.

**1992:** A. Bierman and W. Hondius Boldingh: The relation between tension and exposure time in radiography (Acta Radiologica **35**, 22-25, 1951).

Experiments, carried out under widely varying conditions, have proved that the density of a radiograph remains practically constant when varying kV and mAs in such a way that the value of the product $(kV)^5$ mAs is kept constant. Application of this rule does not give intolerable deviations in the whole field of diagnostics.

**1993:** J. J. Zaalberg van Zelst: Over versterkers met weinig veranderde versterking (T. Ned. Radiogenootschap **16**, 117-135, 1951). (On amplifiers with nearly constant gain; in Dutch).

This article contains a survey of methods providing a reduction of the variation of the gain of amplifiers. The methods using at least some constant network elements are divided into two groups. In those of the first group the influence of the actual change of the constants of the valves is eliminated by corrections of the input or output voltages, in those of the second group the actual change is reduced by regulating voltages derived from an auxiliary signal. It is also shown to be possible to construct an

amplifier having small variations in gain when using elements which are not constant but have a limited variation. The fundamental limits of this precedure are discussed, regarding oscillators to be selective amplifiers of noise with positive feedback.

**1994:** J. A. Kok: On electrostatic plasma oscillations in metals (Physica 17, 543-547, 1951).

In this article an attempt is made to apply to metals the equations of the electrostatic plasma oscillations and ions in electrical discharges in gases. It is pointed out that for electrons in metals the Fermi-Dirac distribution law has to be used and not the Maxwell statistics. The frequencies of the electrostatic ion oscillations have an upper limit of the same order of magnitude as the Debije maximum frequency.

**1995:** J. J. Went: The value of the spontaneous magnetization of binary nickel alloys as a function of temperature (Physica 17, 596-602, 1951).

The $I_s$ versus $T$ curve for pure nickel is more concave towards the $T$-axis than that for any binary nickel alloy, with only one exception, viz. completely ordered alloys, such as slowly-cooled $Ni_3Fe$. For pure nickel the form of the $I_s$ vs $T$ curve can be explained by the occurrence of an order-disorder phenomenon of the magnetic moments where the latter can placed only parallel or anti-parallel. It is suggested that all the other $I_s$ vs $T$ curves must be explained as being a result of this order-disorder phenomenon and the statistical fluctuations of the concentration of dissolved atoms. In the case of $Ni_3Fe$ two different order-disorder phenomena (a crystallographic and a magnetic) act simultaneously.

**1996:** J. Smit: Magnetoresistance of ferromagnetic metals and alloys at low temperatures (Physica 17, 612-672, 1951).

The magnetoresistance of pure Ni and Fe, of Ni-Fe-, Ni-Co-, and Ni-Cu-alloys, and of Heusler's alloy has been measured at room temperature and at temperatures of liquid nitrogen and liquid hydrogen. The behaviour of the pure metals is essentially different from that of the alloys. At low temperatures a maximum in the magnetoresistance is observed for alloys having about one Bohr magneton per atom. The positive difference between the longitudinal and the transversal resistance can be explained by means of the spin-orbit interaction. At low temperatures the pure metals show an increase in resistance with increasing field just as the non-ferromagnetic metals. From this the value of the internal field, acting on the conduction electrons, could be determined and was found to be approximately equal to the induction $B$.

1997: J. A. Haringx: De instabiliteit van inwendig op druk belaste dunwandige cylinders (De Ingenieur **63**, O39-O41, 1951). (The instability of thin-walled cylinders subjected to internal pressure; in Dutch).

It is shown that under certain conditions a thin-walled cylinder may buckle when subjected to internal pressure. The critical value of this pressure can easily be calculated on account of Euler's well-known formula. Most of the formulae for pressure-loaded cylinders given in current textbooks, however, fail to predict this behaviour and should therefore be applied with caution for great lengths and/or high pressures.

1998: J. A. Haringx: De instabiliteit van inwendig op druk belaste ronde balgen (De Ingenieur **63**, O42-O44, 1951. (The instability of cylindrical bellows subjected to internal pressure; in Dutch).

Like thin-walled cylinders, dealt with in the preceding abstract, also bellows may become unstable when loaded by internal pressure. The critical value of this pressure, which is governed through Euler's well-known formula by the rigidity of the bellows with respect to bending, is computed only for rectangularly shaped corrugations. The result found has been checked experimentally.

1999: C. G. Koops: On the dispersion of resistivity and dielectric constant of some semiconductors at audiofrequencies (Phys. Rev. **83**, 121-124, 1951).

Semiconducting $Ni_{0.4}Zn_{0.6}Fe_2O_4$, prepared in different ways, has been investigated. It appeared that the AC resistivity and the apparent dielectric constant of the material show a dispersion which can be explained satisfactorily with the help of a simple model of the solid: there should be well-conducting grains separated by layers of lower conductivity. Dispersion formulae are given. There is good agreement between experiment and theory.

2000: E. J. W. Verwey: Oxidic semi-conductors (from: Semi-conductory Materials, Buttersworth's Scientific Publications Ltd, London, 1951, pp. 151-161).

Non-conducting oxidic materials may be rendered conductive either by preparing solid solutions with oxides of high conductivity (e.g. a poorly conducting spinel with $Fe_3O_4$) or by introducing into the lattice ions of deviating valency (method of controlled valency). A comparison is made between semi-conductors obtained is this way and elemental semi-conductors such as silicon and germanium. The conductive properties of polycrystalline semi-conductors as a function of frequency are discussed, especially the influence of non-conducting thin layers at the grain boundaries.

**2001:** J. Volger: Some properties of mixed lanthanum and manganites (from: Semi-conductory Materials, Butterworth's Scientific Publications Ltd, London, 1951, pp. 162-171).

Discussion of electrical properties of manganites ($XMnO_3$). This material is an example of a polycrystalline semi-conductor prepared according to the principle of controlled valency (replacing the bivalent ion X, e.g. Sr., by a bivalent ion, such as La). There exists a close analogy between tha resistivity and the ferromagnetic properties. Discussion of anomalies te the Curie-temperature, of the influence of frequency, electric field and external magnetic field on the resistivity, of the thermo-electric properties and the Hall effect.

**2002:** H. Bremmer: On the diffraction theory of Gaussian optics (Comm. Pure appl. Math. 4, 61-74, 1951).

The diffraction theory of optical imaging is developed for objects with arbitrary structure. The theory is based on rigorous solutions of the wave equation instead of the conventional approximation of Kirchhoff's formula. The similarity of the wave functions in the object plane and in the corresponding paraxial image plane (Gaussian systems with unlimited aperture) proves to be connected with Neumann's integral theorem for Bessel functions (instead of the Fourier identity as in the Kirchhoff approximations). Another solution does account for the effects of optical aberrations and of limited apertures.

**2003:** H. Bremmer: The W.K.B. approximation as the first term of a geometric-optical series (Comm. Pure appl. Math. 4, 105-115, 1951).

The W.K.B. approximation of the solution of $y'' + k^2(x)\, y = 0$ is derived from a discontinuous model of an inhomogeneous medium. Higher approximations are found by considering multiple reflections. The solutions of different order form a series, the convergence of which is discussed. The well-known insufficiency of the W.K.B. approximation in the neighbourhood of zeros of $k(x)$ can be interpreted as a very slow convergence of the series mentioned.

**2004:** W. Ch. van Geel: On rectifiers (Physica **17**, 761-776, 1951).

Experiments are described in which rectifiers were obtained from combinations of metals, semiconductors and intermediate layers. The following combinations have been considered: (a) the contanct between an excess semiconductor and a deficit semiconductor: (b) the combination aluminium/aluminium oxide/semiconductor; and (c) the combination metal/ resin layer/ semiconductor. In all these cases rectification occurs. The suggestion is put forward that, in all three cases, the contact between two layers with charge carriers of opposite sign is the cause of rectification.

VOL. 7 NR 5 OCTOBER 1952

# Philips Research Reports

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

*R 198* *Philips Res. Rep. 7, 321-343, 1952*

## DENSE RANDOM PACKING OF UNEQUAL SPHERES

by M. E. WISE

513.468

**Summary**

A heap of rigid spheres has a given typical statistical distribution for their radii, and the spheres are densely packed as they might be 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 for which every sphere has to touch the other three. Thus $w$ is defined by 4 radii. Boundary effects are ignored. General equations for $w$ are deduced and a possible numerical solution 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 apply (e.g.) little-known formulae of solid geometry and for integrating numerically in 2, 3 and 4 dimensions.

**Résumé**

Soit donnée une collection d'un grand nombre de sphères rigides, avec une distribution statistique typique pour leurs rayons, les sphères étant empillées de façon dense comme les grains d'une poudre. On définira d'une nouvelle manière „l'empillage dense au hasard" dans les termes d'une fonction de distribution de la probabilité $w$ pour des tétraèdres; chaque tétraèdre a ses sommets au centre de quatre sphères, dont chacune touche les trois autres. Ainsi donc $w$ est defini par quatre rayons. On ne considère pas les effets au bord. On déduit les équations générales pour $w$ et une solution numérique possible pour une distribution logarithmique normale des rayons; ce résultat est utilisé pour obtenir des approximations pour plusieurs des plus intéressantes propriétés géométriques de la configuration. De cette manière, par example, on trouve la densité moyenne, et le nombre moyen d'interstices, et aussi une distribution statistique pour les sphères les plus grandes possibles qui peuvent être logées dans les interstices, et une pour le nombre des sphères qui touchent une sphère de rayon donné. Finalement, on fait une discussion critique du système considéré comme un modèle d'une collection réelle. Les méthodes mathématiques présentent également de l'intérêt: elles utilisent entre autres des formules peu connues de la géométrie dans l'espace et de l'intégration numérique dans 2, 3 et 4 dimensions.

**Zusammenfassung**

Behandelt wird ein Haufen von festen Kugeln, für dessen Halbmesser eine feste typische Verteilungsfunktion gegeben ist. Die Kugeln sind

dicht gepackt wie etwa die Körner in einem Pulver. Die dichte Packung und regellose Packung werden auf einem neuen Wege mittelst einer Wahrscheinlichkeitsverteilung $w$ für Tetraeder definiert; die Spitzen jedes Tetraeders liegen in den Mittelpunkten von vier Kugeln, deren jede die anderen drei berührt; $w$ ist also durch vier Halbmessern bestimmt. Kugeln an der Aussenseite werden nicht betrachtet. Es werden allgemeine Gleichungen für $w$ abgeleitet und eine numerische Lösung gegeben für den Fall, daß die Logarithmen der Halbmesser eine Gausssche Verteilung haben. Diese Lösung wird gebraucht, um zu Näherungswerten für einige der interessantesten geometrischen Eigenschaften des Haufens zu gelangen. Auf diese Weise wird zum Beispiel die mittlere Dichte und die mittlere Anzahl der Zwischenräume ermittelt, sowie auch die Verteilungsfunktion für die größten Kugeln, die sich innerhalb der Zwischenräume noch unterbringen lassen; man findet auch die Verteilungsfunktionen für die Anzahl der Kugeln, die eine andere Kugel von gegebenem Halbmessern berühren. Zum Schluß wird das System als Modell eines reellen gepackten Haufen kritisch diskutiert. Von Interesse sind auch die rechnerischen Methoden: hierbei kommen beispielweise wenig bekannte Formeln aus der Stereometrie und der numerischen Integration in 2, 3 und 4 Dimensionen zur Anwendung.

## 1. Introduction

A densely packed powder is quite unlike a lattice. Usually the volumes of the particles have a statistical distribution, and round any one particle there will be a more or less random sample of the rest. In this paper we assume all the particles are rigid spheres, we start with a typical size distribution, and we try to calculate the most interesting geometrical properties of the corresponding densely packed heap. The original stimulus was a pressing need for theoretical help in two different laboratory problems involving powders. They had in common that the physical properties of the powders depended (from experiments) very much upon the size distributions of the particles and were known to depend on the geometry of packing. One of the aims was then to relate these three things to one another.

So the emphasis is on obtaining a numerical result. But the way of getting it is of general interest, and the approach to the general problem seems new. It is one of those elementary mathematical problems that are nearly insoluble, but if we consider only the densest possible packing (or nearly so) it becomes much easier.

There has been plenty of work on lattice packing, with one sphere radius, or at most two. What we have done is not a generalization of this work. In fact the type of random packing defined here is impossible if the spheres are all equal; then the highest proportion of space that can be filled in an infinite lattice is believed to be 0·74, but with dense random packing we shall see that one can obtain at least 0·8.

With equal spheres, a density of about 0·78 can be obtain *locally*, and even for certain infinite spaces: Boerdijk [1]) has recently shown that spheres can be closely packed round a certain pair of infinite lines that do not intersect; some spheres of course belong to each of the two systems. The type of packing studied in our paper is one of several types given

in his. This deals with equal spheres packed regularly, but not necessarily in a lattice, and it is easy to see from his results why unequal spheres can be packed with a higher density than equal ones.

What we shall call random packing can be demonstrated with plasticine balls. A fairly large ball $A$ is picked out, and others are put on its surface. (In a real powder it is assumed these are kept in place by the pressure of surrounding balls.) The first and second balls are made to touch each other as well as $A$, and a third ball is to touch all three. Every new ball is made to touch $A$ and *at least* two others that touch each other. If $A$ is larger than the others, looking out from its centre we shall see something like fig. 1.

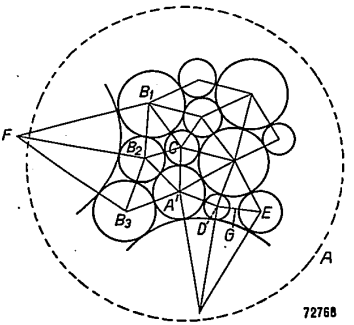



Fig. 1. Small spheres all touching one large one $A$. All the spheres touching $A$ and (say) $C$ form a complete ring round $A$ and $C$. The rings cannot always be completed; for example there is a gap between spheres $D$ and $E$. All lines of centres of touching (or nearly touching) spheres are drawn. These form triangles, and also faces of tetrahedra with vertices at the centre of $A$. (Note: one line of centres has been left out accidentally.)

Almost any size of sphere can touch $A$ and two others: usually just one can be found of the right radius to touch $A$ and three others — such as $B_1$, $B_2$ and $B_3$. Sometimes a gap like $G$ is unavoidable. But if the larger spheres are picked out first it will nearly always be possible to fit the smaller ones in afterwards.

Now consider the lines from the centre of a sphere such as $D$ to all the spheres round $D$ (except $A$ itself) touching sphere $D$ or nearly touching it. (There is no sphere *between* $D$ and $E$.) Then these lines form a network of triangles, each in a different plane. The centres of the spheres are the vertices of a polyhedron and the triangles are its faces. Next join the vertices to the centre of $A$. Then the whole space enclosed by the polyhedron is partitioned into tetrahedra. We call the system densely packed, and so define *dense* packing, when for every tetrahedron *every one of its spheres touches every other one.*

This is so for every tetrahedron in fig. 1 except the one with edge $DE$. And for any real system there will always be a few gaps like $G$, — especially if the statistical distribution of radii has only a small spread. On the other hand imperfect spheres should pack better than perfect ones. For example if the spheres round $D$ except $E$ are already in place, an ellipsoid replacing $E$ might more nearly fill the gap. In what follows, we shall only work with the limiting case, and assume there are no gaps.

Why there is no dense packing when all the spheres are equal can be easily seen from the polyhedron round any one sphere such as $A$. There is no regular polyhedron with the right circum-radius (i.e. radius of the sphere through its vertices) equal to its length and to the diameters of the touching spheres. But if the middle sphere has radius 0·902, and the outer spheres have unit radii, 12 of them can be closely packed round the middle one and their vertices form a regular icosahedron [2]), [5]).

Obviously we are not limitcd to one sphere inside a polyhedron. Next we start from one of the touching spheres like $F$ and again put round it all the spheres touching it: many of course are already there because they touch $A$. Every sphere except those on the outside of the packed heap has in the same way a polyhedron round it, and every tetrahedron is determined uniquely by four radii.

We shall not consider the outside *), and what follows is strictly true within a certain large polyhedron: its vertices are the centres of the outside spheres, and its faces, from the definition of dense packing, are all plane triangles.

The density of the system is obviously defined to be a weighted sum of densities of tetrahedra: for in each tetrahedron the proportion of filled space is the ratio of its volume to the sum of the volumes of the four solid angles, i.e. the four it has in common with its spheres. If the spheres are equal this ratio is 0·78, but we shall soon see that any other tetrahedron gives a greater density.

The density can be made much greater if the interstices are filled. To define an interstice is simple: it is the empty space that is inside the tetrahedron but outside its four mutually touching spheres. With a typical radius distribution (such as we shall work with) there are very few spheres small enough to go into the interstices, and we assume that there are none in them. But a useful practical problem is indeed to find the distribution of the largest spheres that will go into the interstices, that is, of a second distribution with one sphere for each interstice of the original densely packed heap (section 8).

*) For some general theorems concerning packing of spheres into a *finite* cylinder see [3]). This is a completely different approach. Also Rogers [8]) has worked on the packing of unequal circles, and Fejes [9]) on discs packed into or covering a finite area.

It is less easy to define *random* packing rigorously. The physical idea behind it has been suggested — we assemble the spheres at random and rearrange them to make the heap denser. Let us imagine a closed surface, small compared with the total volume of the heap but large enough to contain the centres of many spheres, perhaps 200 on the average. Then in random packing the number and radii of the spheres ought to be a random sample from the heap. Similarly, in *dense* random packing, given the radius distribution, there must certainly be a non-zero probability of obtaining every possible tetrahedron. There is a probability distribution function — in four dimensions — for the four radii. It must have a simple weighting factor for the statistical distribution, and another one for size, since a large sphere must nearly always have more spheres round it and be part of more tetrahedra than a small one. Apart from this the probability ought to be as uniform as possible.

If this tetrahedron probability function can be found, so can everything else about the heap that depends on its packing. We shall put down equations which it must satisfy, suggest a solution, and apply it to find some of the most interesting geometrical properties. A four-dimensional probability function seems formidable in practice, but it is made manageable by a little-known formula of numerical integration. In any case the calculations should delight anyone interested in spherical trigonometry and solid geometry, — several fascinating but little-used formulae are taken out of the museum.

## 2. Mathematical formulation of the problem

The tetrahedron probability function is a function of the four radii of the mutually touching spheres only: if there are $N$ tetrahedra, the number with "first" radius between $r_1$ and $r_1 + \mathrm{d}r_1$, ..., "fourth" radius between $r_4 + \mathrm{d}r_4$ is written: $Nw\,(r_1, r_2, r_3, r_4)\,\mathrm{d}r_1\,\mathrm{d}r_2\,\mathrm{d}r_3\,\mathrm{d}r_4$. Of course $w$ is normalized thus

$$\int_0^\infty\int\int\int w(r_1, r_2, r_3, r_4)\,\mathrm{d}r_1\,\mathrm{d}r_2\,\mathrm{d}r_3\,\mathrm{d}r_4 = 1. \qquad (2.1)$$

It is difficult to define 1st, 2nd, 3rd and 4th, but it is not really necessary! We could let $r_1$, $r_2$, $r_3$ and $r_4$ be in order of increasing distances above a fixed plane. For example, the heap might be formed by particles falling out of a suspension, forming a sediment, and afterwards being packed more closely: the coarser particles fall faster and will be nearer the bottom. But for strictly random packing whatever distinction we make must disappear, i.e. $w$ must be symmetrical in $r_1$, $r_2$, $r_3$ and $r_4$. Just as in statistical mechanics the labels are removed after the groups have been counted.

Next we weight $w$ according to the statistical distribution of $r$, and

IPR2025-00217 EXHIBIT 2010 Page 329 of 485

roughly also by a positive power of $r$; the latter is the simplest plausible function increasing with $r$ that we can have. So we put

$$w(r_1, r_2, r_3, r_4) = P(r)P(r_2)P(r_3)P(r_4)\ (r_1 r_2 r_3 r_4)^p\ W_p(r_1, r_2, r_3, r_4), \quad (2.2)$$

where $P(r)\mathrm{d}r$ is the proportion of spheres (by number) with radii between $r$ and $r + \mathrm{d}r$; $p$ is an unknown positive power, and $W_p(r_1, r_2, r_3, r_4)$ an unknown symmetrical function which (we hope) does not vary too much over the different possible tetrahedra (i.e. different points in the four-dimensional space). There is one necessary condition for $W_p$. It must be such that the tetrahedra fit together. They do so in the way shown in fig. 1. Let us consider all spheres with "first" radius between $\varrho_1$ and $\varrho_1 + \mathrm{d}\varrho_1$. One of them touches spheres of radii $r_2$, $r_3$ and $r_4$ and the corresponding tetrahedron cuts off

$$A(\varrho_1; r_2, r_3, r_4) \quad (2.3)$$

of the volume of that sphere, where $A$ is $1/4\ \pi$ times the corresponding solid angle at the centre of the $\varrho_1$ sphere. For all tetrahedra with this "first" radius, we sum this ratio $A$, over all possible values of the other three radii. Then if there are $M$ spheres and $N$ tetrahedra

$$N\int_0^\infty\!\!\int\!\!\int w(\varrho_1; r_2, r_3, r_4)\ A(\varrho_1; r_2, r_3, r_4)\ \mathrm{d}r_2\,\mathrm{d}r_3\,\mathrm{d}r_4 = \tfrac{1}{4} MP(\varrho_1). \quad (2.4)$$

This must be true for all values of $\varrho_1$. The factor $\frac{1}{4}$ comes from symmetry: for every tetrahedron with "first" radius between $\varrho_1$ and $\varrho_1 + \mathrm{d}\varrho_1$ equal numbers must have 2nd, 3rd and 4th radii in the same range of values of $\varrho_1$.

Instead of thinking of one sphere at a time and summing the contributions from all tetrahedra of which it is a part, we can go round one tetrahedron and sum its solid angles. Doing this for all tetrahedra gives

$$N\int_0^\infty\!\!\int\!\!\int\!\!\int w(r_1, r_2, r_3, r_4)\left\{A(r_1; r_2, r_3, r_4) + A(r_1, r_2; r_3, r_4) + \right.$$
$$\left. + A(r_1, r_2, r_3; r_4) + A(r_1, r_2, r_3, r_4;)\right\}\ \mathrm{d}r_1\,\mathrm{d}r_2\,\mathrm{d}r_3\,\mathrm{d}r_4 = M. \quad (2.5)$$

$A(r_1, r_2; r_3, r_4)$ is $1/4\pi$ times the solid angle at the "second" sphere.

This is important because it gives us the number of tetrahedra, even without knowing the $W_p$ function. For the numerical calculations show that for many tetrahedra the sum of the four $A$'s is about the same; hence for many distributions of sphere radii, the mean number $N/M$ of tetrahedra (and of interestices) per sphere is about the same. From Table I(f) it is about five. Thus the mean number associated with any one sphere is $4 \times 5 = 20$ (cf. Table I (c)); of course it varies statistically. Before trying to solve these equations we shall derive some other interesting general ones.

IPR2025-00217 EXHIBIT 2010 Page 330 of 485

### 3. Contact points

Many physical properties must depend directly on the distribution of pairs of touching particles. For example suppose the spheres are electrically charged and afterwards packed together from a thick suspension. Then for a pair to come together work must be done against a potential barrier and the amount of this work may be a simple function of the two radii. In a different problem, the material of the spheres may conduct more than the medium, and then it must do so across contact points (or in practice small areas round them) *). We need the distribution of pairs, and from this the number and probability distribution of spheres touching any one sphere, and they can be calculated from the tetrahedron frequency function.

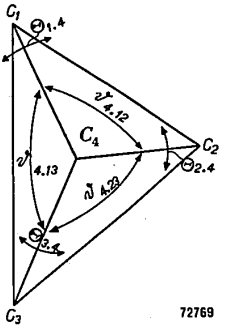



Fig. 2. The notation for some of the angles for the tetrahedron of centres of mutually touching spheres, centres $C_1$, $C_2$, $C_3$ and $C_4$.

In fig. 1 suppose the spheres $A$ and $A'$ have the given radii. Six spheres touch these two, and just six tetrahedra are thereby formed, which have a common edge. The tetrahedron $AA'B_2B_3$, e.g., takes up some of the space round this edge — an amount equal to $1/2\pi$ of the angle (in radians) between its faces through the edge — faces $AA'B_2$ and $AA'B_3$. Then to calculate the probability for any one pair of radii the edge angles for every other two radii — divided by $2\pi$ — are summed. Again we have to be careful about symmetry. If $L$ is the total number of contact points and $LP(\varrho_1, \varrho_2)\mathrm{d}\varrho_1\mathrm{d}\varrho_2$ is the number of pairs with one radius $\varrho_1$ and another radius $\varrho_2$, with $\varrho_2$ "after" $\varrho_1$, $\varrho_1$ and $\varrho_2$ can be respectively "1st" and "2nd",

*) This is purposely given rather vaguely. One may be interested in diffusion or electrical or heat conduction, or in the analogous problem of the dielectric constant of a closely packed system embedded in a medium, etc., etc.

IPR2025-00217 EXHIBIT 2010 Page 331 of 485

or the "1st" and "3rd", "2nd" and "4th", or one of the other three pairs, and hence

$$\frac{1}{6} LP(\varrho_1, \varrho_2) = \frac{1}{2\pi} \int_0^\infty\!\!\int N w(\varrho_1, \varrho_2, r_3, r_4)\ \Theta_{\varrho_1\varrho_2}(r_3, r_4)\ \mathrm{d}r_3\, \mathrm{d}r_4 \qquad (3.1)$$

where $\Theta_{\varrho_1\varrho_2}(r_3, r_4)$ is the angle across the edge $C_1C_2$ in the tetrahedron $C_1C_2C_3C_4$ of centres of spheres with radii $r_1, r_2, r_3$ and $r_4$ respectively.

The number of spheres in the completed ring round such a pair is not necessarily six of course. Their total number (after summing over all rings) is equal to the number of tetrahedra having these two radii, $\varrho_1$ and $\varrho_2$; thus the mean number $R(\varrho_1, \varrho_2)$ of spheres in the ring is given simply by

$$LP(\varrho_1, \varrho_2)\, R(\varrho_1, \varrho_2) = 6N \int_0^\infty\!\!\int w(\varrho_1, \varrho_2, r_3, r_4)\, \mathrm{d}r_3\, \mathrm{d}r_4. \qquad (3.1')$$

The mean of $R$ over all pairs of spheres is thus $6N/L$.

The mean number of spheres touching a sphere of radius $\varrho_1$ is clearly

$$T(\varrho_1) = \frac{L}{P(\varrho_1)} \int_0^\infty P(\varrho_1, \varrho_2)\, \mathrm{d}\varrho_2. \qquad (3.2)$$

Again we can concentrate on one tetrahedron, and sum all its edge angles expressed as fractions of the cylinder round a pair in contact. Doing this for all tetrahedra gives

$$N \iiint \frac{w(r_1, r_2, r_3, r_4)}{2\pi} \{\Theta_{r_1r_2}(r_3, r_4) + \Theta_{r_1r_3}(r_2, r_4) + \Theta_{r_1r_4}(r_2, r_3) + \\ + \Theta_{r_2r_3}(r_1, r_4) + \Theta_{r_2r_4}(r_1, r_3) + \Theta_{r_3r_4}(r_1, r_2)\}\, \mathrm{d}r_1\, \mathrm{d}r_2\, \mathrm{d}r_3\, \mathrm{d}r_4 = L. \qquad (3.3)$$

The sum of these angles, again, is nearly the same for many different tetrahedra; it is very simply related to the sum of the solid angles. In fact it is easy to prove from spherical trigonometry (see next section) that in any one tetrahedron (we now abbreviate $\Theta$ and $A$ in an obvious way)

$$\frac{1}{2\pi} \{\Theta_{1.2} + \Theta_{1.3} + \Theta_{1.4} + \Theta_{2.3} + \Theta_{2.4} + \Theta_{3.4}\} = \\ = 1 + A_{1.234} + A_{2.341} + A_{3.412} + A_{4.123}, \qquad (3.4)$$

i.e., the sum of the solid angles (as ratios of a whole sphere) plus 1 equals the sum of the edge angles (as ratios of a whole circle).

Now summing over all tetrahedra gives

$$\frac{L}{N} = 1 + \frac{M}{N}, \qquad (3.5)$$

whatever is the distribution of tetrahedra — neglecting surface effects. This result also follows from Euler's theorem, by applying it to the polyhedron formed by the centres of spheres round any one sphere *).

Thus the *average* number of spheres that touch any one sphere is about 12. (1 + 5 = 6 contact points per sphere, from equations (3.5), (2.5) and Table I). The number is exactly twelve for face-centred cubic lattice packing of equal spheres, and for the central sphere of the icosahedron packing [1]), [2]). The mean number round any one sphere of course increases rapidly with the radius: we shall be able to estimate how rapidly in the particular distribution chosen. But first we will give briefly the needed spherical trigonometry. The main argument is taken up again in the next section but one.

## 4. Calculation of angles and solid angles in the tetrahedra

This section is interesting in itself, and necessary for any further practical calculations: the formulae are easily found from known ones, but I have not seen them given explicitly anywhere.

The lengths of the edges of a tetrahedron are $r_1 + r_2$, $r_1 + r_3$, $r_2 + r_3$, and so on. The angles between the edges are first needed, and are found by plane trigonometry. Two typical formulae, for $\vartheta_{4\cdot 13}$ or $C_3\ C_4\ C_1$ — see fig. 2 — are:

$$\cos \vartheta_{4\cdot 13} = \frac{2r_4\,(r_1 + r_3 + r_4)}{(r_1 + r_4)\,(r_3 + r_4)} - 1, \tag{4.1}$$

$$\tan^2 \tfrac{1}{2}\,\vartheta_{4\cdot 13} = \frac{r_1\, r_3}{r_4\,(r_1 + r_3 + r_4)}. \tag{4.2}$$

The tetrahedron cuts off a spherical triangle (its sides are in diametric planes) on the surface of each sphere. On sphere nr 4 the sides are equal to the plane angles $\vartheta_{4\cdot 13}$, $\vartheta_{4\cdot 23}$ and $\vartheta_{4\cdot 12}$. The angles of this triangle are those between faces: the angle opposite $\vartheta_{4\cdot 13}$ is between planes $C_2\ C_4\ C_3$ and $C_2\ C_4\ C_1$: thus it can just as well be written $\Theta_{4\cdot 2}$. This is calculated from an 'S' formula:

$$2S_4 = \vartheta_{4\cdot 13} + \vartheta_{4\cdot 23} + \vartheta_{4\cdot 31},$$

$$\tan^2 \tfrac{1}{2}\,\Theta_{4\cdot 2} = \tan^2 \tfrac{1}{2}\,\Theta_{4\cdot 13} = \frac{\sin\,(S_4 - \vartheta_{4\cdot 12})\,\sin\,(S_4 - \vartheta_{4\cdot 23})}{\sin S\,\sin\,(S - \vartheta_{4\cdot 13})} \tag{4.3}$$

The areas of each spherical triangle are obtained from the spherical excess, i.e.

$$A\,(r_4; r_1, r_2, r_3) = \frac{1}{4\pi}\left\{\Theta_{4\cdot 1} + \Theta_{4\cdot 2} + \Theta_{4\cdot 3}\right\} - \frac{1}{4}. \tag{4.4}$$

It is now easy to prove (3.4) from formulae like (4.4) for $A_1$, $A_2$ and $A_3$.

The volume $V$ of a tetrahedron in terms of its edges is the square root of a 5 × 5 determinant, and in terms of the four radii it is given by

$$\begin{aligned} 9V^2 = {} & 2r_1r_2r_3r_4\,(r_1r_2 + r_1r_3 + r_1r_4 + r_2r_3 + r_2r_4 + r_3r_4) - \\ & - (r_1r_2r_3r_4)^2\,\left(\overline{r_1}^2 + \overline{r_2}^2 + \overline{r_3}^2 + \overline{r_4}^2\right). \end{aligned} \tag{4.5}$$

It is very much simpler when two radii are equal. If $r_4 = r_1$,

$$9V^2 = r_1^3\left\{4r_2r_3(r_2 + r_3) - r_1(r_2 - r_3)^2\right\}. \tag{4.6}$$

*) Van Oosterhout of this laboratory produced a neat proof of this. It is also possible to allow approximately for the surface: one must think in terms of solid angles rather than complete tetrahedra.

IPR2025-00217 EXHIBIT 2010 Page 333 of 485

## 5. Numerical calculations for a particular distribution of sphere radii

It is surprisingly easy to allow for the statistical distribution of radii This always comes into the various calculations as a factor in a multipl integral, and a little-known rule for numerical integration greatly simplifie these. The distribution chosen is a Gaussian one, not for $r$ but of log $r$; i should be a function of log $r$ because the geometrical properties of the hea are not then changed if every sphere is magnified by the same constant and it is made Gaussian because in many powders the distribution of equi valent radii (radii of spheres of equal volume) *is* approximately log-normal

The integral over all $x$ of a product of a Gaussian formula $\varphi(x)$ and polynomial in $x$ is a simple "selected-ordinate" formula. If $F(x)$ is a poly nomial of the 5th (or lower) degree and $\varphi(x) = (2\pi)^{-\frac{1}{2}} \exp(-\frac{1}{2}x^2)$, the

$$\int_{-\infty}^{+\infty} F(x)\,\varphi(x)\,dx = \frac{1}{6}\left\{F(-\sqrt{3}) + 4F(0) + F(+\sqrt{3})\right\}. \qquad (5.1$$

This result is attributed to Berger [4].

It is easy to extend this to 2, 3 and 4 dimensions. $F = F(x_1, x_2, x_3, x_4$ must be such that

$$\frac{\partial^6 F}{\partial x_i^6} = 0 \text{ for } i = 1, 2, 3 \text{ and } 4. \qquad (5.2$$

In $n$ dimensions, if $F$ is symmetrical in the $x_i$, for the term that ha $-\sqrt{3}\,p_0$ times, $0\,p_1$ times and $+\sqrt{3}\,p_2$ times, the coefficient of th corresponding $F$ is that in

$$\left\{\frac{1}{6}(x_0 + 4x_1 + x_2)\right\}^n \text{ of } x_0^{p_0}\, x_1^{p_1}\, x_2^{p_2}.$$

Thus in two dimensions writing $F_{01}$ for $F(-\sqrt{3},0)$, $F_{02}$ for $F(-\sqrt{3},+\sqrt{3}$ and so on:

$$36 \int_{-\infty}^{+\infty}\!\!\int F(x_1,x_2)\,\varphi(x_1)\,\varphi(x_2)\,dx_1 dx_2 = F_{00} + F_{22} + 16F_{11} + 8(F_{01}+F_{12}) + 2F_{02} \qquad (5.3$$

Similarly in three dimensions (for four dimensions see "$m_1$" in Table I):

$$216 \int_{-\infty}^{+\infty}\!\!\int\!\!\int F(x_1,x_2,x_3,)\,\varphi(x_1)\,\varphi(x_2)\,\varphi(x_3)\,dx_1\,dx_2\,dx_3 = F_{000} + F_{222} +$$

$$+ 3(F_{002}+F_{022}) + 12(F_{001}+F_{122}) + 24F_{012} + 48(F_{011}+F_{112}) + 64F_{111}. \qquad (5.4$$

For a log-normal distribution with standard deviation $\sigma$, the unit variat is $x$ and $x = (1/\sigma)\log(r/\bar{r})$. We shall take $\bar{r} = 1$. Then (see fig. 3($a$))

$$\varphi(x)\,dx = P(r)\,dr \text{ and } P(r) = (1/r\sigma)\,\varphi(x). \qquad (5.5$$

Thus the selected radii in (7.3) to (7.5) are 1 and $\exp(\pm\sigma\sqrt{3})$. But we cannot use this rule straight away, for in all our multiple integrals $\varphi(x)$ will be multiplied by $e^{p\sigma x}$, which certainly does not satisfy (5.2). But the numerical integration is valid and still convenient if $\exp(p\sigma x_i - \frac{1}{2}x_i^2)$ is included in a new variable. Further, we should expect the tetrahedron

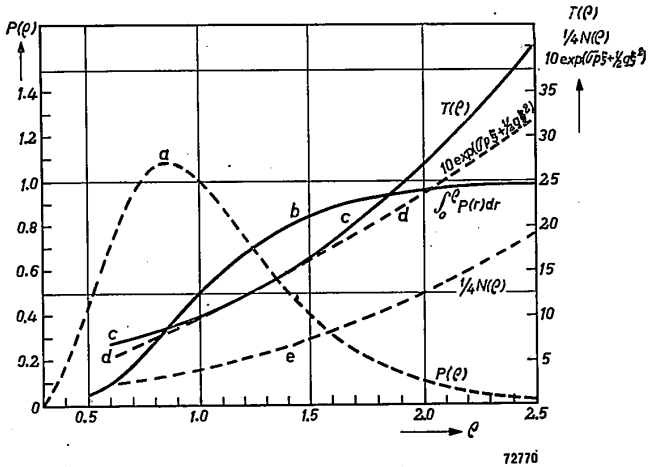



Fig 3. (*a*) Probability distribution function for the sphere radii (by numbers); (*b*) Corresponding proportions of spheres with radii less than $\varrho$; (*c*) mean numbers of spheres touching a sphere of radius $\varrho$; (*d*) the factor $\exp(p\sigma\xi_i+\frac{1}{2}q\xi_i^2)$ of the tetrahedron probability function; (*e*) $\frac{1}{4}\times$ mean number of tetrahedra of centres that a sphere of radius $\varrho$ is part of.

probability function to deviate from a simple $r^p$ law at extreme radii. So besides $r^p$ we put into the weighting factor a second-degree term $\exp(\frac{1}{2}qx_i^2)$. Thus we write

$$w(r_1,r_2,r_3,r_4)\,\mathrm{d}r_1\,\mathrm{d}r_2\,\mathrm{d}r_3\,\mathrm{d}r_4 = W_{p,q}(r_1,r_2,r_3,r_4)\prod_{i=1}^{4} e^{p\sigma x_i+\frac{1}{2}qx_i^2}\,\varphi(x_i)\mathrm{d}x_i. \quad (5.6)$$

The exponent in (5.6) is then changed into a perfect square by putting

$$y_i = (1-q)^{\frac{1}{2}}\,x_i - p\sigma(1-q)^{-\frac{1}{2}}. \quad (5.7)$$

Then $$w = (1-q)^{-2}\exp\left\{\frac{2p^2\sigma^2}{1-q}\right\} W_{p,q}\,(r_1,r_2,r_3,r_4)\prod_{i=1}^{4}\varphi(y_i)\,\mathrm{d}y_i. \quad (5.8)$$

The selected radii must correspond to $y_i = 0$ and $\pm\sqrt{3}$: we write them as $r_0$, $r_0\lambda$ and $r_0/\lambda$, and it is easily seen that

$$r_0 = \exp\{p\sigma^2/(1-q)\}, \tag{5.9}$$

$$\log\lambda = \sigma\sqrt{\frac{3}{1-q}}. \tag{5.10}$$

For $q = 0$ and $\lambda = 2$ we have $\sigma = 0{\cdot}4002$. For the numerical calculations it is convenient to fix $\lambda$, and then $\sigma$ varies, but only slightly, since $q$ must be small. (This is why $\sigma$ is slightly below 0·4 in curve fig. 3($a$)). So all solid angles (as ratios of the whole sphere), and angles across edges (as ratios of a whole circle) have been worked out for all tetrahedra in which each radius is either $\frac{1}{2}r_0$, $r_0$ or $2r_0$. There are 3.4.5.6./4! or 15 different ones. The first few rows explain themselves: row ($f$) and row $m_2$ provide the numerical evidence for $N = 5M$: rows $m_3$, $m_4$ and $m_5$ will be explained later.

We can get some idea of what happens when the standard deviation of log $r$ is changed. For example if $\sigma = 0{\cdot}8$, $\lambda$ is about 4. Then the tetrahedra 1,1,4,4, and $^1/_4$, $^1/_4$, 1,1 have weight 96; 1, 1, 4, 1 has weight 256, and so on. But these have identical angles to those in tetrahedra on the right of the table. Three of these contribute high densities $\Sigma/V$, viz. 0·846, 0·849 and 0·824. On the other hand the mean density for the log-normal distribution with $\sigma = 0{\cdot}4$ is almost certain to be near to 0·8. Clearly the density of dense packing — provided as always that the interstices are empty — increases with $\sigma$, but rather slowly. For the 10 differently shaped tetrahedra $\Sigma/V$ is greatest when two large spheres touch two small ones and when one large sphere touches three small ones. The tetrahedron is then long and narrow and most of its volume is included in the largest sphere. But if *two* large spheres touch, the tetrahedra might be very small or even non-existent.

There is an extremely small probability of finding a sphere small enough to go into an interstice when $\sigma = 0{\cdot}4$, but it is not negligible in a distribution with $\sigma = 0{\cdot}8$. To see this, let three large spheres have radius each $\lambda r_0$ and one small one have radius $r_0/\lambda$. The volume of the tetrahedron of centres is then given by

$$9V^2 = r_0^6\,\lambda^2\,(3 + 6\lambda^2 - \lambda^4) \tag{5.11}$$

and it is imaginary if $\lambda \geqslant 2{\cdot}542$ ($\lambda^2 = 6{\cdot}464$). Then the fourth sphere can be put inside the planar ring formed by the other three. The $W_{p,q}$ factor must thus be zero for such tetrahedra.

In any case we can get denser packing by increasing the variance of the distribution; the practical objection is that this might spoil some other property of the powder. Then it would be better, as said before, to try to fill the interstices with a finer powder deliberately added.

IPR2025-00217 EXHIBIT 2010 Page 336 of 485

TABLE I

| $r_1, r_4, r_2, r_3$ $(r_1 = r_4)$ | 1111 | $11\frac{1}{2}1$ | 1121 | $11\frac{1}{2}2$ | $\frac{1}{2}\frac{1}{2}11$ | 1122 | $\frac{1}{2}\frac{1}{2}12$ | $22\frac{1}{2}1$ | $\frac{1}{2}\frac{1}{2}1\frac{1}{2}$ | 2212 | $\frac{1}{2}\frac{1}{2}22$ | $\frac{1}{2}\frac{1}{2}2\frac{1}{2}$ | $22\frac{1}{2}2$ | $\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}$ | 2222 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $(a)$ | 0 | — | + | + | 0 | | + | — | + | — | 0 | + | — | 0 | 0 |
| $(a') = (r_1 r_2 r_3 r_4)^{-\frac{3}{4}} V/0{\cdot}9428$ | 1 | 0·9861 | 1·0082 | 0·9843 | 1 | | 1·0191 | 0·9402 | 1·0082 | 0·9861 | 1 | 1·0534 | 0·8291 | 1 | |
| $(b)$ $\Theta_{1.2}/2\pi$ | 0·19513 | 0·23406 | 0·17793 | 0·25951 | 0·19879 | | 0·21205 | 0·26410 | 0·17793 | 0·23406 | 0·20613 | 0·17607 | 0·31357 | 0·19513 | |
| $\Theta_{1.3}/2\pi$ | ,, | 16348 | 21728 | 14474 | 19879 | | 19231 | 14806 | 21728 | 16348 | 20613 | 23120 | 11018 | ,, | |
| $\Theta_{1.1}/2\pi$ | ,, | 16348 | 21728 | 17135 | 25000 | | 27305 | 11614 | 21728 | 16348 | 30281 | 23120 | 11018 | ,, | |
| $\Theta_{2.3}/2\pi$ | ,, | 23406 | 17793 | 23291 | 14758 | | 13130 | 29605 | 17793 | 23406 | 10817 | 17607 | 31357 | ,, | |
| $(c)$ $A_{1.234}$ | 0·04387 | 03050 | 05625 | 03780 | 07379 | | 08871 | 01420 | 05625 | 03050 | 10817 | 06642 | 01696 | 0·04387 | |
| $A_{2.341}$ | ,, | 10108 | 01690 | 12596 | 02258 | | 02769 | 16221 | 01690 | 10106 | 01022 | 00570 | 22036 | ,, | |
| $A_{3.412}$ | ,, | 03050 | 05625 | 01121 | 02258 | | 00796 | 04608 | 05625 | 03050 | 01022 | 06642 | 01696 | ,, | |
| $(d)$ $\frac{4}{3}\pi\Sigma(A_i r_i^3)$ | 0·73505 | 0·4363 | 1·2731 | 0·7579 | 0·2665 | 2·1320 | 0·4756 | 1·2991 | 0·1591 | 3·4904 | 0·7981 | 0·2954 | 1·8207 | 0·0919 | 5·880 |
| $(e') = V$ | 0.9428 | 0·5528 | 1·5986 | 0·9280 | 0·3333 | 2·6667 | 0·5713 | 1·4907 | 0·1998 | 4·422 | 0·9428 | 0·3511 | 2·2111 | 0·11785 | 7·542 |
| $(e) = \Sigma/V$ | 0·7796 | 0·7900 | 0·7965 | 0·8160 | 0·7994 | | 0·8325 | 0·8240 | 0·7965 | 0·7900 | 0·8470 | 0·8490 | 0·8240 | 0·7796 | |
| $(f) = \frac{M}{N} = \frac{L}{N} - 1 = A$ *) | 0·17548 | 0·19270 | 0·18405 | 0·21277 | 0·19912 | | 0·21306 | 0·23650 | 0·18405 | 0·19270 | 0·23551 | 0·20501 | 0·27125 | 0·17548 | |
| $6^4 m_1$ | 256 | 256 | 256 | 192 | 96 | | 48 | 48 | 16 | 16 | 6 | 4 | 4 | 1 | |
| $6^4 m_2$ | 44·92 | 49·33 | 47·12 | 40·85 | 19·12 | | 10·23 | 11·35 | 2·945 | 3·083 | 1·413 | 0·820 | 1·085 | 0·175 | |
| $6^3 m_3$ | 0 | 6.468 | 0 | 6·078 | 3·541 | 0 | 2·129 | 1.947 | 0·675 | 0 | 0·3245 | 0·1993 | 0·2204 | 0·0439 | 0 |
| $6^3 m_4$ | 2·808 | 1·464 | 2·700 | 0·9072 | 0·2712 | 0·8854 | 0·0831 | 0·1382 | 0·0169 | 0·1011 | 0 | 0 | 0 | 0 | 0 |
| $6^3 m_5$ | 0 | 0 | 1·0816 | 0·5381 | 0 | 1·0840 | 0·0955 | 0·3408 | 0 | 0·3660 | 0·0307 | 0·0057 | 0·0509 | 0 | 0·0439 |

*) $A = 2A_{1.234} + A_{2.341} + A_{3.412} = 1 + \{\Theta_{2.3} + \Theta_{1.1} + 2\Theta_{1.2} + 2\Theta_{1.3}\}/2\pi$.

$M$ = number of spheres; $N$ = number of tetrahedra, $L$ = number of contact points (each "shared" between two touching spheres).

Geometry of tetrahedra of centres of four mutually touching spheres.
$(a)$ sign of skewness (see end of section 6);
$(a')$ volume of tetrahedron compared with the cube of the geometric mean of the four radii;
$(b)$ $1/2\pi$ times angles across edges;
$(c)$ $1/4\pi$ times solid angles;
$(d)$ sum of four solid angles;
$(d')$ volume of tetrahedron;
$(e)$ ratio of $(d)$ to volume of tetrahedron;
$(f)$ sum of four solid angles as fractions of unit sphere;
$m_1$ numerical coefficient in four-dimensional integral (6.2);
$m_2$ the same multiplied by corresponding values of $A$ in $(f)$;
$m_3, m_4, m_5$ numerical coefficients in three special cases of (6.1).

## 6. Estimating the tetrahedron frequency function

In every sphere the solid angles fit together. This was expressed by (2.4), which is easily transformed into ($i = 3$, 4 or 5)

$$\iiint_{-\infty}^{+\infty} W_{p,q}(r_1, r_2, r_3, r_4)\, A_{1.234}\, \varphi(y_2)\, \varphi(y_3)\, \varphi(y_4)\, dy_2\, dy_3\, dy_4 = m_i = \\ = (1-q)^{3/2} \left( \exp \frac{-2p^2\sigma^2}{1-q} \right) \frac{r_1}{\varphi(y_1)}.$$

Then from (5.5) and (5.7)

$$n_i = \frac{M}{4N} (1-q)^{3/2} \exp \left[ \frac{-qy_1^2}{2(1-q)} - \frac{y_1 p\sigma}{(1-q)^{1/2}} - p^2\sigma^2 \left\{ \frac{1}{2(1-q)^2} + \frac{2}{1-q} \right\} \right]. \quad (6.1)$$

As particular cases of this equation, we can put $y_1 = -\sqrt{3}$, 0 and $+\sqrt{3}$ in turn. Thus we obtain three more equations and the unknowns are the $W_{p,q}$ factors for 15 tetrahedra, and $p$ and $q$. The left-hand side of (6.1) is expanded according to (5.4); the coefficients of the corresponding $W_{p,q}$ are at the foot of Table I, and $m_3$, $m_4$ and $m_5$ are the values of the triple integral for $y_1 = -\sqrt{3}$, 0 and $+\sqrt{3}$ respectively.

There are two more equations: The normalizing condition (2.1) becomes

$$n_1 = \iiint_{-\infty}^{+\infty} W_{p,q} \varphi(y_1) \varphi(y_2) \varphi(y_3) \varphi(y_4) dy_1 dy_2 dy_3 dy_4 = (1-q)^2 \exp \frac{-2p^2\sigma^2}{1-q}. \quad (6.2)$$

Thus $m_1$ is the sum of products of each tabulated coefficient multiplied by the corresponding $W$, just as for the other $m$'s. We also need $M/N$, and this comes from row $m_2$ in the table:

$$n_2 = \iiint_{-\infty}^{+\infty} W_{p,q} \left\{ \sum_{i=1}^{4} A(r_i) \right\} \varphi(y_1) \varphi(y_2) \varphi(y_3) \varphi(y_4) dy_1 dy_2 dy_3 dy_4 = \frac{m_1 M}{N}, \quad (6.3)$$

where $\sum_{i=1}^{4} A(r_i)$ is the sum of the four solid angles for one tetrahedron, each as fractions of a whole sphere — just as in (2.5).

So now there should be at least enough equations to tell us whether any reasonable solutions are possible. "Reasonable" means that the 15 unknown $W_{p,q}$ coefficients must not be too different from each other, they must vary in a systematic way with the four radii and of course they must be positive. Besides this $p$ and $q$ are unknown. $q$ as a correction factor is probably small. But the $(r_1\, r_2\, r_3\, r_4)^p$ factor might be guessed at. If one of the $r$'s is large the factor might be roughly proportional the volume of a shell of fixed thickness round that sphere and so to the square of that $r$. Another guess is that it is related to the chance that any volume is common to a pair of spheres if their centres are distributed at random; that would give $p = 1{\cdot}5$.

Or, other things being equal, the factor might be proportional to the volume of the tetrahedron, and then $p = 0{\cdot}75$. We find in practice for this particular distribution that solutions are impossible if $p$ is greater than about 1·3. To see this we try a reasonable law for $W_{p,q}$, and re-calculate the five $m$'s. The equations (6.1) can then be greatly simplified: $\log \{m_4^2/m_3 m_5\}$ comes to $3\left(\frac{1}{q-1} - 1\right)$ and so gives $q$: then $\log (m_3/m_5)$ and (5·10) give $p[\lambda = 2$, thus $\sigma^2 = 0{\cdot}16p^2(1-q)]$.

The third equation to be obtained from (6.1) is then

$$f(p,q) = -2\log\frac{m_2}{4m_4} + \log(1-q) + \frac{\sigma^2 p^2}{(1-q)^2} = 0. \qquad (6.4)$$

But it was very difficult to make $f(p,q)$ zero and I have found only one way to do it. With any simple law based on the frequencies of single radii it is positive, and also with many laws based on pairs of radii. Instead it seems that the skewness of the group of four radii is important: groups like 2, 2, 2, 1 must have a small weighting factor (always apart from the size factor and the distribution-function factors: these are already allowed for) and groups like 1, 1, 1, 2 a large one. In Table I the signs of the skewnesses are given. Weighting only the positively skew groups did not help, nor did lowering the weighting factors of the negatively skew groups only: $f(p,q)$ was still positive. But $f(p,q)$ is nearly zero if $W_{p,q}$ is 5 for all positively skew groups, 1 for all symmetrical ones and 0·2 for negatively skew ones. Then $q$ comes to 0·011 and $p$ is 1·165; $N/M$ comes to 5·10, and the density is between 0·8 and 0·805. The absolute value of the middle $W_{p,q}$ from (6.2) is then 0·2646 and $\sigma = 0{\cdot}3980$; the probability distribution and percentages are given in fig. 3. The "selected radii" are $\frac{1}{2}r_0$, $r_0$ and $2r_0$ with $r_0 = 1{\cdot}205$.

Of course a real distribution cannot be so simple, but this is at least a possible one, and it seems very plausible. So it is worthwhile to deduce various geometrical properties from it.

## 7. Numerical estimations of the numbers of contact points and of tetrahedra belonging to a sphere of given radius

(i) First we shall calculate the mean number of tetrahedra $N(\varrho_1)$ that a sphere of radius $\varrho_1$ belongs to. It is given by

$$\tfrac{1}{4}MP(\varrho_1)\, N(\varrho_1) = N \int_0^\infty\!\!\int\!\!\int w(\varrho_1, r_2, r_3, r_4)\, \mathrm{d}r_2\, \mathrm{d}r_3\, \mathrm{d}r_4, \qquad (7.1)$$

for $\frac{1}{4}N(\varrho_1)$ is the number whose "first" radius is $\varrho_1$. Hence

$$N(\varrho_1) = \frac{4N}{M} \exp\left\{p\sigma x_1 + \tfrac{1}{2}qx_1^2 + \frac{3p^2\sigma^2}{2(1-q)}\right\} I_N(\varrho_1), \qquad (7.2)$$

here

$$I_N(\varrho_1) = \iiint_{-\infty}^{+\infty} W_{p,q}(\varrho_1, r_2, r_3, r_4)\, \varphi(y_2)\, \varphi(y_3)\, \varphi(y_4)\, dy_2\, dy_3\, dy_4 \qquad (7.3)$$

=0·59425, 0·56872 and 0·81409 for the three selected radii. Hence $N(\frac{1}{2}r_0)$ = 7·40, $N(r_0) = 20{\cdot}0$, $N(2r_0) = 65{\cdot}9$ (fig. 3). (Compare these values with $4A$ from Table I(c).) Over the whole distribution of course every tetrahedron is counted four times and the mean of $N$ is about 20.

i) The mean number of contact points increases in much the same way ith the radius, but relatively faster for large radii. First we find the six equencies for pairs of selected radii. From (3.1) we easily find that

$$\frac{1}{6} LP(\varrho_1, \varrho_2) = \frac{N}{1-q}\left(\exp\frac{2p^2\sigma^2}{1-q}\right) \frac{\varphi(\eta_1)\varphi(\eta_2)}{\sigma^2\varrho_1\varrho_2} I(\varrho_1, \varrho_2) \qquad (7.4)$$

here

$$I(\varrho_1, \varrho_2) = \iint_{-\infty}^{\infty} W_{p,q}(\varrho_1, \varrho_2, r_3, r_4) \frac{\Theta_{1\cdot 2}(r_3, r_4)}{2\pi} \varphi(y_3)\, \varphi(y_4)\, dy_3\, dy_4, \qquad (7.5)$$

nd $y_1$, $y_2$ are equal to $\eta_1$, $\eta_2$, respectively when $r_1 = \varrho_1$, $r_2 = \varrho_2$.

The (normalized) values of $I$ for the six special pairs of radii are given Table II. Its last column gives reduced values: for $P(r_i)\, P(r_j)\, dr_i\, dr_j$ the probability of choosing the corresponding pair of spheres at random om the heap. We expect these reduced values to increase with the radii. hus the high values for 1, 2 and $\frac{1}{2}$, 2 compared with that for 2,2 are teresting. This is entirely due to the skewness effects.

TABLE II

The $I$ function of equation (7.4), the corresponding joint probability nction for contact pairs, and the same as ratio of the corresponding obability for picking the two radii at random; $r_0 = 1{\cdot}205$.

| $r_i/r_0, r_j/r_0$ | $\frac{1}{2}$ $\frac{1}{2}$ | $\frac{1}{2}$ 1 | $\frac{1}{2}$ 2 | 1 1 | 1 2 | 2 2 |
|---|---|---|---|---|---|---|
| $I(r_i, r_j)$ | 0·1870 | 0·1118 | 0·1989 | 0·1005 | 0·1584 | 0·0223 |
| $P(r_i, r_j)$ | 0·2019 | 0·2581 | 0·0537 | 0·5447 | 0·0958 | 0·001507 |
| $\dfrac{P(r_i, r_j)}{P(r_i)\, P(r_j)}$ | 0·3686 | 0·4680 | 2·009 | 0·9812 | 3·563 | 1·157 |

(iii) The mean numbers of spheres $R$ in the rings round the six pairs ar now easily calculated, from (3.1), by a similar substitution, which give

$$R(\varrho_1, \varrho_2) = \int_0^\infty\int W_{p,q}\,(\varrho_1, \varrho_2, r_3, r_4)\,\mathrm{d}r_3\,\mathrm{d}r_4/I\,(\varrho_1, \varrho_2). \qquad (7.5$$

For the six representative pairs:

$$\begin{array}{l} \varrho_1/r_0, \varrho_2/r_0 = \frac{1}{2}\quad\frac{1}{2};\ \frac{1}{2}\quad 1;\ \frac{1}{2}\quad 2;\ 1\quad 1:1\quad 2;\ 2\quad 2; \\ R(\varrho_1, \varrho_2)\ \ = 4{\cdot}087;\ 4{\cdot}206;\ 4{\cdot}583;\ 4{\cdot}839;\ 5{\cdot}971;\ 6{\cdot}843; \end{array}$$

and the mean of $R$ over all pairs is almost exactly five ($\bar{R} = 5{\cdot}015$). $R$ i five for equal spheres, but then the ring cannot be completed [1]). Thi fact, indeed gives perhaps the most convincing proof that dense packin in the sense of this paper is impossible when all the spheres are equal.

For any one pair the distribution of numbers of spheres in the surround ing ring must have a very small spread. A ring of three is rare, even fo the $\frac{1}{2}r_0$, $\frac{1}{2}r_0$ pair; for example if the radii of the spheres in the pair ar equal (to $\varrho_1$) the radii of the spheres in the ring must each be about $6\varrho$ (see equation (5.11)), and the chances of this are negligible for the chose $P(r)$ (fig. 3*b*). So for the $\frac{1}{2}r_0$, $\frac{1}{2}r_0$ pair nearly every ring must have fou spheres, and most of those in $\frac{1}{2}r_0$, $2r_0$ and $r_0$, $r_0$ must have either 4 or 5 One can show in a similar way that even a ring of 4 is rare for the pair $r_0$, $2r_0$ and $2r_0$, $2r_0$.

(iv) Next we find the mean number $T(\varrho_1)$ of spheres touching a sphere c radius $\varrho_1$; using (3.2) this comes to

$$T(\varrho_1) = \frac{12N}{M}\,(1-q)^{-3/2}\left\{\exp\frac{3p^2\sigma^2}{2(1-q)}\,\exp\,(p\sigma x_1 + \tfrac{1}{2}qx_1^2)\right\}\,I(\varrho_1), \qquad (7.6$$

where $x_1 = (\log \varrho_1)/\sigma$ and

$$I(\varrho_1) = \int_{-\infty}^{\infty} I(\varrho_1, \varrho_2)\,\varphi\,(y_2)\,\mathrm{d}y_2. \qquad (7.$$

This gives $I(\frac{1}{2}r_0) = 0{\cdot}1388$, $I(r_0) = 0{\cdot}1120$, $I(2r_0) = 0{\cdot}1425$. Th $I(r)$ does not change very quickly with $r$, and the mean number of conta points, like the mean number of tetrahedra, is mainly determined by t exponential factor in the tetrahedron probability function. We fi $T(\frac{1}{2}r_0) = 6{\cdot}8$; $T(r_0) = 12{\cdot}2$; $T(2r_0) = 35{\cdot}8$. These are at the percenta points 10·2, 68·0 and 98·6 respectively in the radius distribution by numbe of spheres. I have also interpolated for $I(r_0)$ by assuming that it is parabo in log $r$. The results are shown in fig. 3.

IPR2025-00217 EXHIBIT 2010 Page 341 of 485

If they are plotted on a logarithmic scale for radii, it seems that when $\varrho > 1{\cdot}6, T(\varrho)$ is about proportional to $\varrho^{1{\cdot}8}$. We might expect a $\varrho^2$ factor when $T$ is very large [5]). A sphere of unit radius (at the 50-per-cent point) touches on the average only 10 others, and only 34 per cent of the spheres touch more than the overall mean number 12.

It is interesting to compare this with the maximum possible number of touching spheres when they are all equal, but not equal to the middle one. Suppose, then, that all the outer radii are equal to the mean radius of $P(r)$. This mean is about 1·08. Or, given a number $T(\varrho)$ of equal spheres, we wish to know the least possible radius of the middle sphere they all touch. Schütte and Van der Waerden [6]) have estimated this for different $T$ (their $N$) very ingeniously. For $N$ unit touching spheres they find (in effect) for $N = 16$, 24 and 32, $\varrho = 1{\cdot}17$, 1·63, 2·02 respectively. Multiplying each $\varrho$ by 1·08, the corresponding $T$ (i.e., $T(1{\cdot}083\,\varrho)$) from our fig. 3 are 13, 21 and 30 respectively. Of course this comparison is only a rough one, and we should perhaps compare their $N$ also with $T$ for spheres with surface areas all equal to the mean surface area, i.e. with $T(1{\cdot}17\varrho)$. But it seems clear that the effect of the statistical distribution of touching spheres is, by comparison, to lower $T$ when these touch a small or average "central" sphere but perhaps to increase $T$ for a very large central sphere which has fewer very large spheres touching it than would be expected in a random sample.

(v) Surprisingly, the exact mean density cannot be calculated. A weighted average of the ratios $\Sigma/V$ of Table I cannot be far wrong as they vary so little. This average is almost exactly 0·8 (0·802). But strictly we should find the mean $\Sigma$, i.e. the mean over all tetrahedra of each sum of their four solid angles, and then the mean volume $V$ of a tetrahedron. Both $\Sigma$ and $V$ are nearly constant multiples $\vartheta_v$ of the cube of the geometric mean of the four radii given in Table I. So their means can be found by a transformation, similar to (5·6) and (5·7), of the multiple integrals, and now it is in terms of the ratio $\vartheta_v$ and, again, of tetrahedron weight factors $W_{p,q}$ — not for the tetrahedra with radii $\frac{1}{2}r_0$, $r_0$ and $2r_0$ ($r_0 = 1{\cdot}205$), but with radii 1·128 times these, i.e. the middle one is 1·366. If the $W_{p,q}$ are all the same as for the corresponding '$r_0$' tetrahedra, $\Sigma/V$ comes to 0·8044, fortunately very close to the other value.

But an interesting test shows that the $W_{p,q}$ values cannot be exactly the same. If they are, the mean volume of a sphere must be $N\Sigma/M$: this comes to 8·11. But its exact volume, cf. (8·3), is $(4\pi/3)\exp(9\sigma^2/2)$ or 8·543. So, on the average, when the radii are each increased by 13 per cent the $W_{p,q}$ are 5 per cent lower than their extrapolated values.

This might not be so with an improved probability distribution law for the tetrahedra; the most promising way to improve it seems to be to relate it more closely to the volume of a tetrahedron or to our $\vartheta_v$. For all the positively skew groups of radii give relatively large volumes, except for the doubtful one $\frac{1}{2}$, 1, 1, 2. (This is the only group for which $\log r$ and $r$ have different signs of skewness.) And of course a tetrahedron with zero or imaginary volume must have a weight factor zero.

## 8. The size distribution of the largest spheres that can go into the interstices

The problem of filling interstices has been thoroughly studied for regular packing [6]), [7]). There are two different kinds of interstice, inside either a regular tetrahedron or a regular octahedron. Hudson [6]) has calculated densities and radii of the small spheres when many different numbers of them are packed — either loosely or immovably — into these spaces. We only consider tetrahedral interstices of course, and only one sphere in the interstice, the largest possible one. There is in fact just one radius for such sphere that touches all the four externally, and the first step is to find

how to calculate it for a general group of 4 mutually touching spheres, or rather, for a group with not more than one equal pair in the 4. This is another interesting exercise in analytical solid geometry: a method will be given in a separate note and the results are in Table III.

## TABLE III

Radii of the interstitial spheres $r_5$ and their ratio $\vartheta$ to the geometric mean of the four outer spheres; $\vartheta = r_5(r_1^2 r_2 r_3)^{-\frac{1}{4}}$.

| Outer spheres $r_1, r_1, r_2, r_3$ | 1111 | $11\frac{1}{2}1$ | 1121 | $11\frac{1}{2}2$ | $\frac{1}{2}\frac{1}{2}11$ | 1122 | $\frac{1}{2}\frac{1}{2}12$ | $22\frac{1}{2}1$ | $\frac{1}{2}\frac{1}{2}1\frac{1}{2}$ |
|---|---|---|---|---|---|---|---|---|---|
| $10^3 r_5$ | 224·7 | 186 | 261 | 218 | 155 | 309 | 175 | 254 | 130·5 |
| $10^3\vartheta$ | 224·7 | 221 | 219·5 | 218 | 218·5 | 218·5 | 210 | 214 | 219·5 |
| Outer spheres $r_1, r_1, r_2, r_3$ | 2212 | $\frac{1}{2}\frac{1}{2}22$ | $\frac{1}{2}\frac{1}{2}2\frac{1}{2}$ | $22\frac{1}{2}2$ | $\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}$ | 2222 | | | |
| $10^3 r_5$ | 372 | 201 | 145 | 313·5 | 112·3 | 449·4 | | | |
| $10^3\vartheta$ | 221 | 201 | 205 | 221·5 | 224·7 | 224·7 | | | |

Thus for these fifteen tetrahedra the interstitial radius $r_5$ varies between 0·895 and 1·0 times the geometric mean of the four larger radii. This makes it easy to calculate the first few moments of $\log r_5$ and thus to get its approximate distribution. Writing $x_i = (1/\sigma)\log r_i$ (for the original distribution), the $n$th moment $\mu_n'$ of $\log r_5$ is given by

$$\mu_n' = \int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty \left\{\tfrac{1}{4}\sigma(x_1+x_2+x_3+x_4) + \log\vartheta\right\}^n w\,\mathrm{d}r_1\,\mathrm{d}r_2\,\mathrm{d}r_3\,\mathrm{d}r_4 . \qquad (8.1)$$

So for $n = 1$ to 5 the selected-ordinate formula for integration is certainly accurate enough according to condition (5·2). There is no need to give the details. The selected radii are the same, $\frac{1}{2}r_0$, $r_0$ and $2r_0$; $r_0 = 1{\cdot}205$ as before. The mean of $\log r_5$, written $\log \bar{r}_5$, is only 0·0124 greater than $\log(0{\cdot}2247\ r_0)$: *) thus $\bar{r}_5 = 0{\cdot}2742$, the standard deviation is just under half that of the coarse powder, viz. $\sigma_5 = 0{\cdot}178$; the third moment about the mean is $-1{\cdot}155\sigma_5^3$ and the fourth is $(3 + 0{\cdot}540)\sigma_5^4$. A good enough approximation to the distribution (except far in its tails) is then obtained from the Gaussian function $\varphi(x)$ with the same mean and $\sigma_5$, and its third and fourth derivatives. (Gram-Charlier series — Type A.) Then if $P_5(x)\mathrm{d}x$ is the probability that $x,=(1/\sigma_5)\log(r_5/\bar{r}_5)$, is between $x$ and $x+\mathrm{d}x$,

*) When the four radii are equal $r_5/r = -1 + \sqrt{3/2} = 0{\cdot}224749$ as given and used by Hudson[6]: $r\sqrt{3/2}$ is the circum-radius of a tetrahedron with edges $2r$.

$$P_5(x) = \varphi(x) - \frac{\mu_3}{\sigma^3}\frac{\varphi'''(x)}{3!} + \left(\frac{\mu_4}{\sigma^4} - 3\right)\frac{\varphi^{iv}(x)}{4!} =$$
$$= \varphi(x) + 0{\cdot}1924\,\varphi'''(x) + 0{\cdot}0225\,\varphi^{iv}(x). \qquad (8.2)$$

In this distribution the number of radii with extreme values is greater than in the corresponding log-normal distribution, but (from Table IV) the spread of the radii is still not large. Thus the fine-powder distribution overlaps the tail of the coarse one, but only very slightly.

TABLE IV

Frequency distributions for the largest interstitial spheres:
(i) The "type A" series of equation (8.2),
(ii) Log-normal distribution with the same mean and variance (of log $r_5$).

| (i) | | | (ii) | |
|---|---|---|---|---|
| Per cent $<r$ | | $r$ | Per cent $<r$ | |
| 0·48 | | 0·15 | 0·03 | |
| | 6·5 | | | 3·47 |
| 7·0 | | 0·2 | 3·5 | |
| | 22·0 | | | 25·5 |
| 29·0 | | 0·25 | 29·0 | |
| | 35·0 | | | 41·0 |
| 64·0 | | 0·3 | 70·0 | |
| | 30·7 | | | 21·7 |
| 94·7 | | 0·35 | 91·7 | |
| | 2·8 | | | 6·5 |
| 97·5 | | 0·4 | 98·2 | |
| | 1·3 | | | 1·55 |
| 98·8 | | 0·45 | 99·75 | |
| | 1·2 | | | 0·25 |

The percentages by volume (weight) are of interest. The mean volume per sphere in the fine powder is

$$\overline{\frac{4}{3}\pi r_5^3} = \frac{4}{3}\pi \bar{r}_5^3 \left(\exp\frac{9}{2}\sigma_5^2\right)\left\{1 + \frac{9}{2}\mu_3 + \frac{27}{8}(\mu_4 - 3\sigma_5^4)\right\}, \qquad (8.3)$$

and comes to $0{\cdot}973.\ 10^{-2}$. The mean volume of the coarse powder comes to 8·50. Thus with 5·10 times as many interstitial spheres as coarse ones, their total volume is 5·8 per cent of the volume of the coarse powder. This

is only about 30 per cent of the volume of the interstitial space. (In fact for four equal spheres the fifth fills exactly 30·8 per cent of this space.) Obviously this is because the shape of the space inside is so far from spherical.

Thus the density can be brought from 0·8 up to about 0·87, and to improve on this four times as many still smaller spheres — i.e. 20 $M$ of them — must be put into the new interstices. Horsfield [7]) has made calculations for a whole series of filling spheres of this kind, and Hudson [6]) has discussed it: again, for regular packing only.

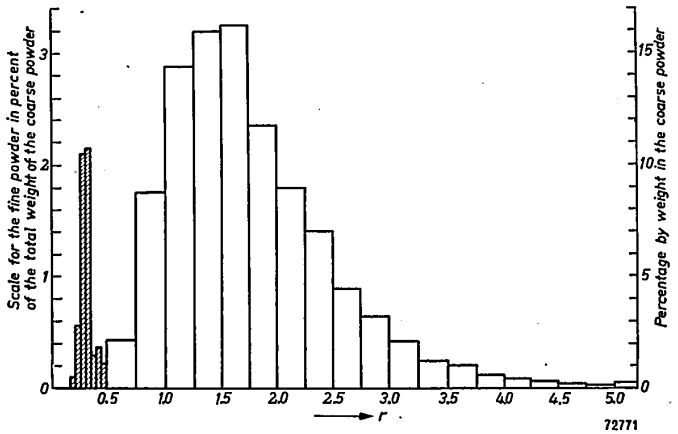



Fig. 4. Histogram of the main distribution and the corresponding distribution of largest interstitial spheres (shaded). Unit of area: 25 = weight of coarse powder (thus the ordinates in the coarse-powder histogram are percentages, and the ordinates for the fine powder are in a scale five times larger since the frequency interval for the fine powder is 0·05).

The two weight distributions (fig. 4) overlap even less than the number distributions: this is because the one for the coarse powder — because of its larger $\sigma$ — is shifted more to larger radii than the one for the fine powder.

## 9. Discussion

First it is useful to recapitulate briefly. We first obtained equations that have to be satisfied by all densely packed systems. They were conditions necessary for dense packing. It was not proved that they were sufficient. This part of argument was kept quite separate from the numerical application to a particular distribution.

However, far more of the text deals with this special case. Of course a method of handling the mathematics had to be found. We worked entirely

IPR2025-00217 EXHIBIT 2010 Page 345 of 485

with probability distributions, and the formulae for the multiple integrals are only suitable for a log-normal distribution, $P(r)$ for the radii. This might limit the range of usefulness of the results. But this is not quite true. First because a distribution that is only approximately Gaussian in log $r$ could be handled in the same way — our $W_{p,q}$ factors and/or the $\exp(p\sigma x_i + \frac{1}{2}qx_i^2)$ factor could just as well include terms correcting $P(r)$. Again, the fact that $P(r)$ is not bounded — thus log $r$ can be infinitely large or small — matters no more than it does in any other statistical applications of Gaussian functions, because for their extreme values they approach zero so rapidly. But this suggests a possibly more serious objection; a powder can always be sieved so that its particle size distribution extends over only a limited range; was it then only for mathematical convenience (or amusement!) that complete Gaussian functions were chosen? The answer is that dense random packing is impossible unless there is an appreciable spread in the distribution of radii. We came upon this point several times: it is clearly seen for equal spheres from Boerdijk's paper [1]) (and the papers referred to). Hudson makes another point as follows: that with a regular close-packed system with small spheres filling the interstices, only a few small spheres in the wrong places are enough to throw the whole system out, which is why it is experimentally very hard to get densities anywhere near the theoretical ones. But even with dense random packing and a statistical distribution it is quite possible that there is no such packing when its standard deviation (of log $r$) has (say) half the value of 0·4 we have chosen.

We have left out the results of very many trials with simple tetrahedron probability functions, which were tested in the "fitting-together" equations. These gave much negative evidence supporting the unexpected solution in terms of a skewness-law, although of course, they do not prove it is the only "simple" solution. But the law probably has physical significance: the largest spheres must be kept apart as much as possible, and it is then easier to complete the network of tetrahedra of centres of spheres round one large one if the touching spheres are nearly all much smaller. Perhaps it is more likely simply that the relatively large tetrahedra must be favoured (as we have suggested at the end of section 8).

The next question is whether this system could be built up in practice. Perhaps a sort of "Maxwell demon" or "giant brain" could do so by shifting the spheres about for long enough. But in the shaking and compressing of a powder it will often happen that a small ball of the right size needed to fill a gap (such as $G$ in fig. 1) is not in the right place to fill it. If the balls were elastic this could be done more often, especially as the model should be valid for balls not quite spherical, but this case seems less important. If the particles are imperfect spheres, that (as discussed)

IPR2025-00217 EXHIBIT 2010 Page 346 of 485

helps to fill the gaps; this is easy to verify experimentally if one makes the plasticine balls oneself! But probably there will always be a few gaps like $G$, and then the problem is to allow for them in the theoretical model.

We might still assume that in a tetrahedron no more than one pair of spheres do not touch each other. Even then five coordinates are needed to define it; the fifth is the width of the gap between this pair, and it has an unknown probability distribution. Yet it is not easy to obtain any more equations except by increasing the number of selected radii to four; this greatly increases the number of terms in every equation, now that they involve 4- and 5-dimensional integrals.

Two things might be tried. First, the gaps influence the density of the system so strongly that we might choose a density and then see what distributions if any will fit it. Secondly we might define randomness more precisely. The spheres round a large one ought to be as nearly as possible a random sample from the heap (subject to the restrictions we have discussed): to be more precise this may mean that the $W_{p,q}$ factors should have the minimum statistical variance. We should still have to guess at the probability distribution function for the gap width. Again, many of the interesting geometrical properties may still be estimated even if the guess is not too good.

All this is much more complex and that is a good reason for publishing our non-rigorous results. It is hoped that this new approach to the problem will at least suggest and help towards further advances. Such a statistical model, for all its limitations and difficulties, does seem to be more like a real densely-packed heap than other models.

**Acknowledgements**

This work was begun some time ago in the Materials Research Laboratory of Philips Electrical Ltd, Mitcham, Surrey, England, when I was in the group led by Dr A. L. G. Rees (in Melbourne, Australia, since 1944). I wish to thank him and the head of the laboratory Dr J. A. M. van Moll. In Eindhoven I have recalculated some of the earlier results and also made all the calculations on contact points. Two of these were suggested by Dr Boerdijk. Apart from this it was particularly interesting to compare the ideas, which we came to independently, on spheres touching one central sphere and on rings of spheres each touching a pair.

*Eindhoven, February 1952*

## REFERENCES


1) A. H. Boerdijk, Philips Res. Rep. **7**, 303-313, 1952; Thesis, Delft University, Stelling*) **25**, Delft, Waltman, 1951.
2) L. Fejes Tóth, Jahrb. Deutsch. Math. Verein. **53**, 66-68, 1943.
3) F. Supnick, Trans. Amer. Math. Soc. **65**, 14-26, 1949.
4) E. Whittaker and G. Robinson, Calculus of Observations, Blackie, 1944, 4th Edition, p. 163.
5) K. Schütte and B. L. Van der Waerden, Math. Ann. **123**, 96-124, 1951.
6) D. R. Hudson, J. Appl. Phys. **20**, 155-162, 1949.
7) H. T. Horsfield, J. Soc. Chem. Ind. **53**, 107T, 1934.
8) C. A. Rogers, Act. Math. Uppsala **86**, 309-321, 1951.
9) L. Fejes Tóth, Proc. Amer. Math. Soc. **1**, 806-812, 1950.


*) One of a number of extra propositions published with the thesis.

IPR2025-00217 EXHIBIT 2010 Page 348 of 485

Philips Res. Rep. 7, 344-400, 1952

# KINETICS OF THE EMULSION POLYMERIZATION OF STYRENE *)

by J. BAKKER

541.127:542.952.623:547.538.141

**Summary**

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 electric 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 $x^{1/4}(1-x)^{-1/4} = t/t_{1/4}$ independently of the amounts of soap or catalyst used, provided the emulsifier is K myristate ($x$ = degree of conversion, $t$ = time, $t_{1/4}$ = 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

$$\frac{1-\varphi}{k}(x-\tfrac{1}{2}) - \frac{\varphi}{k}\ln[2(1-x)] + 1 = t/t_{1/4}$$

where $k$ probably depends on the nature of the soap, but not upon the amounts of soap and catalyst; $\varphi$ is specific for the monomer. With K myristate as emulsifier we found for $k$ a value of $1{\cdot}10 \pm 0{\cdot}07$. 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 $t_{1/4}$ bears a relation to the amounts of soap ($S$) and catalyst ($R$) as $t_{1/4} = \text{const.}\ (SR)^{-1/4}$. 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 $t^* = \text{const.}\ R^{-3/4}$ or, after correction of the values found for $t^*$ on account of theoretical considerations, as $t^* = \text{const.}\ R^{-3/4}$. This applies only as long as the amounts of soap used are less than 15 % by weight of the amount of monomer. When this limit is exceeded the micelles no longer disappear entirely from the solution. From considerations relating to the initiation theory of Smith and Ewart we found another method for determining the number of latex particles, which, combined with the first method, provided a means of calculating the area $\omega_0$ 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 $t_{1/4}$ 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 $t_{1/4}$ is independent of $V$ but directly proportional to $m_0^{1/4}$, within the range covered by our experiments. Further experiments are required before any conclusion can be drawn as to the dependency of $t_{1/4}$ 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 could be reconciled with the conceptions of the termination reaction in the case of emulsion polymerization. In none of our experiments modifier has been used.

*) Thesis, University of Utrecht, July 1952.

IPR2025-00217 EXHIBIT 2010 Page 349 of 485

**Résumé**

La vitesse de polymérisation en émulsion du styrène fut étudiée avec une méthode dilatométrique. En mesurant simultanément le changement de la conductibilité électrique, la fin de la période d'incubation fut déterminée. Les courbes représentant la conversion en fonction du temps étaient sigmoïdes. La difficulté d'une détermination exacte du commencement de la réaction causa une erreur d'environ 5 %. La fonction mentionnée ci-dessus peut être représentée par

$$\left(\frac{x}{1-x}\right)^{1/2} = t/t_{1/2}$$

jusqu'à 50 % de conversion, indépendamment des quantités de savon ou de catalyseur, pourvu qu'on utilise du myristate potassique comme émulsifiant ($x$ est le dégré de conversion, $t$ le temps et $t_{1/2}$ le temps du demi conversion). L'usage d'autres savons, soit laurates, stéarates ou oléates, montre des petites déviations. Au-dessus de 50% de conversion, la fonction peut être représentée par la relation

$$\frac{1-\varphi}{k}(x-\tfrac{1}{2}) - \frac{\varphi}{k}\ln[2(1-x)] + 1 = t/t_{1/2}$$

dans laquelle $k$ ne dépend pas des quantités de savon ou de catalyseur mais probablement du type de savon; $\varphi$ est une valeur spécifique de la monomère. Avec le myristate potassique nous avons obtenu une valeur de 1,10 $\pm$ 0,07 pour $k$. Ce paramètre se rapporte au nombre des particules de latex; par conséquent, on peut calculer ce nombre. Avec des quantités invariables de monomère et d'eau, $t_{1/2}$ dépend des quantités de savon ($S$) et de catalyseur ($R$) comme suit: $t_{1/2} = \text{const.}\ (RS)^{-1/5}$. Le moment où la période d'incubation est finie ($t^*$) ne dépend pas de la quantité de savon, mais comme $t^* = \text{const.}\ R^{-3/5}$ de la quantité de catalyseur, ou après correction des valeurs de $t^*$ expérimentales basée sur des considérations théoriques, comme $t^* = \text{const.}\ R^{-2/5}$. Cette fonction ne peut être appliquée que si le poids du savon ne dépasse pas 15 % du poids de la monomère. En passant cette limite les micelles ne disparaissent pas complètement de la solution. Des considérations basées sur la théorie d'initiation de Smith et Ewart menèrent à une deuxième méthode pour déterminer le nombre des particules de latex. Une combinaison de cette méthode avec la méthode précédente nous procura un moyen de calculer la surface $\omega_0$ occupée par une molécule de savon. Pour le myristate potassique nous trouvâmes: $\omega_0 = 35 \pm 6$ Å$^2$. Le rapport des propriétés des courbes de conversion et ceux de $t_{1/2}$ nous apprit que $t_{1/2}$, outre de quelques qualités spécifiques du savon, doit aussi dépendre de $m_0$, le poids de la monomère, et de $V$, le volume de la phase aqueuse. Expérimentalement, $t_{1/2}$ se trouvait être indépendant de $V$, mais proportionnel à $m_0^{1/5}$. D'autres expériences seront nécessaires avant qu'on puisse tirer des conclusions quant aux propiétés de $t_{1/2}$ hors de la gamme des expériences décrites ici. Quelques déterminations du dégré de polymérisation nous apprîmes que ce dégré fut entièrement établi par transfert de chaîne. Il était possible d'expliquer ce phénomène en rapport avec le mécanisme de la terminaison des chaînes lors de la polymérisation d'émulsion. Pendant les expériences décrites on n'a jamais utilisé de régulateurs.

**Zusammenfassung**

Mit Hilfe einer dilatometrischen Methode wurde die Polymerisationsgeschwindigkeit von Styrol in Emulsion gemessen. Gleichzeitig wurde aus der Änderung der elektrischen Leitfähigkeit das Ende der Inkubationsperiode festgestellt. Die Kurven, die den Umsatz als Funktion der Zeit wiedergeben, sind $S$-förmig. Die Genauigkeit der

Meßergebnisse betrug etwa 5%, hauptsächlich wegen der Schwierigkeit, den Zeitpunkt des Reaktionsanfanges zu ermitteln. Bis 50 % Umsatz konnte diese Abhängigkeit durch

$$\left(\frac{x}{1-x}\right)^{1/3} = t/t_{1/3}$$

wiedergegeben werden, und zwar unabhängig von den angewendeten Mengen Seife oder Katalysator, falls K-Myristat als Emulgator verwendet wurde ($x$ = Umsatz, $t$ = Zeit, $t_{1/3}$ = Halbwertzeit). Bei Anwendung anderer Seifen — Laurat, Stearat oder Oleat — ergaben sich kleine Abweichungen. Oberhalb 50% Umsatz konnte die erwähnte Funktion durch

$$\frac{1-\varphi}{k}(x-\tfrac{1}{2}) - \frac{\varphi}{k}\ln[2(1-x)] + 1 = t/t_{1/3}$$

wiedergegeben werden, wobei $k$ nicht von den Mengen Seife oder Katalysator, dagegen vermutlich wohl von der Seifenart abhängt; $\varphi$ ist eine Kennzahl des Monomeres. Mit K-Myristat erhielten wir für $k$ den Wert 1,10 $\pm$ 0,07. Da dieser Parameter mit der Anzahl Latexteilchen zusammenhängt, konnten wir diese berechnen. Die Halbwertzeit $t_{1/3}$ ist bei konstanten Mengen Monomer und Wasser in Abhängigkeit von den Mengen Seife ($S$) und Katalysator ($R$) durch $t_{1/3}$ = Konst. $(RS)^{-1/3}$ gegeben. Der Zeitpunkt, in dem die Inkubationsperiode endet, war praktisch unabhängig von der Seifenmenge, während er folgendermaßen von der Katalysatormenge abhängt: $t^*$ = Konst. $R^{-1/3}$ oder nach Korrektion der gefundene $t^*$ Werten auf Grund theoretischer Erwägungen: $t^*$ = Konst. $R^{-1/3}$. Dies galt nur, wenn das Gewicht der angewendeten Seife weniger als 15 % vom Gewichte des angewendeten Monomeres betrug. Bei Überschreitung dieser Grenze verschwanden die Mizellen nicht mehr vollständig aus der Lösung. Aus Erwägungen an Hand der Initiationstheorie von Smith und Ewart ergab sich eine zweite Methode zur Bestimmung der Anzahl Latexteilchen. Die Kombination dieser Methode mit der ersten ermöglichte es, die durch ein Seifenmolekül beanspruchte Oberfläche $\omega_0$ zu berechnen. Für K-Myristat wurde $\omega_0 = 35 \pm 6$Å$^2$ erhalten. Betrachtungen über den Zusammenhang zwischen der Reduzierbarkeit der Konversionskurven und den Eigenschaften der Halbwertzeit lehrten, daß $t_{1/3}$ außer von einigen spezifischen Eigenschaften der Seife auch noch von dem Volum $V$ der wäßrigen Lösung und vom Ausgangsgewicht des Monomers $m_0$ abhängen muß. Experimentell ergab sich jedoch daß $t_{1/3}$ innerhalb des untersuchten Gebietes unabhängig von $V$ und verhältnisgleich $m_0^{1/3}$ ist. Die Abhängigkeit außerhalb dieses Gebietes wurde nicht näher untersucht. Einige Messungen ergaben, daß der Polymerisationsgrad vollständig durch Kettenübertragung bestimmt wird. Es gelang, dies mit den Auffassungen, die auf dem Gebiete der Terminationsreaktion bei Emulsions-Polymerisationen bestehen, in Einklang zu bringen. Bei keinem dieser Experimente wurde ein Regler angewendet.

## CONTENTS

I. INTRODUCTION
II. EXPERIMENTAL PART
1. Description of the dilatometric method.
2. Description of the electrical method of determining the disappearance of the soap micelles.

3. Results obtained with the dilatometric method.
4. Results of the conductivity measurements.

II. DISCUSSION

1. The stationary state
2. The number of latex particles.
3. Length of the incubation period.
4. The initial stages of the reaction.
5. Influence of the initial amount of monomer and of the volume of the aqueous phase on $t_{1/2}$.
6. Comparison with data taken from the literature.

List of symbols.
References.

## . INTRODUCTION

Although the art of making high polymeric substances known under the ame of emulsion polymerization [1] [2] was developed a long time ago, it s only in recent years that a closer insight has been obtained into the ature of the reactions taking place in such a process. This is by no means ifficult to understand, as the systems used are of a very complex compoition in so far that they are not only built up of the monomers, diluent nd initiator, but also contain a detergent, the latter being indispensible or the progress of the reaction, which, in comparison with other processes, apidly produces very-high-molecular polymers even at moderate or low emperatures.

We are indebted most to Harkins [3] [4] and his coworkers for unraveling many of the mysteries involved in these reaction systems. Further aluable contributions have been given by Smith and Ewart [5] and by Haward [6].

In accordance with the general opinion we now suppose that the deterent used in these systems not only serves for making an emulsion of the nsoluble monomer but also forms micelles in the solution. These micelles ontain some monomer, a phenomenon called solubilization. It is interesting to note that as a consequence the micelles play a prominent role in the rocess of the polymerization, for the initiator dissolved in the water deomposes into free radicals and these penetrate the micelles where they tart the polymerization of the monomer. It is not improbable that the ree radicals first react with the monomer dissolved in the water phase but re prevented from growing into a long chain as a consequence of the very ow concentration of the monomer present in the water, and the radical, nce captured by a micelle, will grow much faster, since there the monomer

concentration is much more favourable. It being difficult to imagine tha the micelles are by their nature more apt to capture free radicals tha the emulsion particles, the preference of the former to act as a reactio locus must be attributed to the much larger interfacial area which the exhibit as compared with that of the latter. The reaction taking plac in the micelles forces them to grow in volume and to lose their origina structure. To compensate for the lack of detergent in their interfac which now becomes larger and larger, non-reacting micelles are split u and the liberated detergent molecules may move to the places wher they are needed. In the meantime the amount of monomer originall solubilized in the growing micelles will be used up, but this is restore by diffusion of monomer from the relatively large emulsion particles t the minute growing micelles, in the following referred to as latex particles In these latex particles we thus have a mixture of polymer chains an of monomer. In this picture the emulsion particles only serve as a stoc of pure monomer.

From this brief description we draw the conclusion that in the first stage o the conversion of monomer into polymer our system is built up of two parts namely an aqueous phase, containing in solution the initiator, some mono mer and moreover the micelles — some of them growing out to latex particle — and another phase consisting of emulsified monomer. Neither the micelle nor the emulsified monomer persists during the whole time of the con version: in most cases there will first be a moment when the non-reacte micelles have disappeared, consumed by the growing latex particles i their demand for detergent. This leads to the second stage, where a momen is arrived at when also the emulsified monomer is used up as a consequenc of the transport of monomer from this phase to the latex particles. Thi transport is largely governed by the difference in activity of the pur monomer in the emulsion and that of the monomer mixed with polyme in the latex. The third stage thus becomes rather simple, as the system onl consists of an aqueous phase and latex particles.

Several experiments have led to the conclusion that under norma conditions the emulsified phase ceases to exist somewhere about halfwa through the conversion [3) 7) 8)]. Microscopical examination proves to b a reliable method of determining whether the emulsion droplets ar present or not. Other methods are based on the separation of the emulsifie monomer phase by centrifugation, or on the determination of the vapou pressure of the monomer.

The moment of disappearance of the micelles has so far been determine by two methods. The first of these consists in measuring the surface tension [3)] which is found to increase as soon as the micelles disappear. The secon method finds its origin in the sudden decrease of the fluorescence of som

particular dye-stuffs upon the disappearance of the micelles [9]). At the end of the year 1947, Professor Overbeek of Utrecht suggested to us that it should be possible to detect the disappearance of the micelles by measuring the electric conductivity of the reaction mixture. As a matter of fact the electric conductivity of colloidal particles is highly dependent on their dimensions. The small micelles deriving their dimensions from the roughly one hundred molecules of soap of which they exist, pass into a surface layer round the latex particles, which are about 20 times as large in diameter, and this must therefore have a marked influence on the conductivity of the solution. This method was first tried in the emulsion polymerization of styrene stabilized with potassium stearate as detergent and gave a promising result [10]).

These determinations gain much more importance when they are considered in connection with the kinetic behaviour of these polymerizations. There has been no lack of investigations concerning the emulsion polymerization of styrene, but the conclusions that can be drawn from the material available are not very satisfying. The majority of investigators share the opinion that in the beginning the reaction proceeds linearly with time (interpreted as zeroth-order reaction [11])) until a conversion of about 50 % is reached. After that the rate of reaction slows down more and more. Nevertheless one often finds conversion curves that almost entirely lack the linear part and are described as "*S*-shaped". Under certain limiting conditions Kolthoff and Bovey [12]) found something between a first and a second-order reaction when styrene was solubilized with potassium palmitate or stearate. The slowing-down part of the reaction was considered by Corrin [13]) to be identical with bulk polymerization and therefore of order 3/2, but his arguments were hardly convincing [2]). So far as the author knows, nobody has tried to analyze the properties of an *S*-shaped conversion curve.

At this point it is of importance to say something about the experience obtained in this laboratory with respect to conversion rates. About five years ago Klaassens [14]) and co-workers decided to determine the rates of reaction of the emulsion polymerization of styrene in dilute systems following two lines of investigation, first by a gravimetric determination of the amount of polymer formed during the reaction, and secondly by a volumetric determination of the amount of monomer left. The two methods only gave identical results after the reaction had already proceeded for 50 %: before that stage of conversion was reached the first method gave, after a certain induction period, a sudden start and linear progress of the reaction up to about 50% conversion, while the second method showed no induction time whatever and proceeded, from a slow start, faster and faster till at about 25% conversion it reached its maximum velocity and

then slowed down continuously (fig. 1). Thus it became apparent that th two aspects of the conversion curve referred to depend to a great exten upon the method of investigation, and in order to decide which of th methods gives the most reliable results a third method, namely a dilatom etric one, was tried. This method has the advantage that it is no longe necessary to extract samples for analysis, and furthermore it is easil combined with the electrical method for determining the disappearanc of the micelles. The shapes of the conversion curves obtained by th dilatometric method proved to be in conformity with those obtained b the volumetric method.

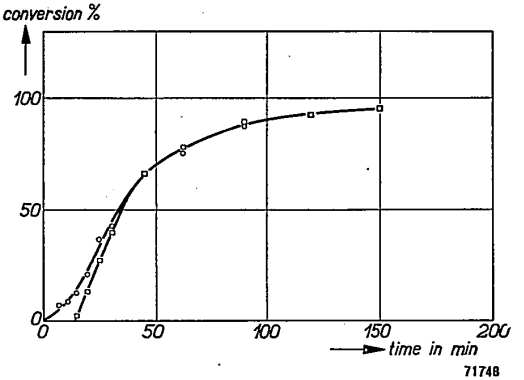



Fig. 1. Conversion-time curves of an emulsion polymerization of 50 g styrene with 4 g potassium stearate and 1 g potassium persulphate in 450 ml water. Temperature 40 °C.
o Determinations by means of the volumetric method.
□ Determinations by means of the gravimetric method.

The programme covered by our experiments, with styrene as sole monomer, was such that information was obtained concerning:

(a) the real shape of the conversion curve and how it is influenced by varying conditions as regards the quantities of soap and catalyst applied;
(b) the moment at which the micelles disappear under these varying conditions.

From the data obtained parameters suitable for a simple characterization of the various results were derived. An attempt was made to investigate whether these results are in conformity with the theories of Smith and Ewart and of Haward. A survey of these theories is given in the sections where the conversion and the other data are discussed from the standpoints suggested by the authors mentioned.

## II. EXPERIMENTAL PART

### 1. Description of the dilatometric method

For measuring the conversion of styrene in emulsion polymerization systems we chose the dilatometric method [2) 15) 16)], as being the most convenient for our purposes. This method is based upon the fact that the specific volumes of monomer and corresponding polymer differ sufficiently to make it possible to determine from the contraction of the original volume of the monomer the extent to which the conversion has proceeded. Of course an important condition underlying this principle is that no change in volume takes place when mixing monomer with polymer. As far as experience goes this condition is always sufficiently fulfilled to make the measurements reliable. The method has been specially tried out with regard to the emulsion polymerization of styrene in Harkins's laboratory [13]).

The main vessel of the apparatus used (fig. 2), which can hold a volume of 450 ml, is fitted with an all-glass stirrer (*a*) at the top of a somewhat enlarged compartment (*b*) that is sealed to the calibrated neck (*c*) of the flask. The compartment (*b*) can be connected with one of the legs of a simple, manometer-like, instrument with which we are able to exert a slight pressure or suction on the liquid in the vessel. With a ground-glass joint a calibrated tube (*d*) — with subdivisions of 0·1 ml — is put on the vessel. The top of this tube has a ground opening (*e*), to take the ground male joint of the filling apparatus. The two small spouts (*f*) serve for passing a small stream of nitrogen through the outlet of the tube, thus preventing penetration of air. Finally the vessel contains two openings in the wall, with ground-glass joints to insert the conductivity cell (*h*) and the thermometer (*i*). The latter consists of a N(egative) T(emperature) C(oefficient) resistor perfectly suited for measuring any temperature variations in the system without appreciable delay. The apparatus was for the greater part immersed in a bath kept at a temperature of 40·0 ± 0·1 °C.

Several precautions were taken to prevent contamination of our reaction system with air, as oxygen is a strong inhibitor of the polymerization reaction [17]). The nitrogen used for replacing the air driven out from the apparatus was specially purified by our factory, and guaranteed to have a maximum oxygen content of less than 1 part per million. During the many polymerizations that were carried out no higher degree of purification was ever found to be necessary. In connection herewith we refer to a publication by Shunmukham c.s. [18]) who found different polymerization velocities according to whether their systems were agitated continuously or not, which can easily be accounted for by their having used nitrogen contaminated with oxygen [19]).

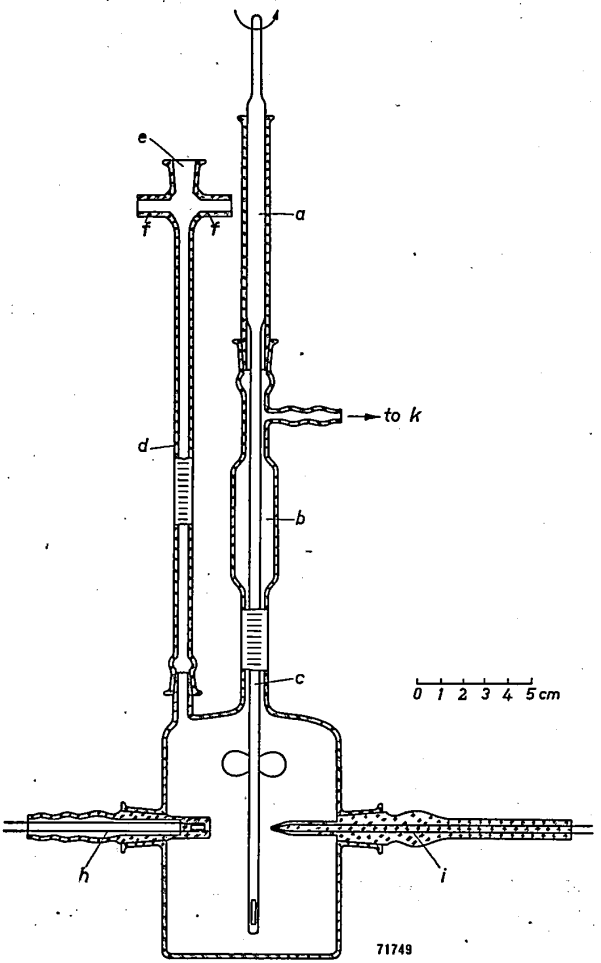



Fig. 2. The dilatometric apparatus.

The recipes used were as follows

50 grams of styrene; in some cases we used about 25 grams;

2·5 g to 15 g of soap; this was prepared by dissolving an amount of

fatty acid in the styrene, 5% more than needed for the preparation of the required amount of soap *); in most of the cases myristic acid was used but in some cases stearic, oleic or lauric acid;
potassium hydroxide solution in an amount sufficient to neutralize 95% of the fatty acid;
0·3 to 1·0 g potassium persulphate;
400 ml water (in some cases about 200 ml) **).
Modifier was never used.

The styrene, after the usual purification, was distilled in nitrogen atmosphere under reduced pressure into a vessel connected by a ground-glass joint to the condensor and already containing a weighed portion of fatty acid. In the meantime the reaction vessel was filled with an adequate quantity of potassium hydroxide solution and enough oxygen-free water as to make the contents of the vessel nearly 400 ml. During an hour a strong stream of nitrogen is blown over the vigorously stirred solution, and at the same time the vessel is brought to a temperature of 40 °C by immersion in the bath. The container with styrene and fatty acid is brought to the same temperature, together with another container consisting of two compartments, one of which holds the persulphate dissolved in 30 ml water while the other contains pure oxygen-free water. All these vessels are blown through with nitrogen while being brought to the required temperature. When that temperature has been reached the container with the styrene is connected by means of its ground-glass joint to the female joint (*e*) of the dilatometer tube and its contents are transferred under nitrogen pressure to the reaction flask, where an emulsion is formed by vigorous agitation. Some minutes later the second container is placed on the dilatometer and the persulphate solution is transferred as quickly as possible to the agitated emulsion. This is followed by the addition of water in such a quantity that, with the help of appropriate manipulations of the "manometer", first the liquid in the dilatometer tube is intimately mixed with the rest of the emulsion, and secondly a fixed point can be

*) To prepare the emulsions we chose the "nascent soap" method followed by Price and Adams [11]). Davis and Bartell [28]) have pointed out that the stabilization of the emulsion is aided by an excess of fatty acid. Our own experience shows that a slight excess of fatty acid, especially in the cases where stearic acid was used, gave less variable conductivities during the induction period of the polymerization. Moreover McBain [29]) has shown that sometimes an excess of fatty acid increases the amount of solubilized monomer, which fact may offer an explanation of the promoting action that an excess of fatty acid has on the rate of reaction of the polymerization found by Price and Adams.

**)The fatty acids were crystallized several times from ethyl alcohol and tested by determining the melting-point and acid number for their purity. The stock potassium hydroxide solution was kept free of carbonate, the distilled water was boiled to remove the atmospheric gases and cooled down under nitrogen. Potassium persulphate was crystallized twice from distilled water and dried in vacuo.

chosen for the level of the liquid on the calibrated neck of the flask, the dilatometer tube being filled to the uppermost calibration line. Naturally the last operations are to be performed as quickly as possible, because the reaction may start as soon as the first drop of persulphate solution reaches the emulsion. However, this seldom happens, probably as a consequence of the introduction of a little air into the mixture, which causes an induction period. In most cases this period did not exceed 15 minutes; sometimes it lasted much longer, but we never found it to exercise any influence upon the kinetic behaviour of the ensuing reaction. From the moment the persulphate was added the dilatometer was read every two minutes, the agitator being stopped for a moment and every time with the help of the "manometer" bringing the level of the liquid in the neck of the flask to the height of the fixed point chosen.

The accuracy of the method has a number of limitations, to be discussed below. The dilatometer tube was graduated in divisions of 0·1 ml, such that readings of 0·01 ml could easily be taken. On the other hand the accuracy with which the fixed point on the neck of the reaction vessel can be adjusted during the repeated measurements was about 0·02 ml. With 50 g of styrene the contraction for complete conversion is slightly less than 9 ml so that, were it not for the occurrence of some disturbances, we might have been able to determine a conversion of about 0·3%. The first disturbance, which was the most troublesome, is caused by the formation of foam on the surface of the liquid and limits the accuracy of the readings to about 0·05 ml. The second disturbance, by far more serious, is a consequence of the generation of the heat of polymerization. Within the first 40% of the conversion a gradual temperature rise is to be observed, which, according to the rate of conversion, in exceptional cases amounted to slightly less than one centigrade. At higher conversions the temperature gradually falls till at about 90% conversion it reaches its original value again. The temperature rise in the beginning causes an expansion of the system which is superimposed on the contraction otherwise caused by the reaction. Consequently the reaction rate seems to be less than it really is. Beyond half conversion the situation is the reverse since the falling temperature increases the contraction that would otherwise be observed. The highest deviation in the contraction to be found as a consequence of the temperature rise can easily be calculated from the volume of the water, taken to be 400 ml, and the expansion coefficient of water at 40 °C, being $4.10^{-4}$/°C. We found 0·15 ml corresponding to a deviation of 1·6% in the extent of the conversion. Consequently the most important parameter, the time of half conversion $t_{1/2}$ will be found at most 3 per cent too high. It was not deemed necessary to correct the values of $t_{1/2}$ found for this error, because the influence of the accidental errors proved to be about 5%.

## 2. Description of the electrical method of determining the disappearance of the soap micelles

In the introduction we have already pointed out that in the course of time two methods have been developed for determining the moment at which the soap micelles have disappeared from the solution. Both methods have the disadvantage of being rather cumbersome, and Professor Overbeek's proposal to try a method based on a change of the electric conductivity was heartily welcomed. Before proceeding to describe this method and the apparatus used, it may be well to outline roughly the considerations on which our expectations are based.

Solutions of colloidal particles as a rule exhibit electrolytic conductivities which depend not only upon the electrokinetic or $\zeta$-potential of the particles but also upon their dimensions. Henry [21]) gave for the electrophoretic velocity of insulating spherical particles the following expression

$$u = \frac{\varepsilon E \zeta}{6\pi\eta} f(\varkappa a),$$

where $u$ = the electrophoretic velocity of the particles,
$\varepsilon$ = dielectric constant of the medium,
$E$ = strength of the external electric field,
$\zeta$ = electrokinetic potential,
$\eta$ = viscosity of the medium,
$\varkappa$ = reciprocal extension of the double layer,
$a$ = radius of the particles.

$f(\varkappa a)$ was calculated, taking only into account the influence of the electrophoretic retardation. Overbeek [20]) and lateron Booth [34]) have extended these calculations reckoning also with the relaxation effect. The function $f(\varkappa a)$ then depends also upon $\zeta$.

In our case it is probably not too bold to assume that $\zeta$ will not vary very much with the transformation of the micelles into latex particles. Moreover $\eta$ is assumed to remain constant during this reaction. Therefore a change is only to be expected in $u$ as a consequence of a change in $\varkappa a$. While the total charge of the particles remains the same, we may expect a change in the electric conductivity of the systems too.

A tentative calculation may give an insight into the magnitude of the difference between the electrophoretic velocity of the micelle compared with that of a latex particle. For the concentrations of soap and electrolyte (catalyst) used we find for $\varkappa$ on the average a value of $5.10^6$ cm$^{-1}$, using the well-known Debye-Hückel formula. In the main we may assume that the radius of a micelle will be of the order of 30 Å, whereas the radius of a latex particle will be about 20 times as large, so for the micelle $\varkappa a = 1{\cdot}5$

IPR2025-00217 EXHIBIT 2010 Page 360 of 485

and for the latex particle $\varkappa a = 30$. Taking $\zeta = 50$ mV, then from tables of $f(\varkappa a,\zeta)$, computed by Overbeek, we find that

$$u_{\text{micelle}} : u_{\text{latex particle}} = 1{\cdot}1 : 1{\cdot}4.$$

This effect of "$a$" upon the electrophoretic velocity should be reflected in a higher conductivity of the system. But the contribution of the counterions of the double layer to the conductivity will certainly also be influenced by the changing conditions. This may be such that the increase of the conductivity caused by the higher velocity of the latex particles is not only diminished but even reversed in sign. Possibly this accounts for the fact that the conversions with 2·5 g myristate showed an increase of the resistance, whereas in all the cases of conversions with higher amounts of soap there was a decrease of the resistance.

The measurements were not performed in the usual way with the help of a Wheatstone bridge. From disturbances encountered in the beginning of the experiments it was concluded, wrongly as it subsequently appeared, that the cell capacity was not constant, owing perhaps to the action of soap or polystyrene films deposited on the platinum electrodes. To avoid the necessity of compensating for this "varying" cell capacity in each measurement — which was carried out every two minutes — Dr Zaalberg van Zelst of this laboratory devised the following network, with which all our measurements were carried out. A diagram is shown in fig. 3. A capacitor ($C$) of about 10 $\mu$F, one terminal grounded, the other leading via a rheostat *) to one of the secondary terminals of a transformer, forms together with the rheostat the first branch of the network. The second branch is formed

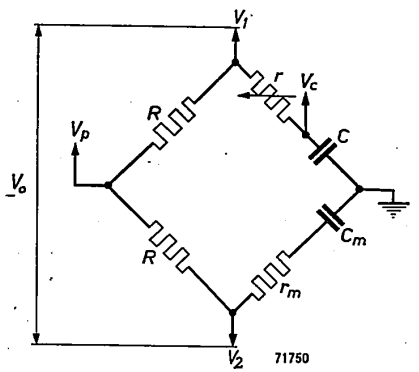



Fig. 3. Network for determining the conductivity.

*) This was a resistor box composed of five decades, the lowest one divided into steps of 0·1 Ω each, and produced by Nedoptifa, Utrecht, Netherlands.

by the conductivity cell ($r_m - C_m$), also one end grounded, the other connected to the other secondary terminal. Moreover two equal branches — each consisting of an induction-free resistor ($R$) of about 5000 Ω and of negligible capacity — are likewise connected between the secondary terminals of the transformer. Now all we have to observe is the phase shift between the voltage $V_c$ on the non-grounded terminal of the capacitor $C$ and the voltage $V_p$ on the junction between the two equal resistors $R$. When we adjust the resistance of the rheostat such that the shift becomes zero, $r$ equals $r_m$, regardless of the values of $C$ or $C_m$.

This is easily demonstrated in the following way. Let the current through the branch containing the capacitor and the rheostat be $i$; then the voltage $V_c$ will be equal to

$$V_c = \frac{i}{j\omega C},$$

where $\omega/2\pi$ is the frequency of the current. The voltage on the junction between the rheostat and one of the resistors $R$ will be

$$V_1 = i\left(r + \frac{1}{j\omega C}\right),$$

while the voltage on the junction between the non-earthed pole of the conductivity cell and the other resistor $R$ is

$$V_2 = -i\left(r_m + \frac{1}{j\omega C_m}\right).$$

From $V_1$ and $V_2$ we deduce the value of $V_p$

$$\begin{aligned} V_p &= \tfrac{1}{2}(V_1 + V_2) \\ &= \tfrac{1}{2} i\left(r - r_m + \frac{1}{j\omega C} - \frac{1}{j\omega C_m}\right). \end{aligned}$$

As soon as $r = r_m$ the voltages $V_c$ and $V_p$ will be mutually in phase.

The frequency of the current in our experiments was about 1000 c/s; the input was supplied by a Philips standard-signal generator GM 2307 connected to the primary terminals of the transformer. As a phase discriminator we used a cathode-ray oscilloscope, Philips GM 5655. $V_c$ was applied to the horizontal plates and $V_p$ to the vertical ones, each after separate amplification of about 30 times. The oscilloscope itself gave an additional gain of 500 times. Thus the total amplification was about 15000 times, which enabled us to limit the voltage on the cell to about 70 mV. As long as $r$ does not equal $r_m$, the oscilloscope shows an ellipse which degenerates into a straight line as soon as $r = r_m$.

The elements of the apparatus were so designed that we could detect a variation of 0·1 Ω in the resistance. This was sufficiently accurate for our work, where the resistances were of the order of some hundreds of ohms and the variations due to the reaction amounted to roughly 10% of the initial values.

The effect of the temperature rise, mentioned in the previous section, tends to augment the variation observed. For one degree of temperature rise generally diminishes the resistance by about 2%, a value verified by carrying out an experiment where styrene was replaced by ethyl benzene and the temperature of the bath varied.

The electrodes of the conductivity cell were not platinized. The cell constant was about 1·5.

### 3. Results obtained with the dilatometric method

The conversion $x$ was calculated with the aid of the formula

$$x = \Delta V/(\varphi_m - \varphi_p)m_0,$$

where $\Delta V$ is the contraction observed in the system, $m_0$ is the initial weight of monomer in grams, $\varphi_m$ and $\varphi_p$ are the respective specific volumes in ml per gram monomer and polymer. For $\varphi_m$ and $\varphi_p$ we used the values $\varphi_m = 1{\cdot}125$ ml.g$^{-1}$ [22]) and $\varphi_p = 0{\cdot}946$ ml.g$^{-1}$ [23]) at 40 °C so that $\varphi_m - \varphi_p = 0{\cdot}179$ ml.g$^{-1}$.

$\Delta V$ attains its highest value at $x = 1$; this was repeatedly checked by letting the reaction mixture stand overnight and in every case we found the theoretical value confirmed within an accuracy of a few per cent.

The great difficulty remaining concerns the determination of the moment at which the reaction actually starts, i.e. our "zero point". In all the cases where the reactions are of a linear character this causes no trouble, but none of the conversions measured by us proceeded linearly.

To determine the zero point we thus have to extrapolate the curve found, a procedure which is apt to introduce an uncertainty of several minutes. Having regard to the rate of conversion — generally 90 % was reached within one hour and a half — it was desired to avoid this discrepancy. Fortunately the characteristics found in the measurement of the electrolytic conductivity provided a means of limiting this discrepancy to a few minutes.

The first series of experiments, of which all the conversion data are collected in table I, were performed with potassium myristate as emulsifying agent in varying quantities, and potassium persulphate in varying amounts too. The amount of styrene was kept constant at 50 ± 0·5 g and a quantity of water was added such as to make the total volume 450 ± 1 ml. The quantities used and the observed induction times $t_i$ are given for each experiment. Bearing in mind that the zero point of the time scale coincides

| $S$ | 2·5 | | | | 3·0 | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $R$ | 0·30 | | | | 0·30 | | | | 0·65 | | | | 1·0 | | | |
| $m_0$ | 49·6 | 50·3 | | | 49·1 | 50·0 | | | 49·1 | 49·3 | | | 49·8 | 48·5 | | |
| $t_i$ | 0 | 0 | Mean | | 22 | 72 | Mean | | 12 | 2 | Mean | | 6 | 8 | Mean | |
| $x$ | $t$ | | $t$ | $\tau$ | $t$ | | $t$ | $\tau$ | $t$ | | $t$ | $\tau$ | $t$ | | $t$ | $\tau$ |
| 0·1 | $23_5$ | 24 | 24 | 0·31 | 21 | $22_5$ | 22 | 0·32 | 16 | $16_5$ | 16 | 0·31 | $13_5$ | 15 | 14 | 0·33 |
| 0·2 | $37_5$ | $38_5$ | 38 | 0·49 | $33_5$ | 34 | 34 | 0·49 | 26 | $26_5$ | 26 | 0·50 | 21 | $22_5$ | 22 | 0·52 |
| 0·3 | 51 | $52_5$ | 52 | 0·66 | $44_5$ | 46 | 45 | 0·65 | $33_5$ | 35 | 34 | 0·65 | 28 | 28 | 28 | 0·67 |
| 0·4 | $63_5$ | $65_5$ | $64_5$ | 0·83 | $55_5$ | $56_5$ | 56 | 0·81 | $42_5$ | $43_5$ | 43 | 0·83 | 35 | $34_5$ | 35 | 0·83 |
| 0·5 | $76_5$ | 80 | 78 | | 67 | $70_5$ | 69 | | 52 | $52_5$ | 52 | | 42 | $42_5$ | 42 | |
| 0·6 | 92 | 96 | 94 | 1·20 | 80 | 84 | 82 | 1·19 | 62 | 63 | $62_5$ | 1·20 | $50_5$ | 51 | 51 | 1·21 |
| 0·7 | $110_5$ | | $110_5$ | 1·42 | $94_5$ | 103 | 99 | 1·44 | 75 | $74_5$ | 75 | 1·44 | 61 | 61 | 61 | 1·45 |
| 0·8 | 137 | | 137 | 1·76 | $114_5$ | | $114_5$ | 1·67 | 92 | $89_5$ | 91 | 1·75 | 74 | $73_5$ | 74 | 1·76 |
| 0·9 | | | | | | | | | 118 | $111_5$ | 115 | 2·21 | 96 | $90_5$ | 93 | 2·22 |

| $S$ | 4·0 | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $R$ | 0·30 | | | | | | 0·65 | | | | 1·0 | | | | | |
| $m_0$ | 49·1 | 49·6 | 49·8 | 49·5 | | | 50·2 | 50·5 | | | 49·0 | 49·6 | | | | |
| $t_i$ | 151 | | | | Mean | | $2_5$ | 50 | Mean | | 10 | 4 | Mean | | | |
| $x$ | $t$ | | | | $t$ | $\tau$ | $t$ | | $t$ | $\tau$ | $t$ | | $t$ | $\tau$ | | |
| 0·1 | 20 | $21_5$ | $23_5$ | $20_5$ | $21_5$ | 0·33 | 14 | 14 | 14 | 0·32 | $11_5$ | 11 | 11 | 0·31 | | |
| 0·2 | $31_5$ | 34 | $35_5$ | $32_5$ | $33_5$ | 0·52 | $21_5$ | $22_5$ | 22 | 0·50 | $17_5$ | $17_5$ | $17_5$ | 0·49 | | |
| 0·3 | 41 | 44 | 45 | $43_5$ | $43_5$ | 0·67 | 29 | $29_5$ | 29 | 0·66 | $23_5$ | $23_5$ | $23_5$ | 0·65 | | |
| 0·4 | 51 | $54_5$ | $54_5$ | $54_5$ | $53_5$ | 0·83 | 36 | 37 | $36_5$ | 0·83 | $29_5$ | $29_5$ | $29_5$ | 0·82 | | |
| 0·5 | $62_5$ | $65_5$ | 65 | 65 | $64_5$ | | 44 | $44_5$ | 44 | | 36 | $35_5$ | 36 | | | |
| 0·6 | $75_5$ | 79 | 77 | 78 | $77_5$ | 1·20 | 53 | $53_5$ | 53 | 1·21 | 43 | $42_5$ | 43 | 1·20 | | |
| 0·7 | 91 | 95 | | 94 | $93_0$ | 1·45 | 64 | $65_5$ | 65 | 1·48 | $51_5$ | $50_5$ | 51 | 1·42 | | |
| 0·8 | 112 | 112 | | 115 | 113 | 1·75 | 79 | 82 | $80_5$ | 1·83 | 63 | 61 | 62 | 1·72 | | |
| 0·9 | | 150 | | $151_5$ | 151 | 2·32 | 108 | | 108 | 2·46 | $79_5$ | 78 | 79 | 2·20 | | |

| $S$ | 5·0 | | | | | | | | | | | | 6·0 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $R$ | 0·30 | | | | | | 0·65 | | | | 1·0 | | 0·3 | | | |
| $m_0$ | 50·0 | 50·0 | 50·5 | 50·0 | | | 50·1 | 49·6 | | | 49·0 | | 49·9 | 50·0 | | |
| $t_i$ | 10 | $10_5$ | 10 | 62 | Mean | | 8 | 8 | Mean | | 0 | | 13 | 36 | Mean | |
| $x$ | $t$ | | | | $t$ | $\tau$ | $t$ | | $t$ | $\tau$ | $t$ | $\tau$ | $t$ | | $t$ | $\tau$ |
| 0·1 | 17 | $17_5$ | $18_5$ | 19 | 18 | 0·32 | 12 | $13_5$ | 13 | 0·34 | 11 | 0·34 | $15_5$ | $16_5$ | 16 | 0·30 |
| 0·2 | 27 | $27_5$ | $29_5$ | 29 | 28 | 0·49 | 19 | 20 | $19_5$ | 0·51 | 17 | 0·52 | 25 | 26 | $25_5$ | 0·48 |
| 0·3 | $35_5$ | 37 | 39 | 39 | 38 | 0·67 | 25 | 26 | $25_5$ | 0·66 | 22 | 0·67 | $34_5$ | $34_5$ | $34_5$ | 0·65 |
| 0·4 | $44_5$ | $46_5$ | 48 | 49 | 47 | 0·82 | $31_5$ | 32 | 32 | 0·83 | $27_5$ | 0·84 | $43_5$ | 44 | 44 | 0·83 |
| 0·5 | 54 | 56 | $57_5$ | 59 | 57 | | 38 | 39 | $38_5$ | | $32_5$ | | $52_5$ | 53 | 53 | |
| 0·6 | 65 | 67 | 69 | 70 | 68 | 1·19 | $45_5$ | $45_5$ | $45_5$ | 1·18 | 39 | 1·20 | 63 | 64 | $63_5$ | 1·20 |
| 0·7 | 78 | 80 | $81_5$ | 84 | 82 | 1·44 | 54 | 54 | 54 | 1·40 | 47 | 1·45 | 76 | $75_5$ | 76 | 1·44 |
| 0·8 | 96 | 97 | 100 | 102 | 99 | 1·74 | 65 | 65 | 65 | 1·69 | $56_5$ | 1·74 | $94_5$ | $91_5$ | 93 | 1·76 |
| 0·9 | $120_5$ | 125 | 128 | | 124 | 2·18 | 84 | $82_5$ | $83_5$ | 2·12 | $71_5$ | 2·20 | 125 | 121 | 123 | 2·32 |

| $S$ | 6·0 | | | | | | | | | | 7·0 | | 10·0 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $R$ | 0·65 | | | | 1·0 | | | | 1·4 (calc.) | | 0·30 | | 0·30 | | | | |
| $m_0$ | 49·9 | 49·7 | | | 49·9 | 49·7 | | | 49·5 | | 49·5 | | 48·5 | 50·6 | 50·7 | | |
| $t_i$ | $6_5$ | 47 | Mean | | 24 | 5 | Mean | | $8_5$ | | $14_5$ | | $11_5$ | 0 | 0 | Mean | |
| $x$ | $t$ | | $t$ | $\tau$ | $t$ | | $t$ | $\tau$ | $t$ | $\tau$ | $t$ | $\tau$ | $t$ | | | $t$ | $\tau$ |
| 0·1 | 11 | 12 | $11_5$ | 0·33 | 9 | 9 | 9 | 0·32 | 8 | 0·33 | $16_5$ | 0·33 | 13 | 13 | 13 | 13 | 0·31 |
| 0·2 | 17 | $18_5$ | 18 | 0·51 | 15 | $13_5$ | 14 | 0·50 | $12_5$ | 0·51 | $25_5$ | 0·51 | $20_5$ | 21 | $20_5$ | $20_5$ | 0·49 |
| 0·3 | $22_5$ | $24_5$ | $23_5$ | 0·67 | $19_5$ | 18 | 19 | 0·68 | $16_5$ | 0·67 | $33_5$ | 0·67 | $27_5$ | $27_5$ | $27_5$ | $27_5$ | 0·66 |
| 0·4 | $28_5$ | 30 | 29 | 0·83 | $24_5$ | $22_5$ | $23_5$ | 0·84 | $20_5$ | 0·84 | $41_5$ | 0·83 | 34 | 35 | $34_5$ | $34_5$ | 0·82 |
| 0·5 | 34 | 36 | 35 | | $29_5$ | 27 | 28 | | $24_5$ | | 50 | | 41 | $42_5$ | 42 | 42 | |
| 0·6 | 40 | $42_5$ | 41 | 1·17 | $35_5$ | 32 | 34 | 1·21 | 29 | 1·19 | 59 | 1·18 | $48_5$ | 51 | $50_5$ | 50 | 1·19 |
| 0·7 | $46_5$ | $50_5$ | $48_5$ | 1·39 | 42 | 38 | 40 | 1·43 | 34 | 1·39 | $70_5$ | 1·41 | $58_5$ | 62 | $60_5$ | 60 | 1·43 |
| 0·8 | $55_5$ | 61 | 58 | 1·66 | $50_5$ | 46 | 48 | 1·72 | $40_5$ | 1·66 | $85_5$ | 1·71 | $71_5$ | $77_5$ | $74_5$ | $74_5$ | 1·78 |
| 0·9 | 72 | 76 | 74 | 2·12 | 66 | 61 | $63_5$ | 2·27 | 52 | 2·14 | 117 | 2·34 | 94 | $109_5$ | $100_5$ | 101 | 2·40 |

| $S$ | 10·0 | | 15·0 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $R$ | 1·0 | | 0·30 | | | | | 1·0 | | |
| $m_0$ | 50·2 | | 50·4 | 47·5 | 49·5 | | | 49·9 | | |
| $t_i$ | 17 | | 0 | 0 | 0 | | | | | |
| | | | | | | Mean | | | | |
| $x$ | $t$ | $\tau$ | $t$ | | | $t$ | $\tau$ | $t$ | $\tau$ | $s(t)$ |
| 0·1 | $8_5$ | 0·35 | 11 | 10 | 10 | 10 | 0·29 | 6 | 0·33 | 0·7 |
| 0·2 | 13 | 0·53 | 18 | 16 | $16_5$ | 17 | 0·49 | 9 | 0·50 | 0·9 |
| 0·3 | 17 | 0·69 | 24 | 21 | $22_5$ | $22_5$ | 0·65 | 12 | 0·67 | 1·1 |
| 0·4 | 21 | 0·86 | $29_5$ | $26_5$ | 29 | 28 | 0·81 | 15 | 0·83 | 1·2 |
| 0·5 | $24_5$ | | 36 | $32_5$ | 35 | $34_5$ | | 18 | | 1·5 |
| 0·6 | 29 | 1·18 | $42_5$ | $39_5$ | 42 | 41 | 1·19 | 21 | 1·17 | 1·8 |
| 0·7 | $34_5$ | 1·41 | $50_5$ | 49 | 50 | 50 | 1·45 | 25 | 1·39 | 2·2 |
| 0·8 | $41_5$ | 1·70 | $61_5$ | $61_5$ | $62_5$ | 62 | 1·80 | 31 | 1·72 | 2·3 |
| 0·9 | 53 | 2·16 | 82 | $83_5$ | $87_5$ | 84 | 2·44 | $40_5$ | 2·25 | 4·3 |

$m_0$ = initial amount of monomer in grams,
$S$ = amount of soap in grams,
$R$ = amount of catalyst in grams,
$t_i$ = length of induction time in minutes,
$s(t)$ = standard deviation of $t$.
The meanings of the other symbols are to be found in the list of symbols.

# TABLE II

Polymerizations with different soaps, etc.

K myristate

| $S$ | 2·0 | | | | 3·0 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $R$ | 0·15 | | | | 0·3 | | | | | |
| $m_0$ / $V$ | 20·1 / 225 | 24·8 / 225 | 22·5 / 225 | | 26·6 / 425 | 24·5 / 425 | 25·5 / 425 | | 50·1 / 200 | |
| | | | Mean | | | | Mean | | | |
| $x$ | $t$ | | $t$ | $\tau$ | $t$ | | $t$ | $\tau$ | $t$ | $\tau$ |
| 0·1 | 23 | 23 | 23 | 0·35 | $20_5$ | $14_5$ | $17_5$ | 0·36 | $21_5$ | 0·29 |
| 0·2 | $35_5$ | 35 | 35 | 0·53 | 29 | $23_5$ | 26 | 0·54 | $33_5$ | 0·46 |
| 0·3 | $45_5$ | 45 | 45 | 0·68 | $36_5$ | $31_5$ | 34 | 0·71 | $46_5$ | 0·64 |
| 0·4 | $55_5$ | $55_5$ | $55_5$ | 0·84 | $43_5$ | $38_5$ | 41 | 0·85 | 59 | 0·81 |
| 0·5 | 65 | $66_5$ | 66 | | $50_5$ | 46 | 48 | | $72_5$ | |
| 0·6 | 77 | 79 | 78 | 1·18 | 59 | 55 | 57 | 1·19 | 88 | 1·22 |
| 0·7 | 92 | 91 | 92 | 1·40 | 70 | 67 | 69 | 1·44 | 106 | 1·46 |
| 0·8 | 112 | 116 | 114 | 1·73 | $88_5$ | 84 | 86 | 1·79 | 129 | 1·78 |
| 0·9 | | | | | | 113 | 113 | 2·35 | | |
| $k$ | 1·09 | | | | 1·05 | | | | 1·03 | |

| $S$ | 4·0 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $R$ | 0·3 | | | | | | | | | |
| $m_0$ / $V$ | 22·4 / 425 | 22·8 / 425 | 23·7 / 425 | 23·0 / 425 | | 49·9 / 200 | 49·0 / 200 | 49·5 / 200 | | |
| | | | | Mean | | | | Mean | | |
| $x$ | $t$ | | | $t$ | $\tau$ | $t$ | | $t$ | $\tau$ | |
| 0·1 | 15 | $14_5$ | $15_5$ | 15 | 0·37 | $20_5$ | $22_5$ | $21_5$ | 0·34 | |
| 0·2 | 22 | $21_5$ | 23 | 22 | 0·54 | $31_5$ | 34 | 33 | 0·52 | |
| 0·3 | 28 | 28 | 29 | 28 | 0·69 | $41_5$ | $44_5$ | 43 | 0·68 | |
| 0·4 | $33_5$ | 34 | 35 | 34 | 0·84 | 51 | $54_5$ | 53 | 0·83 | |
| 0·5 | 39 | $40_5$ | $41_5$ | $40_5$ | | 62 | 65 | $63_5$ | | |
| 0·6 | 45 | | $48_5$ | 47 | 1·16 | $73_5$ | 77 | 75 | 1·18 | |
| 0·7 | 53 | | $57_5$ | 55 | 1·36 | 89 | 91 | 90 | 1·42 | |
| 0·8 | $63_5$ | | 71 | $67_5$ | 1·67 | 108 | 110 | 109 | 1·72 | |
| 0·9 | $83_5$ | | 91 | $87_5$ | 2·16 | 139 | 137 | 138 | 2·17 | |
| $k$ | 1·20 | | | | | 1·11 | | | | |

| | K laurate | | | | K stearate | | | | | | K oleate | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $S$ | 5·0 | | | | 4·0 | | 5·0 | | 10·0 | | 5·0 | |
| $R$ | 0·65 | | | | 1·0 | | 0·3 | | 0·3 | | 0·3 | |
| $m_0$ / $V$ | 49·3 / 400 | 49·8 / 400 | 49·5 / 400 | | 50 / 400 | | 50·4 / 400 | | 50·4 / 400 | | 49·7 / 200 | |
| | | | Mean | | | | | | | | | |
| $x$ | $t$ | | $t$ | $\tau$ | $t$ | $\tau$ | $t$ | $\tau$ | $t$ | $\tau$ | $t$ | $\tau$ |
| 0·1 | $19_5$ | $19_5$ | $19_5$ | 0·36 | 12 | 0·34 | $13_5$ | 0·26 | 12 | 0·31 | $12_5$ | 0·25 |
| 0·2 | 29 | 28 | 29 | 0·54 | 18 | 0·51 | $23_5$ | 0·44 | 19 | 0·49 | $21_5$ | 0·43 |
| 0·3 | 37 | $36_5$ | 37 | 0·69 | 23 | 0·66 | 33 | 0·62 | 26 | 0·67 | $29_5$ | 0·58 |
| 0·4 | 45 | 45 | 45 | 0·83 | $28_5$ | 0·81 | $42_5$ | 0·80 | 32 | 0·82 | 39 | 0·77 |
| 0·5 | 53 | 54 | 54 | | 35 | | 53 | | 39 | | $50_5$ | |
| 0·6 | 64 | 65 | 65 | 1·20 | $42_5$ | 1·21 | $64_5$ | 1·22 | 47 | 1·20 | 66 | 1·31 |
| 0·7 | 77 | 78 | 78 | 1·45 | $52_5$ | 1·50 | $77_5$ | 1·46 | 57 | 1·46 | 87 | 1·72 |
| 0·8 | 100 | 98 | 99 | 1·83 | 66 | 1·89 | 94 | 1·79 | 71 | 1·82 | $120_5$ | 2·38 |
| 0·9 | | 123 | 123 | 2·28 | 100 | 2·88 | $120_5$ | 2·27 | 95 | 2·44 | | |
| $k$ | 1·00 | | | | 0·93 | | 1·04 | | 1·00 | | 0·60 | |

The meanings of the symbols are to be found in a list at the end of the paper.

with the moment at which the reaction starts, we find in the table the time at which several whole multiples of 10% conversion were reached. These data are derived from the curves drawn through the thirty to forty points representing the values determined during the actual experiment. Most of the experiments were duplicated, some were triplicated and two were carried out in quadruplicate. In the further analysis we used the arithmetic mean of the determinations and, moreover, to get an idea of the degree of accuracy reached we calculated the standard deviation of the times along each row in the table (thus with $x$ constant) by means of the formula

$$s = \left\{ \frac{\Sigma (n_\lambda - 1) s_\lambda^2}{\Sigma (n_\lambda - 1)} \right\}^{1/2}$$

where $n_\lambda$ is the number of determinations in each set of conversions under the same conditions, and $s_\lambda$ is the standard deviation of the times included in each set. It appears that $s$ is about 5 % of the mean value of $t$. It is further to be noted that the deviations of the individual points from the curve which we drew between them are generally much less, from which it may be concluded that the main error in $t$ is caused by the inaccuracy in the determination of the zero point.

IPR2025-00217 EXHIBIT 2010 Page 368 of 485

Beside each column in table I containing the mean values of $t$, anothe column is added giving values of $\tau$, that is these times divided by the co responding half-time value $t_{1/2}$, the time at which a conversion of 50 % i reached. Within our limits of accuracy, this at once reveals the strikin uniformity of $\tau = t/t_{1/2}$, which is independent of the circumstances goverr ing the reaction, i.e. the amounts (in grams) of soap ($S$) or catalyst ($R$ used. See table III.

TABLE III

Dependency of $\tau$ upon $x$, averaged over several values of $S$, for various amounts of catalyst ($R$)

| $R$ \ $x$ | 0·1 | 0·2 | 0·3 | 0·4 | 0·6 | 0·7 | 0·8 | 0·9 |
|---|---|---|---|---|---|---|---|---|
| 0·3 | $0{\cdot}31_5$ | $0{\cdot}49_5$ | 0·66 | 0·82 | $1{\cdot}19_5$ | $1{\cdot}43_5$ | 1·75 | $2{\cdot}33_5$ |
| 0·65 | $0{\cdot}32_5$ | $0{\cdot}50_5$ | 0·66 | 0·83 | 1·19 | 1·43 | $1{\cdot}73_5$ | $2{\cdot}19_5$ |
| 1·0 | 0·33 | 0·51 | 0·67 | $0{\cdot}83_5$ | $1{\cdot}19_5$ | $1{\cdot}42_5$ | $1{\cdot}72_5$ | $2{\cdot}21_5$ |
| Mean | 0·32 | 0·50 | $0{\cdot}66_5$ | 0·83 | $1{\cdot}19_5$ | 1·43 | 1·74 | 2·26 |
| s.d. | $0{\cdot}01_5$ | $0{\cdot}01_3$ | $0{\cdot}01_1$ | $0{\cdot}01_2$ | $0{\cdot}01_3$ | $0{\cdot}02_4$ | $0{\cdot}04_5$ | $0{\cdot}10_0$ |

A slight influence of $R$ on the values of $\tau$ may be traced, but to a fir approximation this can be ignored; such an influence of the amount soap could not be detected.

Generally speaking $t$ is not only a function of $x$ but also of the paramete determining the initial conditions of the reaction system, viz. the amou of monomer $m_0$, the volume of the water $V$, $S$ and $R$ and moreover of th temperature, which will not be taken into account here.

Hence

$$t = f(x; m_0, V, R, S).$$

But the facts pointed out above indicate that this function may be writte in the form

$$t = f_1(x; m_0, V)\, f_2(R, S),$$

since now

$$\tau = \frac{t}{t_{1/2}} = \frac{f_1(x; m_0, V)\, f_2(R, S)}{f_1(\tfrac{1}{2}; m_0, V)\, f_2(R, S)} = \frac{f_1(x; m_0, V)}{f_1(\tfrac{1}{2}; m_0, V)},$$

an expression independent of $S$ and $R$. Of course the reduction of th time values with respect to a value of $t$ at any other fixed value of instead of $t_{1/2}$ provides similar results. From this exposition it becom

clear that $t_{1/2}$ proves to be a suitable parameter for characterizing the particular conversions.

Before considering the dependency of this parameter upon $R$ and $S$, we introduce two functions replacing

$$f_1(x; m_0, V)/f_1(\tfrac{1}{2}; m_0, V).$$

It proved that for $x \leqslant 0{\cdot}5$ we could use the empirical*) relation

$$\Psi_1(x) \equiv \left(\frac{x}{1-x}\right)^{1/2}$$

and for $x \geqslant 0{\cdot}5$ the theoretical **) relation

$$\Psi_2(x, k) = \frac{1-\varphi}{k}(x - \tfrac{1}{2}) - \frac{\varphi}{k}\ln\{2(1-x)\} + 1,$$

where $\varphi = \varphi_p/\varphi_m = 0{\cdot}84$ and $k$ is a constant, probably still depending upon $m_0$, $V$ and the temperature.

From table IV it appears that, each in its own interval, the two expressions proposed lend themselves excellently for describing $\tau$ as a function of $x$.

TABLE IV

The values of $\Psi_1$, $\Psi_2$ and $\tau$ for various values of $x$

| $x$ | 0·0 | 0·1 | 0·2 | 0·3 | 0·4 | 0·5 | 0·6 | 0·7 | 0·8 | 0·9 | 1·0 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\Psi_1(x)$ | 0·0 | 0·33 | 0·50 | 0·66 | 0·82 | 1·00 | 1·23 | 1·53 | 2·00 | 3·00 | $\infty$ |
| $\Psi_2(x,k=1)$ | 0·34 | 0·45 | 0·56 | 0·70 | 0·84 | 1·00 | 1·21 | 1·46 | 1·82 | 2·42 | $\infty$ |
| $\Psi_2(x,k=1{\cdot}10)$ | 0·40 | 0·49 | 0·60 | 0·72 | 0·85 | 1·00 | 1·18 | 1·42 | 1·74 | 2·29 | $\infty$ |
| $\tau$ (mean) | 0·00 | 0·32 | 0·50 | 0·66 | 0·83 | 1·00 | 1·19 | 1·43 | 1·74 | 2·26 | $\infty$ |

The values of $\tau$ (mean) are taken from table III.

The values of $k$, determined from the mean values of the various sets of $t$ by a graphical method, are to be found in table V (figs 4 and 5).

Besides the conversions with myristate we carried out some others with potassium stearate, oleate and laurate and also some with diminished amounts of water or of monomer. The results of these experiments are given

*) In connection herewith it may be mentioned that Corrin [13]) found that the same function proved to be proportional to the time as soon as $x > 0{\cdot}5$. Dr Klaassens of this laboratory, when attempting to verify this relation for his own conversions, found that it holds very well even for $x < 0{\cdot}5$ when applying it to $S$-shaped curves.

**) Thanks to the much simpler conditions prevailing in the systems as soon as $x$ becomes higher than 0·5, this relation could be derived on theoretical grounds, as will be shown in section III, 1.

in table II. When suitable values of $k$ are chosen they all confirm the secon function for conversions higher than $x = 0{\cdot}5$; on the other hand they d not conform to the results which would be expected from $\Psi_1$ *).

There remains to be considered the influence that the applied condition have on the chosen parameters. As far as $k$ is concerned it may be said tha the results fluctuate round a value of 1·10, with a standard deviatio of 0·07 (table V). It was not possible to detect any uniformity in th discrepancies depending upon the amounts of soap or catalyst.

TABLE V

Values of $k$ determined from the conversions with K myristate

| R \ S | 2·5 | 3·0 | 4·0 | 5·0 | 6·0 | 7·0 | 10 | 15 | Mean |
|---|---|---|---|---|---|---|---|---|---|
| 0·3 | 1·08 | 1·06 | 1·06 | 1·10 | 1·06 | 1·11 | 1·05 | 1·00 | 1·09 |
| $0{\cdot}6_5$ | | 1·08 | 1·00 | 1·15 | 1·21 | | | | 1·10 |
| 1·0 | | 1·08 | 1·16 | 1·10 | 1·12 | | 1·17 | 1·16 | 1·11 |
| 1·4 | | | | | 1·25 | | | | |

Average $k = 1{\cdot}10 \pm 0{\cdot}07$

*) The analysis of the curves outlined above introduces a discontinuity in the fir derivative $dx/d\tau$ in the neighbourhood of $x = 0{\cdot}5$ when calculating this from $\Psi_1$ from $\Psi_2$ but in reality no such discontinuity exists.

In the case of oleate we could avoid the introduction of the discontinuity by writin $\Psi_1$ in a more general form

$$\Psi_1(x, a) \equiv \left\{\frac{ax}{1-(2-a)x}\right\}^{1/3} = \tau.$$

Then $|dx/d\tau|_{\tau=1}$ calculated from $\Psi_1$ is equal to $a/2$,

whereas $|dx/d\tau|_{\tau=1}$ calculated from $\Psi_2$ is equal to $k/(1+\varphi)$.

Putting this two values equal yields

$$a = \frac{k}{0{\cdot}92}.$$

For oleate $k = 0{\cdot}6$, thus $a = 0{\cdot}65$ and the values of the new function

$$\Psi_1(x, a = 0{\cdot}65) \equiv \left\{\frac{0{\cdot}65\ x}{1-1{\cdot}35\ x}\right\}^{1/3}$$

computed for successive values of $x$ are now respectively

| $x$ | 0·1 | 0·2 | 0·3 | 0·4 | 0·5 |
|---|---|---|---|---|---|
| $\Psi_1(x,a)$ | 0·27 | 0·42 | 0·57 | 0·75 | 1·0 |

agreeing well with the values of $\tau$ found for this case.

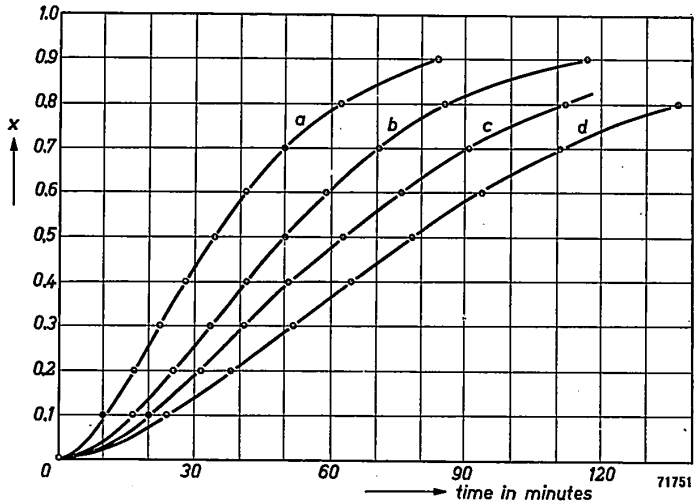



Fig. 4*a*. Conversion-time curves of emulsion polymerizations at 40 °C of 50 g styrene with 0·3 g $K_2S_2O_8$ in 400 ml water, respectivily with (*a*) 15 g, (*b*) 7 g, (*c*) 4 g, (*d*) 2·5 g potassium myristate.

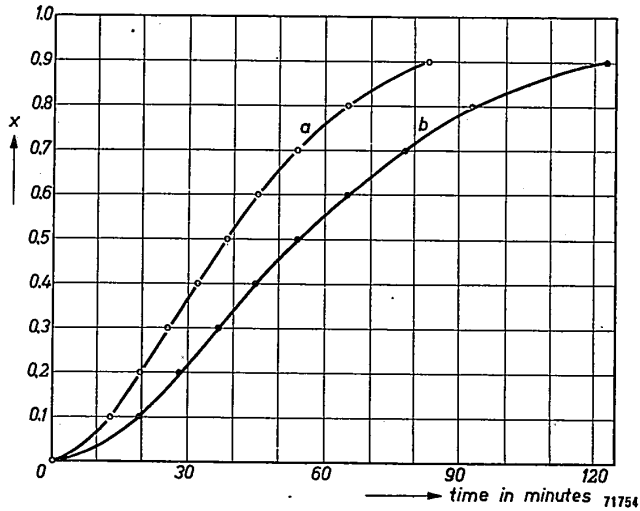



Fig. 5. Conversion-time curves of emulsion polymerizations at 40 °C of 50 g styrene with 0·65 g $K_2S_2O_8$ in 400 ml water, respectively with (*a*) 5 g potassium myristate, (*b*) 5 g potassium laurate.

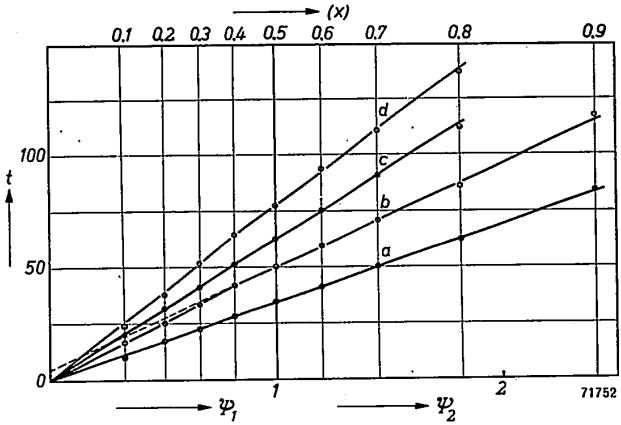



Fig. 4*b*. Analysis in terms of $\Psi_1$ and $\Psi_2$ of the curves given in fig. 4*a*.

In the diagram lines are drawn parallel to the ordinate and cutting the abscissa at di tances from the origin equal to the values of $\Psi_1$ (in the interval from $x = 0$ up to $x = 0$· and the values of $\Psi_2$ $(x, k = 1)$ (in the interval $x = 0{\cdot}5$ up to $x = 0{\cdot}9$) belonging to t values of the argument $x$ shown at the top of the diagram. Plotting on these lines the tir values corresponding to the respective degrees of conversion, we get a number of point of which those to the left of the line $\Psi_1 = \Psi_2 = 1$ should fall on a straight line drav through the origin to the point $(1, t_{1/_2})$ while those to the right should fall on a straight li through the same point $(1, t_{1/_2})$ but not necessarily passing through the origin. The extensi of the latter line cuts the abscissa at a point removed from the point (1,0) by a distan equal to $k$. This is easily understood as

$$\Psi_2(x,k) \equiv \frac{1-\varphi}{k}(x-\tfrac{1}{2}) - \frac{\varphi}{k}\ln[2(1-x)] + \frac{1}{k} - \left(\frac{1}{k}-1\right) = \frac{1}{k}\Psi_2(x,k=1) - \left(\frac{1}{k}-1\right) =$$

and consequently

$$\frac{\mathrm{d}\Psi_2(x,k=1)}{\mathrm{d}t} \equiv \frac{\mathrm{d}\Psi_2}{\mathrm{d}t} = \frac{k}{t_{1/_2}}.$$

It is immediately evident that if the time values are subject to an error $\delta t$ as a consequer of an uncertainty with respect to the zero point, there will be an error in the value o to an extent of $\delta k = (k/t_{1/_2})\delta t$.

The half-time value displays a more obvious behaviour in this respe (table VI).

## TABLE VI

Values of $t_{1/2}$ determined from the conversions with K myristate

| R \ S | 2·5 | 3·0 | 4·0 | 5·0 | 6·0 | 7·0 | 10 | 15 |
|---|---|---|---|---|---|---|---|---|
| 0·3 | 78 | 69 | $64_5$ | 57 | 53 | 50 | 42 | $34_5$ |
| $0{\cdot}6_5$ | | 52 | 44 | $38_5$ | 35 | | | |
| 1·0 | | 42 | 36 | $32_5$ | 28 | | $24_5$ | 18 |
| 1·4 | | | | | $24_5$ | | | |

The values of $t_{1/2}$ are expressed in minutes. The values of $S$ and $R$ in g.

Plots on double-logarithmic paper at once reveal that these $t_{1/2}$ values are proportional to the reciprocal square root of the amounts of soap or catalyst (figs 6 and 7). On the average we found

$$t_{1/2} = (71{\cdot}0 \pm 2{\cdot}7)\,(RS)^{-1/2}\ \text{min}$$

using potassium myristate and persulphate to polymerize 50 g styrene dispersed in 400 ml water.

The experiments performed with potassium stearate show the same property (table VII).

## TABLE VII

Values of $t_{1/2}$ of the conversion with K stearate

| R \ S | 4 | 5 | 10 |
|---|---|---|---|
| 0·3 | | 53 | 39 |
| 1·0 | 35 | | |

Here we found on the average

$$t_{1/2} = (67{\cdot}5 \pm 2{\cdot}5)\,(RS)^{-1/2} \text{ min} \quad \text{and} \quad k = 0{\cdot}99 \pm 0{\cdot}06$$

polymerizing 50 g of styrene in 400 ml water.

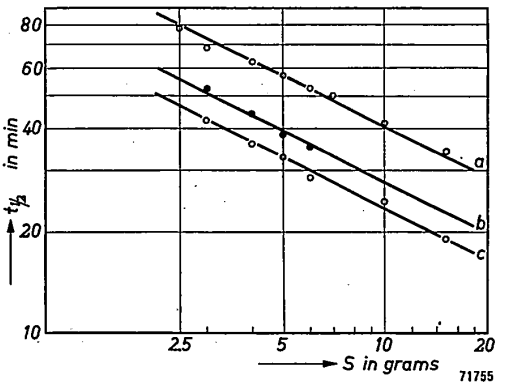



Fig. 6. Dependency of $t_{1/2}$ upon the amount of soap (K myristate); (*a*) 0·3 g, (*b*) 0·65 g (*c*) 1·0 g of catalyst.

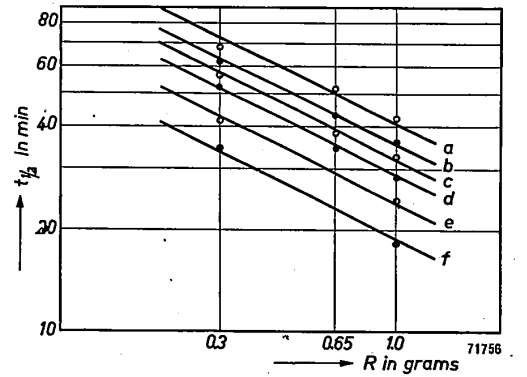



Fig. 7. Dependency of $t_{1/2}$ upon the amount of catalyst; (*a*) 3 g, (*b*) 4 g, (*c*) 5 g, (*d*) 6 g (*e*) 10 g, (*f*) 15 g of K myristate.

Analogous findings have already been announced with respect to th initial velocities [11] [24]). From the discussion, however, it will become clea that the use of $t_{1/2}$ offers great advantages compared with the velocity itself

## 4. Results of the conductivity measurements

Concerning the conductivity determinations the following is to be remarked. When we are confronted with an induction period, where the reaction is probably inhibited by the presence of materials which kill the radicals or stop the reacting chains, the resistance of the system might be expected to remain constant until the moment that the inhibitors are used up. Such a behaviour is indeed found when carrying out experiments with potassium stearate [10]), whereas in the systems with potassium myristate the resistance often oscillates in a rather irregular manner about a mean value, sometimes showing a tendency to increase, although likewise irregularly or sometimes even slightly decreasing. But as soon as the reaction starts a steep and quite linear decline of the resistance sets in. Figs 8 and 9 show for K myristate that this rapid decline is not followed by a rather abrupt transition to a temporarily constant resistance, as found with stearate, but only by an arbitrary deviation from linearity.

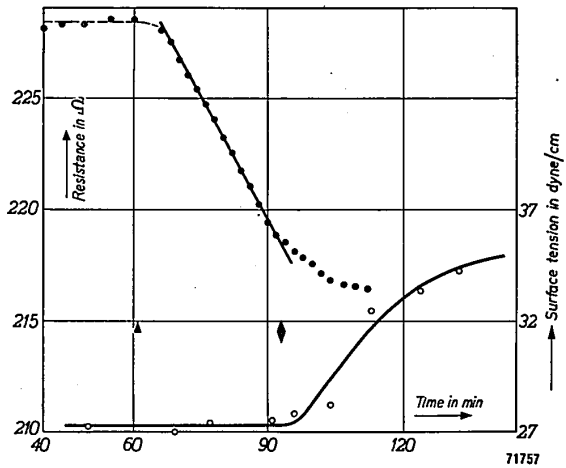



Fig. 8. Dependency of the conductivity and the surface tension upon time; dots: conductivity; circles: surface tension; ▲ starting point; ♦ end point.

The difficulty is now to decide where the two main points of the curve, namely that indicating the start of the reaction and the other indicating the moment where the micelles disappear, are to be localized. Probably we may assume that the steep part of the curve will be preceded by a smooth transition to the first part. The time gap to be bridged by this transition never exceeds more than a few minutes. In all the cases where inhibition periods are absent we found at once a sudden decrease of the resistance.

As regards the determination of the second point, it was found that the best way to obtain reproducible results was to determine the point where the steep fall loses its linear character. This was checked by carrying out some experiments where the change of the resistance as well as the surface tension of one and the same system were measured. The latter quantity was measured with a stalagmometer and the result showed that reasonable confidence could be placed in the criterion chosen.

The values of $t^*$, the time at which the micelles have disappeared, are given in table VIII. No values are shown for amounts of soap larger than 6 grams, because with higher amounts of this substance $t^*$ becomes larger than $t_{1/2}$, the moment of disappearance of emulsified monomer, and so impossible to detect. We must assume then to have empty micelles existing next to latex particles containing all the remaining monomer. The reactions with 2·5 g soap, which showed an increase of the resistance, have been ignored here on account of this abnormal behaviour.

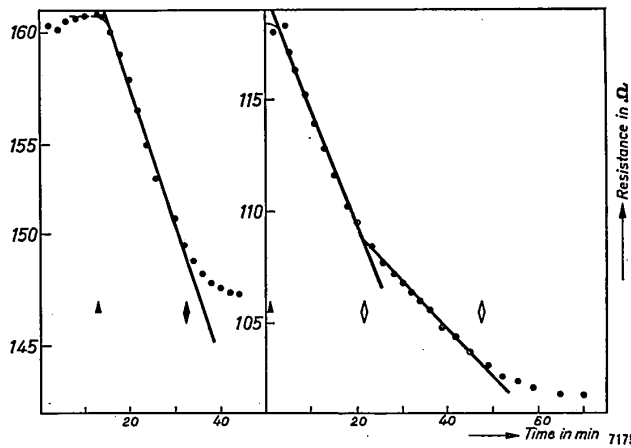



Fig. 9. Conductivity-time curves. To the left for the case of $S = 5$, $R = 1{\cdot}0$, to the right for the case of $S = 15$, $R = 0{\cdot}3$. ▲ starting point; ♦ end point; ◊ characteristic points, to be discussed in section III, 3.

To complete the information given we remark that

(a) the relative change of the resistance from the beginning of the reaction up to $t^*$ is of the order of 8% in the case of potassium myristate;

(b) the change per minute during this interval is of the order of 0·5 Ω.

The amount of soap seems to have hardly any influence on the value of $t^*$. To a first approximation the influence of $R$ can be described by

$$t^* = \text{const.}\ R^{-3/5}\ .$$

## TABLE VIII

Values of $t^*$ (in minutes) for the conversions with K myristate

| $S$ | 3·0 | | | 4·0 | | | 5·0 | | | 6·0 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $R$ | 0·3 | $0{\cdot}6_5$ | 1·0 | 0·3 | $0{\cdot}6_5$ | 1·0 | 0·3 | $0{\cdot}6_5$ | 1·0 | 0·3 | $0{\cdot}6_5$ | 1·0 |
| $t^*$ | 32<br>35 | 22<br>24 | 19<br>19 | 32<br>*31*<br>*33* | 27<br>28<br>*28* | (32)<br>23<br>*20*<br>*21* | 32<br>37<br>34<br>35<br>*32*<br>*32*<br>*34* | 22<br>(27)<br>*20* | 22<br>*20* | 30<br>30 | 21<br>24<br>*20* | (26)<br>(24)<br>*16*<br>*18* |
| Average $t^*$ | $33_5$ | 23 | 19 | 32 | 28 | 21 | 34 | 21 | 21 | 30 | 22 | 17 |

The figures in italics were obtained from experiments where the conversions were not measured step-by-step but only incidentally as a check. The figures between parentheses are not used for further calculations, being obviously erroneous. $S$ and $R$ are expressed in g.

It is difficult to say what degree of accuracy is to be ascribed to the values found for $t^*$. The standard deviation, calculated from the several values is about 1·6. This may indicate that the probable error is of the order of a few minutes.

## III. DISCUSSION

### 1. The stationary state

To find a theoretical basis for the interpretation of the results obtained, attention is first confined to that part of the reaction where the conditions are by far the simplest, i.e. the part where the emulsified monomer phase no longer exists but all the remaining monomer is contained in the latex particles and mixed with the polymer meanwhile formed. Our discussion starts from ideas developed by Smith and Ewart [5]), taking a system of isolated soap-stabilized particles, each containing monomer, able to polymerize under the influence of radicals. These radicals, generated outside the particles in an aqueous phase, enter the particles by mere accident. The radical then generally starts a new reaction chain, but in the case where a growing macromolecule is already present in the particle the question is to be asked whether the radical will grow into a new chain or react with the growing macromolecule, such as to terminate its growth. The last alternative is the more likely the smaller the dimensions of the particles and the higher the rate of the termination reaction. It will be shown below for our systems that, when a radical enters a particle containing

a growing macromolecule, the chain is stopped before the radical has itself built up a chain of any importance, the particle then remaining inactive until it is again struck by a radical, which then without hindrance starts a new chain. If the average time elapsing between the entry of two successive radicals into a particle is long compared with the average time needed for a radical to stop any active chain present, a peculiar stationary state may arise.

At this point it may be of some interest to see how far the above considerations can be applied to our systems. First of all it is required to know the diameter of the latex particles, which can be determined from kinetic data, as will be discussed later on, as well as from electron-microscopic observations *). Both methods gave an average diameter of 600 Å, which is in agreement with the general experience that the diameter seldom exceeds 1000 Å.

Imagine a particle of a volume $u$ and containing two growing chains ($M^\bullet$). The mutual reaction by which the chains terminate may be denoted by

$$2\,M^\bullet \xrightarrow{k_t} P,$$

where $P$ stands for the macromolecules, the number of which present in the particle may be $p$; $k_t$ is the constant of the termination reaction.

If there were $z$ growing chains in the particle then generally the rate at which $p$ increases would be

$$\frac{dp}{dt} = k_t \frac{z(z-1)}{u}.$$

In our case $z = 2$, the reaction between the two chains giving rise to an increase in the number of macromolecules of $\Delta p = 1$.

Hence

$$\frac{\Delta p}{\Delta t} \approx k_t \frac{2}{u},$$

where $\Delta t$ may be conceived to be the average time elapsing before the chains react. Thus

$$\Delta t = \frac{u}{2k_t}.$$

Now $u = (\pi/6)(600)^3.10^{-27}$ l and for styrene at 40 °C the value of $k_t = 7.10^6$ l. mole$^{-1}$. sec$^{-1}$ Thus $\Delta t$ is of the order of 1/200 sec.

The growth of a reacting particle may be symbolized by

$$M_n^\bullet + M \xrightarrow{k_p} M_{n+1}^\bullet, \text{ etc.},$$

hence

$$-\frac{dM}{dt} = k_p[M],$$

where $k_p$, the constant of the propagation reaction, is taken equal to $6.10^{-18}$ ml.min$^{-1}$

*) From electron-microscopic photographs, with myristate conversions, in the case where $S = 6$ and $R = 0{\cdot}3$ the average diameter of the latex particles appeared to be 390 Å, while in the case of $S = 6$ and $R = 0{\cdot}65$ this diameter was 805 Å. As a reference emulsion 580 G, lot nr 3584, of Dow Chemical Company with a particle diameter of 2500 Å was used. Calculation of particle dimensions from kinetic data, to be described elsewhere, had the following result. For the first case we found a diameter of 620 Å, while for the second case 520 Å was found. Although the experimental and the calculated values do not agree well, in every case the diameter of the particles proved to be of the order of 600 Å.

per radical *). The mean value of $[M]$ may be taken equal to 0·5 g.ml$^{-1}$ and thus

$$-\frac{dM}{dt} = 3.10^{-18}\ \text{g.min}^{-1}.$$

$$\approx 18\,000 \text{ molecules per minute.}$$

Every alternate radical entering will have hardly any time to react twice with a styrene molecule, and its contribution to chain formation may therefore be ignored. To obtain the number of minutes elapsing between the entry of two successive radicals we have to know the total number of particles $n^*$ and the production of radicals per minute $\varrho$. From Smith's data [30]) we deduce a value for the constant of the monomolecular decomposition of persulphate $k_d = 1{\cdot}1.10^{-3}$ h$^{-1}$ and from Bartlett's data [31]) we find $k_d = 1{\cdot}9.10^{-3}$ h$^{-1}$ at 40 °C taking for the activation energy of this reaction a value of 30 kcal. mole$^{-1}$. Our own measurement gave a decomposition of 24 % during a time interval of 7 days; a determination after four days gave 15 % decomposition which is in agreement with the monomolecular behaviour of the reaction. From these data $k_d = 1{\cdot}6.10^{-3}$ h$^{-1}$ is obtained; further computation with this value gives for the production of radicals per gram persulphate

$$\varrho = 12.10^{16}\ \text{min}^{-1}.$$

Since the number of latex particles having diameters of 600 Å and originating from 50 g of styrene is

$$n^* = 4{\cdot}2.10^{17},$$

we obtain for the average time interval between two successive entries of radicals into a particle

$$\frac{4{\cdot}2.10^{17}}{12.10^{16}} = 3{\cdot}5 \text{ min } **),$$

thus differing by a factor of about 40 000 from the time needed for the termination of two radicals present at the same time in one latex particle.

We shall approach the stationary state from a different angle to that followed by Smith and Ewart. Suppose we have at a certain moment $n_1$ particles, struck an odd number of times by successive radicals and thus containing a growing chain, and $n_0$ particles struck an even number of times and thus inactive, all in a medium generating per unit of time $\varrho$ radicals. How will the numbers $n_1$ and $n_0$ have been altered some time later?

---

*) The values of the reaction constants are taken from a selection given by Küchler [2]). The value of $k_p$ is essentially Burnett's [25]).

**) From this value of 3·5 min a kinetic chain length of about 60 000 may be expected. The average degree of polymerization however will not reach this value, as a consequence of the transfer reaction. The transfer constant for styrene at 40 °C is about $k_f = 3.10^{-22}$ ml.min$^{-1}$ per radical [2]), a value limiting the average degree of polymerization to

$$\overline{P}_n = \frac{k_p}{k_f} = 20\,000.$$

Viscosity measurements of solutions in benzene gave on the average a value for the intrinsic viscosity of 8·1 from which with the aid of the formula [33])

$$[\eta] = 0{\cdot}754.10^{-4}\ M^{0{\cdot}783}$$

a mean molecular weight of $2{\cdot}6.10^6$ was found, in good agreement with the molecular weight calculated when taking the transfer reaction into account.

During a short time $dt$, the number of generated radicals will be $\varrho dt$ and the same number of them will have entered $\varrho dt$ particles. A particle of the $n_1$ group will pass over to the $n_0$ group when it is struck, and vica versa. Thus we obtain the following differential equation

$$dn_1 = -\frac{n_1}{n_1+n_0}\varrho dt + \frac{n_0}{n_1+n_0}\varrho dt,$$

supposing that the entry of a radical into a particle is independent of the dimensions of the particles. Since $n_1 + n_0$ is constant and equal to $n^*$ we get

$$\frac{dn_1}{dt} = \frac{n^* - 2n_1}{n^*}\varrho,$$

$$n_1 = \tfrac{1}{2}n^* + C\exp\left(-2\frac{\varrho t}{n^*}\right). \qquad (1)$$

To find the boundary conditions we consider the initiation, according to Smith and Ewart. The simplest of their calculations leading to $n^*$ assumes that during the stage marked by the presence of micelles each of the radicals generated in the aqueous phase converts a micelle into a latex particle. Taking $t^*$ to be the moment at which the micelles disappear, $n^*$ will be equal to $\varrho t^*$ and at the same moment equal to the number of active particles $n_1$, none of the latex particles meanwhile formed being struck a second time. The latter simplification certainly leads to a value of $n^*$ that is too high and likewise in reality $n_1$ will be smaller than $n^*$. Provisionally we choose

$$|n_1|_{t=t^*} = n^* = \varrho t^*,$$

which enables us to eliminate $C$ from (1); this yields

$$n_1 = \tfrac{1}{2}n^*\left\{1 + \exp\frac{2(t^*-t)}{t^*}\right\}. \qquad (2)$$

When $t \gg t^*$ obviously $n_1$ approaches the value $\frac{1}{2}n^*$. For a further investigation into the properties of our system it is immaterial whether the stationary state $n_1 = \frac{1}{2}n^*$ is ever reached, the point of importance being rather in how far the actual state differs from that stationary state. With the accuracy of our measurements of about 5%, the point of interest to us is what time has elapsed since $t^*$ before $n_1$ differs by, say, only 4 % from the stationary number, thus before $n_1$ reaches the value of 0·52 $n$. Formula (2) yields for this 1·6 $t^*$.

Imagine a case where $t_{n_1} - t^* \geqslant 1{\cdot}6\,t^*$ and let us apply to this the stationary condition $n_1 = \frac{1}{2}n^*$. The rate of polymerization in any of the $n_1$ active particles is a linear function of the concentration of the monomer

incorporated; thus the decrease of monomer per unit of time in the whole system will be

$$-\frac{\mathrm{d}m}{\mathrm{d}t} = n_1 k_p c\,,$$

if $c$, the concentration of the monomer in situ, on the average is the same in every particle whether active or not, and $k_p$ is the constant of the propagation reaction per radical. For convenience $c$ will be expressed in g. ml$^{-1}$ and $k_p$ in ml.min$^{-1}$. Now $m = (1-x)m_0$ and $n_1 = \frac{1}{2}\,n^*$, so we may also write

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{n^* k_p c}{2m_0}.$$

Under the conditions prevailing beyond 50 % conversion an expression for $c$ is easily arrived at, since, $m_0 - m$ being the amount of polymer formed,

$$c = \frac{m}{\varphi_m m + \varphi_p(m_0 - m)} = \left\{\varphi_m - \varphi_p + \frac{\varphi_p}{1-x}\right\}^{-1}.$$

Substituting this in the equation for $\mathrm{d}x/\mathrm{d}t$ we obtain

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{n^* k_p}{2m_0\varphi_m}\left\{1 - \frac{\varphi_p}{\varphi_m} + \frac{\varphi_p/\varphi_m}{1-x}\right\}^{-1}$$

$$= \frac{n^* k_p}{2v_0}\left\{1 - \varphi + \frac{\varphi}{1-x}\right\}^{-1},$$

where $v_0 = \varphi_m m_0$ is the initial volume of monomer and $\varphi = \varphi_p/\varphi_m$. Integration yields:

$$(1-\varphi)x - \varphi \ln (1-x) + C = \frac{n^* k_p}{2v_0}\,t.$$

To eliminate $C$ we apply the condition that $t = t_{1/2}$ as $x = \frac{1}{2}$, which yields

$$(1-\varphi)\,(x - \tfrac{1}{2}) - \varphi \ln\{2(1-x)\} = \frac{n^* k_p}{2v_0}\,(t - t_{1/2})$$

$$= k\,(\tau - 1)\,,$$

putting

$$k \equiv \frac{n^* k_p t_{1/2}}{2v_0}.$$

When replacing $\varphi$ by $0{\cdot}946/1{\cdot}125 = 0{\cdot}84$ and dividing both sides of the equation by $k$, we ultimately obtain

$$\Psi_2(x, k) \equiv \frac{0{\cdot}16}{k}\,(x - \tfrac{1}{2}) - \frac{0{\cdot}84}{k} \ln\{2(1-x)\} + 1 = \tau,$$

which is the formula already used in section 3 of part II, where it is seen that it is fairly well confirmed by all the experiments mentioned, thus not only lending support to the stationary-state theory of Smith and Ewart but also justifying our assumptions for the propagation reaction.

Nevertheless there is an element of surprise in this for among the actual experiments there are several where $t_{1/2} - t^* \ll 1{\cdot}6\ t^*$. Obviously the respective systems have obtained the stationary state already shortly after the end of the incubation period *). On account of the great similarity in the behaviour of the various conversions it seems that, as regards the value of $|n_1/n^*|_{t=t^*}$, all the systems are to be found in corresponding states. This means that the boundary condition $|n_1/n^*|_{t=t^*} = 1$, introduced according to the initiation theory of Smith and Ewart, does not hold for any of the systems under consideration but may be far less than unity.

## 2. The number of latex particles

Investigations carried out in the section devoted to the stationary state led to an expression which is easily transformed into

$$n^* = \frac{2kv_0}{k_p t_{1/2}}, \tag{1}$$

a relation perfectly suited to compute from purely kinetic data relative numbers of particles, since both $k$ and $t_{1/2}$ can be found experimentally; assuming a value for $k_p$, even absolute numbers may be determined.

When only the amounts of soap ($S$) and catalyst ($R$) are varied, $k$ may to a first approximation be considered constant, whilst

$$t_{1/2} = \text{const. } S^{-3/5} R^{-1/5}.$$

From this it immediately appears that

$$n^* = \text{const. } S^{3/5} R^{1/5},$$

a result different from the theoretical relation derived by Smith and Ewart

$$n^* = \text{const. } S^{3/5} R^{2/5}.$$

Table IX gives particle numbers calculated from the actual values of $k$ and $t_{1/2}$, $v_0$ taken to be 56 ml and $k_p = 6{\cdot}10^{-18}$ ml.min$^{-1}$, a value taken from Burnett **).

*) The period during which the latex particles are formed and which ends at $t^*$ when the micelles disappear, or alternatively the number of particles no longer increases, will be called "incubation" period, rather than "initiation" period.

**) From photo-initiated bulk polymerization of styrene Burnett [25]) found

$$k_p = 1{\cdot}44.10^6 \exp(-3150/T)\ \text{l.mole}^{-1}.\text{sec}^{-1}.$$

TABLE IX

Number of latex particles times $10^{-17}$

| R \ S | 3 | 4 | 5 | 6 | 7 | 10 | 15 |
|---|---|---|---|---|---|---|---|
| 0·3 | 2·9 | 3·0 | 3·6 | 3·8 | 4·2 | 4·7 | 5·4 |
| $0{\cdot}6_5$ | 3·9 | 4·3 | 5·6 | 6·5 | | | |
| 1·0 | 4·8 | 6·0 | 6·3 | 7·5 | | 8·9 | 12 |

Attention is drawn to the fact that $n^*$ is the total number of particles contained in the system; to obtain the number per ml the values given in the table have to be divided by about 500. The numbers so found are of the order of $10^{14}$ to $10^{15}$ particles, in good agreement with the number that is generally supposed to be present [4]).

Another method for determining the number of latex particles from kinetic data, largely based upon Smith and Ewart's initiation theory, will now be developed. Smith and Ewart [5]) derived a relation between the total number of latex particles formed during the incubation period, which ends at the moment $t^*$, and the amounts of soap and catalyst used. They followed two lines of approach, both giving the function $n^* = \text{const.}\ S^{3/5}R^{2/5}$ and only differing as regards the numerical constant. Their derivations are based on the following assumptions: at the moment $t^*$, when the micelles disappear, all the available soap is adsorbed at the surface of the latex particles in a monomolecular, close-packed layer; the growth in volume of a latex particle once formed is constant or, to put it in another way, the kinetic activity of the monomer in the particles does not vary during $t < t^*$. The simplest of the derivations is moreover based on the assumption that the rate at which the particles are formed is constant up to $t^*$ and equal to the amount of radicals produced per unit of time. This implies at the same time that a latex particle, once formed, grows uninterruptedly up to $t^*$. This simplification, leading to a particle number which is certainly too high and which may be regarded as the uppermost limit, was avoided in the other derivation where the particles once formed serve also as a target for the radicals. Nevertheless, here too some approximations had to be made, permitting only the calculation of the lowest limit for the particle number. The proportion of these limiting numbers was found to be 0·53/0·37, so that

$$n^*_{\min} = 0{\cdot}70\ n^*_{\max}\ .$$

The main difference between the two methods lies in the extent to which the distribution of the particle sizes is taken into account, and in the following it will be shown that a simple calculation, assuming that on the average the particles have equal sizes, leads to a number $\bar{n}^*$ lying halfway between the two limiting values, viz.

$$\bar{n}^* = 0{\cdot}86\, n^*_{\max}\,.$$

Thus it may be taken that $\bar{n}^*$ will never deviate by more than 10% from the actual number of particles. To find the relation between $\bar{n}^*$ and $n^*_{\max}$ only the first calculation of Smith and Ewart, leading to the number $n^*_{\max}$ has to be used.

During a short period $dt$ the number of radicals generated will be $\varrho dt$, each of them converting a micelle into a latex particle. A micelle struck at the time $t$ and growing unchecked up to the time $t^*$ contains an amount of polymer

$$\delta p^* = k_p c\,(t^* - t)\,;$$

$c$ is the activity of the monomer in the following taken to be identical with the volume concentration. The volume of this particle will be

$$\delta v^* = \frac{\varphi_p \delta p^*}{1 - \varphi_m c},$$

because

$$\delta v^* = \varphi_m \delta m^* + \varphi_p \delta p^* \quad \text{and} \quad \delta m^* = c\,\delta v^*,$$

where $\delta m^*$ is the amount of monomer in the particle. This volume is therefore

$$\delta v^* = \frac{\varphi_p k_p c}{1 - \varphi_m c}\,(t^* - t)\,;$$

the factor $\varphi_p k_p c/(1 - \varphi_m c)$ is identical with the constant $\mu$ used by Smith and Ewart. Now the surface area of the particle will be

$$\begin{aligned}\delta a^* &= (36\pi)^{1/3}\,\delta v^{*2/3} \\ &= (36\pi)^{1/3}\,\mu^{2/3}\,(t^* - t)^{2/3}.\end{aligned}$$

The $\varrho dt$ particles formed between the times $t$ and $t + dt$ will have a surface area

$$da^* = (36\pi)^{1/3}\,\mu^{2/3}\,\varrho(t^* - t)^{2/3}\,dt,$$

which, integrated between the limits $t = 0$ and $t = t^*$, yields the total surface area of the latex particles at $t^*$:

$$a^* = (36\pi)^{1/3}\,\tfrac{3}{5}\,\mu^{2/3}\,\varrho t^{*5/3},$$

and, since $\varrho t^* = n^*_{\max}$ (the total number of latex particles),

$$a^* = \tfrac{3}{5}\,(36\pi)^{1/3}\,(\mu/\varrho)^{2/3}\,n^{*2/3}_{\max}. \qquad (2$$

This is the equation derived by Smith and Ewart with their first method of calculation.

In analogy with this derivation we found a second relation which enabled us to eliminate $\mu/\varrho$ from (2):

$$\delta v^* = \mu(t^* - t)$$

and consequently

$$\mathrm{d}v^* = \mu\varrho(t^* - t)\,\mathrm{d}t,$$

which, integrated between $t = 0$ and $t = t^*$, yields the total volume of the particles

$$v^* = \tfrac{1}{2}\,\mu\varrho t^{*2} = \tfrac{1}{2}\,(\mu/\varrho)\,n^{*2}_{\max}.$$

Comparison with (2) leads to the simple relation

$$a^* = \tfrac{3}{5}\,2^{2/3}\,(36\pi)^{1/3}\,v^{*2/3}\,n^{*1/3}_{\max}$$

or

$$n^*_{\max} = \left(\frac{5}{3}\right)^3 \frac{1}{2^2} \cdot \frac{a^{*3}}{36\pi\, v^{*2}} = 1{\cdot}16\,\bar{n}^*$$

or

$$\bar{n}^* = 0{\cdot}86\,n^*_{\max}.$$

By taking $\bar{n}^*$ for the actual number of particles we were able to work out all our further calculations.

The surface area $a^*$ is easily obtained from the amount of soap since if the area occupied by one molecule is $\omega_0$ then $a^* = \omega_0\,(N/W)\,S$, where $N$ is the Avogadro number and $W$ is the molecular weight of the soap.

The next question to be answered is how to find the value of $v^*$. It is readily seen that

$$v^* = \frac{\varphi_p p^*}{1 - \varphi_m c^*}$$

and since $p^* = x^* m_0 = x^*\, v_0/\varphi_m$ and $\varphi_p/\varphi_m = \varphi$ we get

$$v^* = \frac{\varphi x^*}{1 - \varphi_m c^*}\, v_0.$$

Now $c^*$ is to be found from kinetic data in the following way. At the end of the incubation period the total number of particles is $n^*$ and experience with the reaction rates has made it probable that, as regards $|n_1/n^*|_{t=t^*}$, the systems are to be found in corresponding states. All that remains is to ascertain the value of $|n_1/n^*|_{t=t^*}$: in the calculation to follow we have deliberately put this at $\frac{1}{2}$, although a higher value might

be preferred *). Hence

$$\left(\frac{dx}{dt}\right)^* = \frac{n^* k_p c^*}{2m_0}$$

or

$$\left(\frac{dx}{d\tau}\right)^* = \frac{n^* k_p t_{1/2}}{2v_0}\,\varphi_m c^* = k\varphi_m c^*.$$

For the conversions given in table I we found that during times $< t_{1/2}$ th relation

$$\left(\frac{x}{1-x}\right)^{1/2} = \tau$$

was fairly well satisfied, with the result that

$$\frac{dx}{d\tau} = 2x^{1/2}(1-x)^{3/2}$$

and thus

$$\varphi_m c^* = \frac{2}{k}\, x^{*1/2}\,(1-x^*)^{3/2}.$$

Consequently

$$a^* = (36\pi)^{1/3}\, v^{*2/3}\, n^{*1/2}$$

now becomes

$$a^* = (36\pi)^{1/3}\, \varphi^{2/3} \left\{\frac{kx^*}{k-2x^{*1/2}(1-x^*)^{3/2}}\right\}^{2/3} v_0^{2/3}\, n^{*1/2}.$$

Instead of determining another set of particle numbers with this equation, we shall eliminate $n^*$ from this and equation (1), thus leaving a relation between quantities known from literature or found from experiments. This relation, after replacing $a^*$ by $\omega_0(SN/W)$, then takes the following form

---

*) A rough calculation for the case $S = 6$, $R = 0{\cdot}3$, where $\tau^* = 0{\cdot}7$ or $t_{1/2} - t^* \approx \frac{1}{2}\, t^*$, gave $|n_1/n^*|_{t=t^*} \approx 0{\cdot}6$. On the other hand the derivations just performed, where we took $n^* = \bar{n}^* = 0{\cdot}86\; n^*_{max}$ and where $n^*_{max} = \varrho\, t^*$ the number of radicals generated during the incubation period, lead to $|n_1/n^*|_{t=t^*} \approx 0{\cdot}8$.

As we shall see later on, the actual number of radicals generated during the incubation period is about 4 times the number of latex particles formed. Although the major part of this excess may disappear through processes other than the polymerization there is a good chance that a minor part favours the approach to the stationary state. This would be all the more likely considering that during the incubation period the latex particles presumably have a lower potential barrier than the micelles, so that the ionic $SO_4^-$ radicals may be expected to show a slight preference for the latex particles.

By putting $|n_1/n^*|_{t=t^*} = \frac{1}{2}$ the final results of our further calculations work out at about 10% higher than when putting this at 0·8.

IPR2025-00217 EXHIBIT 2010 Page 387 of 485

$$\omega_0 = (72\pi)^{1/3} \frac{W}{N} \frac{\varphi^{2/3} v_0 f^{*2/3}}{S k_p^{1/3} t_{1/2}^{1/3}}, \tag{3}$$

where

$$f^* = \frac{k^{3/3} x^*}{k - 2x^{*1/3}(1 - x^*)^{3/3}}.$$

Fig. 10 shows two examples of $f^*$ in dependency of $x^*$, namely for $k = 1{\cdot}0$ and $k = 1{\cdot}25$, the extreme values of this parameter so far encountered.

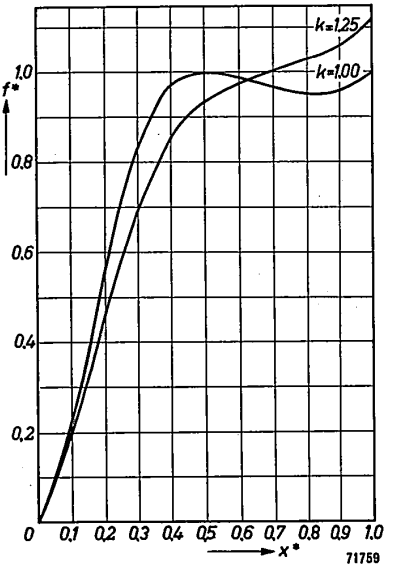



Fig. 10. The function $f^*(x^*, k)$ for $k = 1{\cdot}0$ and $k = 1{\cdot}25$.

When taking $k_p = 6.10^{-18}$ ml.min$^{-1}$, $N = 6.10^{23}$, the molecular weight of potassium myristate $W = 266$ and expressing $\omega_0$ in Å$^2$, then

$$\log \omega_0 = 2{\cdot}871 - \log S - \tfrac{1}{3} \log t_{1/2} + \tfrac{2}{3} \log f^* ,$$

$v_0$ being in all these calculations 56 ml and $\varphi = 0{\cdot}84$. The values of $\omega_0$ computed from this equation are given in table X.

The average value $\omega_0 = 35 \pm 6$ Å$^2$ proves to be in good agreement with the value given by Harkins [4]), which is about 36 Å$^2$. It is quite clear that in principle this method might be used for determining $k_p$. The value of $k_p$ used here is found from bulk polymerizations. Smith [30]) obtained a value 4·5 times larger from emulsion polymerizations in a manner quite different

from that outlined above. The value of $\omega_0$ calculated when using Smith's $k_p$ is 0·6 times the value that we have so far found, thus 21 Å$^2$. This differs appreciably from Harkins's value, but in view of the inaccuracies inherent in our assumptions this does not seem to be the appropriate place to decide whether the $k_p$ in emulsion polymerizations does indeed differ from the $k_p$ in bulk polymerizations, as suggested by Burnett [25]).

TABLE X

Values of $\omega_0$ in Å$^2$

| R \ S | 3 | 4 | 5 | 6 |
|---|---|---|---|---|
| 0·3 | 38 | 33 | 30 | 25 |
| 0·65 | 36 | 46 | 30 | 29 |
| 1·0 | 37 | 41 | 39 | 31 |

Recapitulating, it is perhaps not too bold to assume that the operations performed corroborate what has been postulated concerning the stationary state and the limiting effect which the amount of soap may have upon the number of latex particles.

### 3. Length of the incubation period

In a former section it was already pointed out that the conversions with more than 6 g soap are in so far incomparable with the others in that probably not all the micelles are used up. The incubation period should then end owing to the solubilized monomer in the micelles existing after 50 % conversion being extracted by the latex particles, which have also absorbed all the emulsified monomer, and consequently the micelles are deprived of their power to form new latex particles. In such cases it is not the soap which limits the number of latex particles, but the typical physico-chemical conditions.

We already have some indications that the conditions outlined above may be in conformity with our experience with these conversions. First of all some indication is to be found in the conductivity measurements. When using more than 6 g soap we found that the conductivity curve showed a knee situated in most cases between $x = 0{\cdot}35$ and $x = 0{\cdot}5$, where a much less steep, though quite linear, decline sets in often extending beyond $x = 0{\cdot}8$ after which it changed arbitrarily as already sketched before (fig. 9). Secondly, calculations as carried out in the previous section lead to values of $\omega_0$ which are certainly far too low (table XI). From the curious shape of the $f^*(x^*, k)$ curve it is easily understood that it is immaterial whether we

use the value of $x^*$ as determined by the knee or the much higher value determined from the moment where the arbitrary change sets in. From these values we can deduce how much of the soap has been effective in forming latex particles, namely $S_{\text{eff}} = (\omega_0/35)\,S$. The figures found are as follows (table XII). It is seen that on the average no more than 7 g of soap is used.

TABLE XI

Values of $\omega_0$ for $S>7$

| $S$ | 10 | | 15 | |
|---|---|---|---|---|
| $R$ | 0·3 | 1·0 | 0·3 | 1·0 |
| $\omega_0$ | 20 | 15 | 24 | 19 |

TABLE XII

Values of $S_{\text{eff}}$ for $S>7$

| $S$ | 10 | | 15 | |
|---|---|---|---|---|
| $R$ | 0·3 | 1·0 | 0·3 | 1·0 |
| $S_{\text{eff}}$ | 6 | 7 | 6·5 | 8 |

A third argument in support of our supposition will now be discussed. Elimination of $t_{1/2}$ from equation (3) of the preceding section, with the aid of

$$t_{1/2} = \text{const.}\ (SR)^{-1/2}$$

yields

$$\omega_0 = \text{const.}\ \frac{(SR)^{1/4}}{S}\, f^{*1/2}.$$

Hence we obtain

$$f^* = \text{const.}\ S^{3/4}\, R^{-1/4}.$$

This function provides the possibility to find values for $x^*$ assuming that this value is known for a particular case. Let us take for this $S = 7$, $R = 0{\cdot}3$ and let $x^* = 0{\cdot}5$. Then, if we assume for this and all the other cases $k = 1{\cdot}10$, $|f^*|_{S=7,\,R=0{\cdot}3}$ will be equal to 0·96 and generally

$$f^* = 0{\cdot}96 \left(\frac{0{\cdot}3}{7^5}\right)^{1/4} S^{3/4}\, R^{-1/4}$$

$$= 0{\cdot}062\, S^{3/4}\, R^{-1/4}.$$

When at least $R$ is kept $= 0{\cdot}3$ g, amounts of soap $> 7$ g give rise to values of $f^*$ not only belonging to values of $x^* > 0{\cdot}5$, but even surpassing every value which $f^*$ may ever attain. This clearly shows that the condition used for the derivation of this function, namely that the amount of soap is the factor limiting the number of latex particles, no longer applies.

The following values of $t^*$ were calculated *) for amounts of $S < 7$ and compared with the experimental ones (table XIII).

*) The calculations were made as follows. The value of $x^*$ may be determined from a graph for any $f^*$, obtained for a couple of particular values of $S$ and $R$; with the aid of the function $\Psi_1$ a corresponding value of $\tau^*$ is to be found and then $t^*$ is calculated from $\tau^*$ and the $t_{1/2}$ of the particular case.

TABLE XIII

| | Calculated values of $t^*$ | | | | | Experimental values of $t^*$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| $R$ \ $S$ | 3 | 4 | 5 | 6 | 7 | 3 | 4 | 5 | 6 |
| 0·3 | **30** | **31** | $33_5$ | $36_5$ | *49* | **33** | **32** | **34** | 30 |
| 0·65 | 19 | $19_5$ | **20** | **21** | $22_5$ | 23 | 28 | **21** | **22** |
| 1·0 | 15 | 15 | $15_5$ | **16** | *17* | 19 | 21 | 21 | **17** |

Half of the calculated values are in reasonable agreement with the experimental figures. For the other half the discrepancies sometimes attain serious dimensions, for instance in the case of $S = 4$, $R = 0{\cdot}65$. This indicates that not all the values of $t^*$ found from the conductivity measurements are as reliable as was at first deduced from the standard deviation.

From Smith and Ewart's analysis of the incubation period the conclusion can be drawn that

$$t^* = \text{const. } S^{2/5} R^{-3/5}.$$

It is obvious that neither the experimental, nor the calculated values of $t^*$ obey this equation in every respect. As long as $S < 6$ we find

$$t^* = \text{const. } R^{-3/5}$$

for the calculated $t^*$.

Here it seems appropriate to make a remark concerning the efficiency with which the radicals are used in making the latex particles. From some experiments *) we have obtained for the rate with which the radicals are produced at 40 °C

$$\varrho = 12.10^{16} R \text{ min}^{-1}.$$

During the incubation period $\varrho t^*$ radicals are generated and $n^*$ latex particles formed, thus $n^*/\varrho t^*$ is the fractional part of the number of radicals used for making latex particles. On the average one of out of four radicals generated during the incubation period gives rise to the formation of a latex particle. (Table XIV).

The last three columns of table XIV show the mean time elapsing between the entry of two successive radicals into the same particle during the stationary period. On the whole this time interval — which is of course a minimum value — is about 7 min.

*) Mentioned on the third page of section III, 1.

TABLE XIV

| | Particle number . $10^{-16}$ | | | Duration of incubation period in min | | | Efficiency of the radicals | | | Time elapsing between two successive entries of radicals into a particle | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| S \ R | 0·3 | 0·65 | 1·0 | 0·3 | 0·65 | 1·0 | 0·3 | 0·65 | 1·0 | 0·3 | 0·65 | 1·0 |
| 3 | 29 | 39 | 48 | 33 | 23 | 19 | 1/4 | 1/5 | 1/5 | 8 | 5 | 4 |
| 4 | 30 | 43 | 60 | 32 | 28 | 21 | 1/4 | 1/5 | 1/4 | $8_5$ | $5_5$ | 5 |
| 5 | 36 | 56 | 63 | 34 | 21 | 21 | 1/3 | 1/3 | 1/4 | 10 | 7 | 5 |
| 6 | 38 | 65 | 75 | 30 | 22 | 17 | 1/3 | 1/3 | 1/3 | $10_5$ | 8 | 6 |

## 4. The initial stages of the reaction

It remains to elucidate the dependency of $t_{1/2}$ upon the amounts of soap or catalyst, a difficult task, in so far that the conditions during the incubation period, certainly determining the value of $t_{1/2}$, are too intricate to obtain any result without making several simplifications, as already discussed. To avoid speculations as to the conditions prevailing during the incubation period, only the earliest stage is considered and this has proved to be sufficient. The following general equation for the reaction velocity already used,

$$\frac{dx}{dt} = n_1 \frac{k_p c}{m_0}, \tag{1}$$

provides a relation between the conversion rate and $n_1$, the number of active latex particles, a number which can only be readily established when we have to do with the stationary state. But we may assume that the shorter the time elapsing since the reaction started, the more closely the number of active particles will approach the total number of particles produced during that interval. Now the reaction by which the latex particles are generated may be formulated as follows:

$$R + S(oapmicelle) \xrightarrow{k_i} L(atex\ particle)$$

This equation needs some further explanation, for everyone would assume that the catalyst first decomposes into radicals, which afterwards initiate micelles and force them to develop into latex particles. Several experiments have shown that in pure water the decomposition of potassium persulphate is strictly monomolecular, indicating that combination of the primary radicals can be left out of consideration.

$$S_2O_8^{2-} \xrightarrow{k_d} 2SO_4^{\cdot-}$$

and further $$SO_4^{\cdot-} + H_2O \xrightarrow{k_a} HSO_4^- + OH^{\cdot},$$
$$2\,OH^{\cdot} \longrightarrow H_2O + O.$$

This gives rise to a quasi-stationary concentration of the sulphate radicals in pure water

$$[SO_4^{\cdot-}]_{aq} = \frac{2k_d}{k_a}[S_2O_8^{2-}].$$

When persulphate solution is added to an emulsion of the styrene the decomposition reaction will probably not be influenced and our number of reactions will only be extended by

$$SO_4^{\cdot-} + S(oapmicelle) \xrightarrow{k_i'} L(atex\ particle)$$

bearing in mind that the hydroxyl radicals are rather ineffective for the initiation, as found for instanee by Bartlett [32]).

The equation for the rate of latex particle formation may now be written as

$$\frac{d[n_1]}{dt} = k_i'[S]_{mic.}[SO_4^{\cdot-}],$$

where $[S]_{mic.}$ denotes the concentration of the micelles. Thus in the very first instants of the reaction, to which our further consideration will be confined,

$$\frac{d[n_1]}{dt} = \frac{2k_d}{k_a} k_i'\,[S]_{mic.}[S_2O_8^{2-}].$$

Putting $(2k_d/k_a)k_i' \equiv k_i$ we obtain the rate equation which could have been derived directly from the initiation reaction formulated above.

To derive an expression which relates the original concentration of micelles to the amount of soap, we assume that the whole of this amount is available for their production. When each micelle is built up of $\nu$ molecules the number of micelles is

$$\frac{N}{W\nu} S,$$

where $N$ is the Avogadro number and $W$ the molecular weight of the soap. The following differential equation is now valid for times short enough to assume that hardly any micelles have been used for supplying soap to the growing latex particles:

$$\frac{d}{dt}\left(\frac{n_1}{V}\right) = k_i \frac{N}{W\nu} \frac{S}{V} \frac{R}{V},$$

where $V$ is the volume of the aqueous phase. Integration yields

$$n_1 = k_i \frac{N}{W\nu} S \frac{R}{V} t,$$

and inserting this expression for $n_1$ in (1) we obtain

$$\frac{dx}{dt} = k_i \frac{N}{W\nu} \frac{SR}{m_0 V} k_p\, c\, t, \qquad (2)$$

from which it immediately becomes evident that

$$\left|\frac{dx}{dt}\right|_{t=0} = 0,$$

making it impossible to explain how, as is sometimes reported, the conversion rates can be constant right from the start.

Replacing $t$ by $\tau t_{1/2}$ and making some rearrangement, we obtain

$$\frac{1}{\tau}\frac{dx}{d\tau} = \frac{k_i k_p N}{W\nu} c \frac{SR}{m_0 V} t_{1/2}^2.$$

The only parameter of this equation influenced by a variation of $\tau$ is $c$, which for the case where $\tau$ is equal to zero becomes the concentration of the monomer in the undisturbed micelles, in the following denoted by $c_m$. Hence

$$\left|\frac{1}{\tau}\frac{dx}{d\tau}\right|_{\tau=0} = \frac{k_i k_p N}{W\nu} c_m \frac{SR}{m_0 V} t_{1/2}^2.$$

From this it is clear that $|(1/\tau)\,(dx/d\tau)|_{\tau=0}$, which is also equal to $|d^2x/d\tau^2|_{\tau=0}$, has a finite value, differing from zero and further on called $D$. Experimentally we found in the first stage of the reaction

$$\tau = \left(\frac{x}{1-x}\right)^{1/2},$$

hence

$$D = 2,$$

independent of $R$ and $S$, at least when the other conditions governing the polymerizations of table I remain unaltered. Thus $c_m$ must be expected to depend only on the kind of soap and not on the amount, for only then does $t_{1/2}$ become dependent upon $R$ and $S$, in the manner found experimentally, viz.

$$t_{1/2} = \left(\frac{D}{k_i k_p}\frac{W\nu}{N c_m} m_0 V\right)^{1/2} \left(SR\right)^{-1/2}. \qquad (3)$$

This formula, however, explicitly shows much more than we aimed at, since it seems to show that $t_{1/2}$ depends also upon $V$, $m_0$ and a number of factors typical for the kind of soap.

### 5. Influence of the initial amount of monomer and of the volume of the aqueous phase on $t_{1/2}$

Hitherto it has been commonly held that the volume of the aqueous phase has practically no influence on the reaction velocity, if care is taken not to use so much water as to go below the critical micelle concentration

(*C.M.C.*) of the soap. It was therefore very surprising to find that formula (3) of the previous section shows an explicit dependency of $t_{1/2}$ upon $V$. Some preliminary experiments were carried out to see whether it was indeed justified to abandon the former opinion. A new dilatometer vessel was built with a volume of 250 ml, which limits the volume of the aqueous phase to about 200 ml. At the same time we investigated the dependency of $t_{1/2}$ upon $m_0$. The results so far obtained are such that we find ourselves faced with a very intriguing problem, which we have not yet been able to solve.

Let us at first see how $t_{1/2}$ depends upon $m_0$. Experiments were carried out with 3 g and 4 g potassium myristate, in each case with 0·3 g $K_2S_2O_8$ and 425 ml water, but with about 25 g monomer. Full particulars about these conversions are to be found in table II. In table XV a comparison is made between the half-time values found for the various series of experiments.

TABLE XV

Influence of $m_0$ on $t_{1/2}$

| | $S = 3$, $R = 0{\cdot}3$ | | $S = 4$, $R = 0{\cdot}3$ | |
|---|---|---|---|---|
| $m_0$ | 25·5 | 49·6 | 23·0 | 49·5 |
| $t_{1/2}$ | 48 | 69 | $40_5$ | $64_5$ |

These results are in perfect agreement with the relation

$$t_{1/2} = \text{const.}\ m_0^{1/2},$$

for then

$$|t_{1/2}|_{S=3,\ m_0=25\cdot5} = 69\left(\frac{25\cdot5}{49\cdot6}\right)^{1/2} = 49_5$$

and

$$|t_{1/2}|_{S=4,\ m_0=23\cdot0} = 64_5\left(\frac{23\cdot0}{49\cdot5}\right)^{1/2} = 44\ .$$

The following experiments, with the volume of the aqueous phase halved, confirmed that $V$ has no influence on $t_{1/2}$. Particulars are to be found in tables II and XVI.

The dependency of $t_{1/2}$ upon the extensive parameters $m_0$, $V$, $R$ and $S$, can now be formulated as follows

$$t_{1/2} = \text{const.}\ m_0^{1/2}\, S^{-1/2}\, R^{-1/2}\, V^0.$$

TABLE XVI

Influence of the water volume on $t_{1/2}$

| | K myristate | | | |
|---|---|---|---|---|
| | $S = 3, R = 0{\cdot}3$ | | $S = 4, R = 0{\cdot}3$ | |
| $V$ | 400 | 200 | 400 | 200 |
| $t_{1/2}$ | 69 | $72_5$ | $64_5$ | $63_5$ |

However, $t_{1/2}$ and the constant are certainly intensive parameters of the system, but it seems that, as a consequence of this relation, by halving the system a new value of $t_{1/2}$ should be found equal to $\sqrt{2}$ times the $t_{1/2}$ of the original system, which is quite incredible. Of course we tried to establish such a curious behaviour, but we did not succeed. Table XVII gives an example.

TABLE XVII

Dependency of $t_{1/2}$ upon the external dimensions

| | $S = 4, R = 0{\cdot}3$ $m_0 = 49{\cdot}6, V = 400$ | $S = 2, R = 0{\cdot}15$ $m_0 = 22{\cdot}5, V = 225$ |
|---|---|---|
| $t_{1/2}$ | 64 | 66 |

This means that the expression which relates $t_{1/2}$ to $m_0$, $V$, $R$ and $S$ does not hold over the whole range covered by our experiments. We found namely for $m_0 = 25$, $R = 0{\cdot}3$, $S = 3$ and $V = 425$, $t_{1/2}$ to be 48 min, so we must expect for $m_0 = 25$, $R = 0{\cdot}3$, $S = 3$ and $V = 225$, $t_{1/2}$ to be also 48 min; we found, however, for $m_0 = 50$, $R = 0{\cdot}6$, $S = 6$ and $V = 400$, $t_{1/2}$ to be 37 min. While in the last case we did not do anything else than double the amounts of the second case, and another experiment has shown that this cannot effect $t_{1/2}$, we might have expected to find there also $t_{1/2}$ to be 48 min. So far we have no satisfactory explanation for these facts. *)

*) Supplementary experiments, not listed in table II, with $m_0 = 23$, $S = 4$, $R = 0{\cdot}3$ and $V = 225$ gave a $t_{1/2} = 33$ min, a value to be expected from that found for the experiments with $m_0 = 50$, $S = 4$, $R = 0{\cdot}3$ and $V = 400$, as long as the $(R\,S)^{-1/2}$ rule is not violated. The relations now arising for the several cases with $S = 4$ and $R = 0{\cdot}3$ are now such that going from $m_0 = 50$, $V = 400$ ($t_{1/2} = 64$) to $m_0 = 25$, $V = 225$ ($t_{1/2} = 33$) via $m_0 = 50$, $V = 200$ ($t_{1/2} = 64$) thus by diminishing firstly $V$ and then $m_0$, $t_{1/2}$ jumps proportionally to $V^0 m_0^1$; whereas going from the first to the second case via $m_0 = 25$, $V = 425$ ($t_{1/2} = 41$), thus by diminishing firstly $m_0$ and then $V$, $t_{1/2}$ jumps proportionally to $V^{1/2} m_0^{1/2}$. At last two experiments with about 75 g monomer, 175 ml water, 4 g K myristate and 0·3 g persulphate gave an average $t_{1/2} = 80$ min. Compared to the case $m_0 = 50$, $V = 200$, $S = 4$ and $R = 0{\cdot}3$ with a $t_{1/2} = 64$, this again indicates a jump of $t_{1/2}$ proportional to $m_0^{1/2}$.

## 6. Comparison with data taken from the literature

By far the richest source of information about the influence of the amounts of soap or catalyst is to be found in the publications by Kolthoff c.s.

We confine our attention to the experiments performed with mixtures of styrene and butadiene, which all show a non-linear behaviour of the first part of the conversion curve. In such cases Kolthoff was obliged to choose a new criterion for the reaction rate: the conversion reached within a certain fixed interval of time since the start, short enough to ensure that the conversion does not rank higher than a few tens per cent.

First of all we shall investigate whether the uniformities found for the stationary state in the case of pure styrene hold with these mixtures. From the material published by Kolthoff, Meehan and Carr [26]) examples have been selected such that the conversions could be measured from the diagrams up to 90 %. Estimating the specific volume of the monomer mixture to be about 1·5 ml.g$^{-1}$, whilst that of G.R.-S. rubber is 1·06 ml.g$^{-1}$, the following values for $\Psi_2(x, k = 1)$ have been calculated.

| $x$ | 0·5 | 0·6 | 0·7 | 0·8 | 0·9 |
|---|---|---|---|---|---|
| $\Psi_2(x,k=1)$ | 1·00 | 1·19 | 1·42 | 1·73 | 2·25 |

The graphs of fig. 11 have been plotted from the data given in table XVIII.

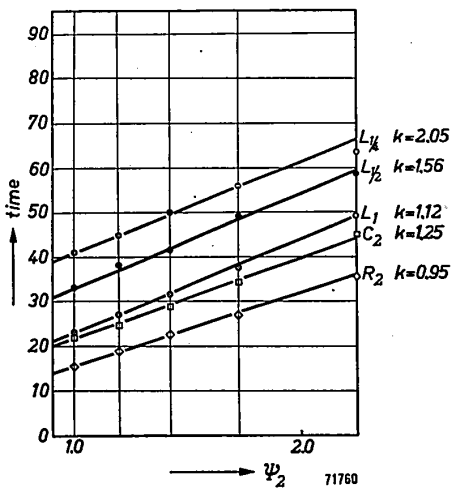



Fig. 11. Analysis of some G.R.-S. conversions (75 butadiene and 25 styrene) in terms of $\Psi_2(x, k)$. $L$ = laurate, $C$ = caprate, $R$ = Rubber Reserve soap; the indices represent the relative soap concentrations.

## TABLE XVIII

Data of G.R.-S. polymerizations with different kinds of soap

| | Sodium laurate | | | Sodium caprate | Sodium rosinate |
|---|---|---|---|---|---|
| $x$ \ $S$ | 1 | $\frac{1}{2}$ | $\frac{1}{4}$ | 2 | 2 |
| 0·5 | 23 | 33 | 41 | 23 | $15_5$ |
| 0·6 | 27 | 38 | 45 | 25 | 19 |
| 0·7 | $31_5$ | $42_5$ | 50 | 29 | $22_5$ |
| 0·8 | $37_5$ | 49 | 56 | $34_5$ | 27 |
| 0·9 | $49_5$ | $58_5$ | $63_5$ | 45 | $35_5$ |
| $k$ | 1·12 | 1·56 | 2·05 | 1·25 | 0·95 |
| $k/t_{1/_2}$ | 0·49 | 0·47 | 0·50 | 0·57 | 0·61 |

The time values in the table are expressed in multiples of 26 min. The amounts of soap are expressed as fractions or as multiples of the normal amount "1", which is 5 % soap relative to the amount of monomer.

The diagram shows that our arguments apply remarkably well to the case of G.R.-S. But at the same time we encounter a new element: in our own cases it was $k$ which proved to be independent of the amount of soap, here it is $k/t_{1/_2}$ which shows that behaviour, as is at once evident from the fact that the lines in the diagram run parallel. *) Now

$$\frac{k}{t_{1/_2}} = \frac{n^* k_p}{2v_0}$$

and the conclusion that $n^*$ is independent of the amount of soap is obvious. Kolthoff, on the contrary, when finding that the velocities beyond 50 % conversion are independent of the amount of soap — a conclusion likewise to be drawn from the parallel running of the lines in fig. 11 — assumes that as a consequence the velocities are independent of the number of

*) In connection herewith reference is made to the subscript of fig. 4*b*, where we find that

$$\frac{d\Psi_2}{dt} = \frac{k}{t_{1/_2}}$$

or

$$\frac{d\Psi_2}{dx}\frac{dx}{dt} = \frac{k}{t_{1/_2}},$$

hence

$$\frac{dx}{dt} = \frac{k}{t_{1/_2}}\left(\frac{d\Psi_2}{dx}\right)^{-1} = \frac{n^* k_p}{2v_0}\left(\frac{d\Psi_2}{dx}\right)^{-1}.$$

latex particles present, and from the results of his seeding experiments he seems to be right[27]). Although it is highly probable that the rate-determining factors of the G.R.-S. system are different from those of the styrene system, both systems can be described, at least for conversions exceeding 50 %, with the same relation between $t$ and $x$, which is rather peculiar. In connection herewith it may be stressed that the conversions below 50 % cannot at all be described with the simple relation found for styrene.

The next point to discuss is the relation between $t_{1/2}$ and the amount of soap. In fig. 12 curves have been plotted for $t_{1/2}$ as a function of the relative soap concentration for different kinds of soap, using the values of table XIX.

TABLE XIX

Data concerning G.R.-S. polymerizations, taken from Kolthoff, Meehan and Carr[26])

| | R.R. soap | | | | | | Na rosinate | | | | | Na myristate | | | Na caprate | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $S$ | 2 | 1 | 1/2 | 1/4 | 1/8 | 1/16 | 2 | 1 | 1/2 | 1/4 | 1/8 | 1 | 1/2 | 1/4 | 2 | 1 |
| $t_{1/2}$ | $10_5$ | 17 | 25 | 36 | 45 | 57 | $15_5$ | $23_5$ | 33 | $45_5$ | $54_5$ | 20 | 30 | 41 | 23 | 32 |
| $\Delta x$ | 0·33 | 0·18 | 0·09 | 0·05 | 0·03 | $0·01_5$ | 0·15 | 0·08 | 0·04 | $0·01_5$ | 0·01 | 0·12 | 0·06 | $0·02_5$ | | |
| $G$ | 1240 | 620 | 305 | 155 | 75 | 35 | 980 | 350 | 160 | 75 | 35 | 385 | 160 | 50 | | |
| fig. 12 | $t_{1/2} \sim S^{-0·5}$ (a) | | | | | | $t_{1/2} \sim S^{-0·5}$ (c) | | | | | $t_{1/2} \sim S^{-0·5}$ (b) | | | | |
| fig. 13 | $\Delta x \sim S^{0·9}$ (a) | | | | | | $\Delta x \sim S^{1·0}$ (c) | | | | | $\Delta x \sim S^{1·10}$ (b) | | | | |
| fig. 14 | $\Delta x \sim G^{0·84}$ (b) | | | | | | $\Delta x \sim G^{0·86}$ (c) | | | | | $\Delta x \sim G^{0·76}$ (a) | | | | |

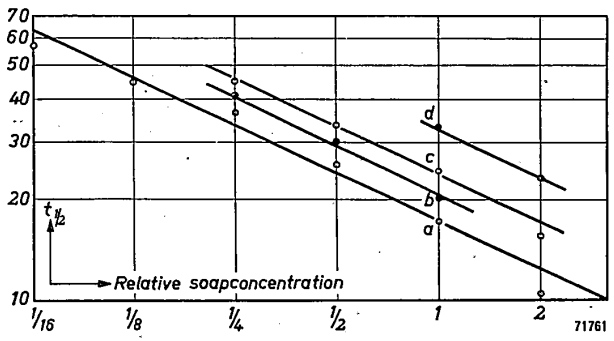



Fig. 12. Dependency of $t_{1/2}$ upon the soap concentration, for G.R.-S. polymerizations. (*a*) Rubber Reserve soap, (*b*) Sodium myristate, (*c*) Sodium rosinate, (*d*) Sodium caprate

Here again the dependency of the half-time values upon the amounts of soap proves to be

$$t_{1/2} = \text{const. } S^{-1/2},$$

just as is the case with pure styrene.

According to Kolthoff the initial velocity should be directly proportional to the amount of originally solubilized monomer, which will be denoted by $G$. At first sight there seems to be no doubt as to the tenability of this assumption, but a closer examination reveals that the deviations are greater than could be ascribed to mere accident. Let $\Delta x$ be the conversion reached after three hours and, following Kolthoff, taken to represent the initial velocity. From table XIX and fig. 14 it then appears that

for R.R. soap $\Delta x = \text{const. } G^{0.84}$,
for sodium rosinate $\Delta x = \text{const. } G^{0.86}$ and
for sodium myristate $\Delta x = \text{const. } G^{0.76}$.

On the other hand, plots of $\Delta x$ against $S$ (fig. 13) reveal that

for R.R. soap $\Delta x = \text{const. } S^{0.9}$,
for Na rosinate $\Delta x = \text{const. } S^{1.0}$ and
for Na myristate $\Delta x = \text{const. } S^{1.10}$, so that on the average $\Delta x$ is indeed directly proportional to $S$ but not to $G$.

To obtain a basis for further reasoning equation (2) in the previous section will be integrated, assuming, with regard to the small conversions to which our considerations are confined, that $c$ remains constant and equal to $c_m$. Then

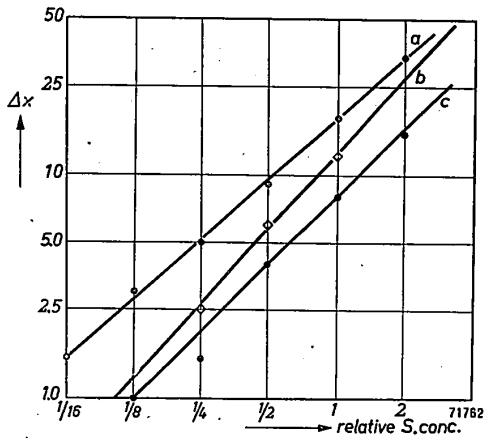



Fig. 13. Dependency of $\Delta x$ (in %) upon the soap concentration for G.R.-S. polymerizations. (*a*) Rubber Reserve soap, (*b*) Sodium myristate, (*c*) Sodium rosinate.

$$\Delta x = \frac{k_i k_p N}{2 W v} c_m \frac{R S}{m_0 V} (\Delta t)^2, \qquad (1)$$

$\Delta t$ being fixed at three hours. We now put $c_m = \gamma/\chi$, where $\gamma$ may be the amount of monomer solubilized by a single micelle and $\chi$ a proportionality factor. But $N/Wv$ is the number of micelles built up from one gram of soap, so that

$$\gamma \frac{N}{Wv} = g$$

is the amount of monomer solubilized per gram of soap. The total amount of solubilized material $G$ thus equals $gS$. Equation (1) can now be modified into

$$\Delta x = \frac{k_i k_p}{2} \frac{g}{\chi} \frac{R}{m_0 V} (\Delta t)^2 S$$

or alternatively into

$$\Delta x = \frac{k_i k_p}{2} \frac{1}{\chi} \frac{R}{m_0 V} (\Delta t)^2 G.$$

Now the former equation proved to be more readily satisfied than the latter, so we must conclude that varying amounts of soap have no influence upon the factor $g/\chi$ — which is presumably equivalent to $c_m$ — whereas they certainly influence $g$ and $\chi$ separately. The dependency of $g$ upon the dilution for several soap solutions is depicted in the diagram of fig. 15, the material from which it is composed being entirely taken from Kolthoff.

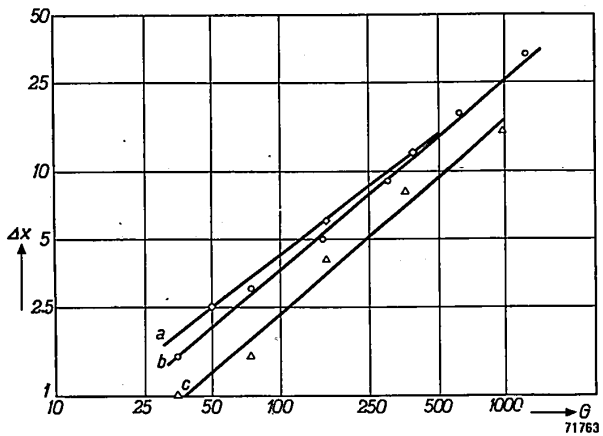



Fig. 14. Dependency of $\Delta x$ (in %) upon $G$. (*a*) Sodium myristate, (*b*) Rubber Reserve soap, (*c*) Sodium rosinate.

In connection herewith it may be of some interest to investigate the influence of the kind of soap upon $t_{1/2}$. We use the formula (3) from the previous section, modifying it in the same manner as we did with formula (1). This yields

$$t_{1/2}^2 = \frac{D}{k_i k_p} \frac{\chi}{g} \frac{m_0 V}{R S}.$$

Unfortunately we have to restrict this investigation to our own experiments, for in these cases the factor $D$ was easily obtained. In the experiments with styrene-butadiene mixtures, on the contrary, $D$ is difficult to evaluate from the diagrams.

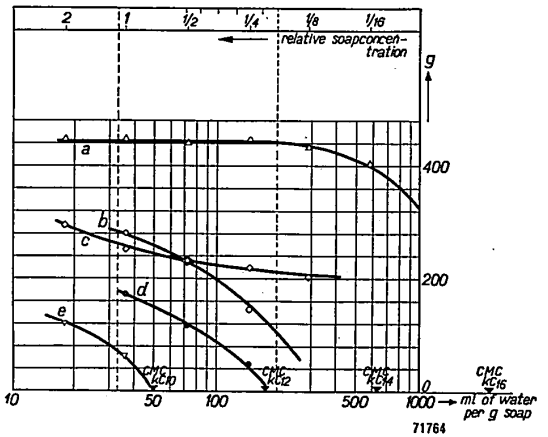



Fig. 15. Dependency of the solubilization upon the dilution of the soap solution. (*a*) Rubber Reserve soap, (*b*) Sodium myristate, (*c*) Sodium rosinate, (*d*) Sodium laurate, (*e*) Sodium caprate. *C.M.C.* indicates the critical micelle concentration.

First a comparison is made between myristate and laurate

| | Laurate | Myristate |
|---|---|---|
| $S$ | 5 | 5 |
| $R$ | 0·65 | 0·65 |
| $t_{1/2}$ | 54 | 38 |
| $g$ | 94 | 194 |

$S$ = amount of soap in g,
$R$ = amount of catalyst in g,
$t_{1/2}$ = half-time value in min,
$g$ = amount of solubilized styrene in mg per g soap (fig. 15),
$(t_{1/2})_L : (t_{1/2})_M = 1{\cdot}4$,
$(g_M : g_L)^{1/2} = 1{\cdot}4$.

In this case the influence of the kind of soap can be solely ascribed to the solubilization, thus $\chi$ retains the same value both for laurate and for myristate.

The next comparison is between stearate and myristate

| | Stearate | | Myristate | |
|---|---|---|---|---|
| $S$ | 5 | 10 | 5 | 10 |
| $R$ | 0·3 | 0·3 | 0·3 | 0·3 |
| $t_{1/2}$ | 53 | 39 | 57 | 42 |
| $g$ | 446 | 446 | 194 | 256 |

$(t_{1/2})_{St} : (t_{1/2})_M = 0{\cdot}93$, $(g_M : g_{St})^{1/2} = 0{\cdot}66$ } $S = 5$

$(t_{1/2})_{St} : (t_{1/2})_M = 0{\cdot}93$, $(g_M : g_{St})^{1/2} = 0{\cdot}76$ } $S = 10$

The influence of the kind of soap on $t_{1/2}$ can be expressed by a fixed number, but in this case no relation to the solubilization can be traced and therefore $\chi$ must vary when changing over from stearate to myristate. From this considerations it appears that not only $g$ and $\chi$ but also $g/\chi$ mostly depend upon the kind of soap; but whereas $g$ and $\chi$ also mostly depend upon the dilution, $g/\chi$ never does. Unfortunately there is no straightforward method available for determining $g/\chi$, or alternatively $c_m$, although this quantity seems to be more appropriate than $g$ for the discussion of the velocities.

**Acknowledgement**

The author is greatly indebted to all who have helped towards the completion of this investigation, particularly Professor Dr J. Th. G. Overbeek for his guidance and advice at all times, Drs J. J. Zaalberg van Zelst and K. H. Klaassens of this laboratory, both of whom have yielded valuable contributions. Thanks are no less due to those who undertook so much of the practical work, namely Mr P. van den Ban, who was a great help during the first difficult stages; Mr J. Goorissen to whose experimental skill the perfection of the methods is in no small degree due; and last but not least Miss M. A. Huijsman and Mr C. J. Schoot, who performed a great many of the measurements. The fruitful discussions with a number of colleagues have likewise been of inestimable value to the author. Finally the author wishes to express his sincere thanks to Dr E. J. W. Verwey, Director Philips Research Laboratories, for his help and the interest shown by him in the work.

*Eindhoven, July 1952*

## LIST OF SYMBOLS

$a^*$ total surface area of latex particles at $t^*$.
$c$ concentration of monomer in a latex particle.
$c^*$ concentration of monomer in a latex particle at $t^*$.
$c_m$ concentration of monomer in undisturbed micelles.
$D$ numerical value of $|(1/\tau)\cdot(dx/d\tau)|_{\tau=0}$.
$g$ weight of monomer solubilized by one gram of soap.
$G$ total weight of solubilized monomer.
$k$ proportionality constant equal to $n^*k_p t_{1/2}/2v_0$.
$k_i$ rate constant of the initiation reaction.
$k_p$ rate constant of the propagation reaction.
$m$ weight of monomer (in grams).
$m_0$ initial weight of monomer (in grams).
$n$ number of latex particles.
$n_1$ number of active latex particles.
$n_0$ number of inactive latex particles.
$n^*$ number of latex particles at and beyond $t^*$.
$N$ Avogadro's number.
$R$ grams of catalyst used.
$S$ grams of soap used.
$t$ time elapsed since the start of the reaction.
$t_i$ induction time.
$t^*$ moment of disappearance of soap micelles.
$t_{1/2}$ half-conversion time.
$v_0$ initial volume of monomer (in ml).
$V$ volume of water used (in ml).
$W$ molecular weight of the soap.
$x$ degree of conversion.
$x^*$ degree of conversion at $t^*$.
$\gamma$ weight of monomer solubilized in a single micelle.
$\mu$ growth factor of latex particle.
$\nu$ number of molecules of soap in a single micelle.
$\varphi_m$ specific volume of monomer.
$\varphi_p$ specific volume of polymer.
$\varphi$ the ratio $\varphi_p/\varphi_m$.
$\varrho$ number of radicals produced per unit of time.
$\tau$ reduced time $t/t_{1/2}$.
$\chi$ the ratio $\gamma/c_m$.
$\omega_0$ area occupied by a single molecule of soap.

## REFERENCES


1) W. P. Hohenstein and H. Mark, J. Polym. Sci. **1**, 127, 549, 588 (1946).
2) L. Küchler, Polymerisationskinetik, Springer, Berlin (1951).
3) W. D. Harkins, J. Am. chem. Soc. **69**, 1428 (1947).
4) W. D. Harkins, J. Polym. Sci. **5**, 217 (1950).
5) W. V. Smith and R. H. Ewart, J. chem. Phys. **16**, 592 (1948).
6) R. N. Haward, J. Polym. Sci. **4**, 273 (1949).
7) J. H. Baxendale, M. G. Evans and J. Kilham, Trans. Faraday Soc. **42**, 668 (1946).
8) E. J. Meehan, J. Am. chem. Soc. **71**, 628 (1949).
9) M. L. Corrin, H. B. Klevens and W. D. Harkins, J. chem. Phys. **14**, 480 (1946).
10) K. H. Klaassens, J. Bakker, C. J. Schoot and J. Goorissen, J. Polym. Sci. **7** 457 (1951).
11) C. C. Price and C. E. Adams, J. Am. chem. Soc. **67**, 1674 (1945).
12) F. A. Bovcy and I. M. Kolthoff, J. Polym. Sci. **5**, 487 (1950).
13) M. L. Corrin, J. Polym. Sci. **2**, 257 (1947).
14) K. H. Klaassens a.o., not published.
15) C. F. Fryling, Ind. eng. Chem. An. Ed. **16**, 1 (1944).
16) H. Wiener, J. Polym. Sci **7**, 1 (1951).
17) F. A. Bovey and I. M. Kolthoff, Chem. Revs **42**, 491 (1948).
18) S. R. Shunmukham, V. L. Hallenbeck and R. L. Guile, J. Polym. Sci. **6**, 691 (1951).
19) C. J. Schoot, J. Bakker and K. H. Klaassens, J. Polym. Sci. **7**, 657 (1951).
20) J. Th. G. Overbeek, Thesis, Utrecht (1941); Advances in Colloid Science, III, Interscience, New York (1950).
21) D. C. Henry, Proc. roy. Soc., London, **A133**, 106 (1931).
22) E. R. Blout, W. P. Hohenstein and H. Mark, Monomers, Interscience, New York (1949).
23) T. Alfrey, G. Goldfinger and H. Mark, J. app. Phys. **14**, 700 (1943).
24) I. M. Kolthoff and W. J. Dale, J. Am. chem. Soc. **67**, 1672 (1945).
25) G. M. Burnett, Trans. Faraday Soc. **46**, 772 (1950).
26) I. M. Kolthoff, E. J. Meehan and C. W. Carr, J. Polym. Sci. **6**, 73 (1951).
27) I. M. Kolthoff, E. J. Meehan and C. W. Carr, J. Polym. Sci. **7**, 577 (1951).
28) J. K. Davis and F. E. Bartell, J. phys. Coll. Chem. **47**, 40 (1943).
29) J. W. Mc Bain and K. J. Lissant, J. phys. Coll. Chem. **55**, 655 (1951).
30) W. V. Smith, J. Am. chem. Soc. **70**, 3695 (1948).
31) P. D. Bartlett and K. Nozaki, J. Polym. Sci. **3**, 216 (1948).
32) P. D. Bartlett and J. D. Cotman Jr., J. Am. chem. Soc. **71**, 1419 (1949).
33) R. H. Ewart c.s., loc. cit. W. V. Smith, J. Am. chem. Soc. **68**, 2059 (1946)
34) F. Booth, Proc. roy. Soc., London, **A203**, 514 (1950).

VOL. 7 NR 6 DECEMBER 1952

# Philips Research Reports

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

R 200 *Philips Res. Rep.* 7, 401-420, *1952*

## INTRINSIC EFFICIENCIES OF PHOSPHORS UNDER CATHODE-RAY EXCITATION

by A. BRIL and H. A. KLASENS

535.376:621.385.832

**Summary**

The intrinsic efficiencies of phosphors and their voltage and temperature dependence were measured with cathode-ray excitation in a demountable tube. The light outputs of thin layers of phosphors settled on glass were also determined. The light output at the glass-side goes through a maximum with increasing layer thickness. For sulphide mixtures the maximum light output at the glass side was found to be only 30% of that of a thick layer at the gun side. Further losses of light occur due to settling and current saturation. As a result only 25% of the total light that can be obtained from a phosphor is radiated from the sulphide screen of a cathode-ray tube. The value of the scattering and absorption coefficients were determined for sulphide and silicate screens using the formulas of Hamaker. Measurements on aluminized screens showed that the maximum gain of 2·5 to 3 which can be reached theoretically is not obtained because of light absorption in the aluminium layer.

**Résumé**

Les rendements intrinsèques des phosphors et leur dépendance du voltage et de la température ont été mesurés à l'aide d'une excitation à rayons cathodiques dans un tube démontable. Le rendement lumineux de couches minces de phosphors, déposés sur un verre, a été déterminé aussi. Le rendement lumineux du côté du verre passe par un maximum, à mesure que l'épaisseur des couches augmente. Pour les mélanges à sulphide le rendement maximum se trouve être seulement 30% de celui d'une couche épaisse du côté du canon. D'autres pertes de lumière se produisent à cause du dépôt et de la saturation due au courant élevé. Par conséquent ce n'est qu'un quart de la lumière totale qu'on puisse obtenir d'un phosphor, qui est rayonné de l'écran de sulphide d'un tube à rayons cathodiques. Les valeurs des coefficients de dispersion et d'absorption ont été déterminées pour les écrans de sulphide et de silicate, à l'aide des formules de M. Hamaker. Des mesures faites sur des écrans couverts d'une couche d'aluminium ont montré que le profit maximum de 2,5 à 3, qu'on peut atteindre théoriquement, n'est pas réalisé à cause de l'absorption de la lumière par la couche d'aluminium.

**Zusammenfassung**

Es werden die intrinsiken Ausbeuten der Phosphore und ihre Spannungs- und Temperaturabhängigkeit gemessen mit Hilfe von

Kathodenstrahlerregung in einer demontierbaren Röhre unter Verwendung dicker Schichten. Zugleich wurde auch die Lichtausbeute auf Glas aufgebrachter dünner Schichten bestimmt. Die Lichtausbeute an der Glasseite geht durch ein Maximum. Für Sulfidgemische wurde als Maximum der Lichtausbeute an der Glasseite ein Wert gefunden von nur 30% der Lichtausbeute einer dicken Schichte an der Elektronenbeschußseite (Kathodenseite). Weitere Verluste sind Niederschlagsbildung und Stromsättigung zuzuschreiben. Das Ergebnis ist, daß nur 25% des Gesamtlichtes, das von einem Phosphor erhalten werden kann, von dem Sulfidschirm einer Kathodenstrahlröhre ausgestrahlt wird. Die Werte der Streuungs- und Absorbtionskoeffizienten wurden für Sulfid- und Silikatschirme unter Verwendung der Formeln von Hamaker bestimmt. Messungen an aluminierten Schirmen zeigten, daß die theoretisch erreichbare maximale Verstärkung von 2,5 bis 3 × nicht erhalten werden kann infolge Lichtabsorption in der Aluminiumschicht.

## 1. Introduction

Many measurements on the efficiencies of phosphors with cathode-ray excitation have already been carried out by various authors [1]), but generally their results do not agree. This is partly due to differences in the experimental conditions (voltage, etc.), partly also because most of the measurements were carried out on thin cathode-ray screens, where the layer thickness is an important factor.

It was our purpose to determine the absolute efficiencies of a number of phosphors applied in a fairly thick layer in a demountable tube, as a function of voltage and temperature. Furthermore we have endeavoured to correlate these efficiencies with the light outputs of thin screens of these phosphors such as used in cathode-ray tubes. Here we may mention an investigation of Longini [2]), who went just the other way round: From the measurements of the efficiencies of tubes he calculated the intrinsic efficiencies of the phosphors used in the tubes. The disadvantage of this procedure, as he mentioned himself, is that the work consists for 90% in making corrections.

Many of the earlier measurements were carried out in a voltage range where the efficiency $\eta$ increases with voltage $V$. Usually it was found that $\eta$ increased with $V$ according to

$$\eta = cV^n \quad \text{with} \quad n > 1.$$

(See, e.g., Nottingham [3]), Martin and Headrick [4]), Piore and Morton [5]), Strange and Henderson [6]). )

It is obvious that this can only hold over a limited voltage range since it would lead to efficiencies exceeding 100% at high voltages. An explanation of this law is given by Fano [7]) by assuming a less effective surface layer on the phosphor. At low voltages for which the depth of penetration is small, the efficiency is low. At higher voltages the electrons penetrate into the more perfect interior of the lattice, and the efficiency is higher.

Another point of importance is the fraction of the energy of the primary electron beam that is really absorbed by the phosphor lattice [8]), [9]). Some of the primary electrons are reflected elastically and others inelastically, giving only part of their energy to the phosphor. Very little is known on this subject and we have not made any corrections for this loss of energy or for the loss due to secondary electrons. The efficiencies below are based on the assumption that all the energy of the primary electrons is absorbed by the phosphor.

## 2. Experimental procedure

The measurements were carried out in a demountable cathode-ray tube pumped continually by an oil diffusion pump backed with a rotary oil pump. The tube (fig. 1) is analogous to that described earlier by Bril [10]), only now the upper part *A* and the lower tube *C* are both made of metal and separated by the glass part *B* covered at the inner side with a conducting layer. To be able to put high voltages up to 60 kV on the tube, the high-voltage connections are placed in the chamber *A* which is evacuated during the measurements. The upper part of the tube is of the same construction as that used in the Philips electron microscope [11]). All joints are made with the aid of rubber rings.

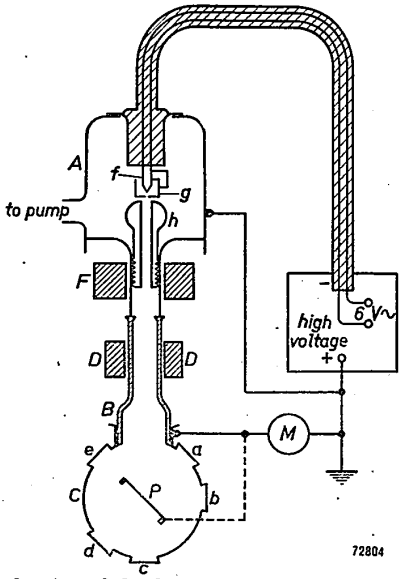



Fig. 1. Schematic drawing of the demountable cathode-ray tube used for the efficiency measurements. The lower part *C* is the cross-section of a long tube perpendicular to the plane of drawing.

The electron gun is kept at a high negative potential. As a cathode an oxide-coated 0·4-mm Ni-Cr filament $f$ is used. It is surrounded by a metal cap $g$ with a small opening through which the electrons pass. The beam current is regulated by varying the current through the filament. The distance between the cap $g$ and the anode $h$ can be varied.

The cross tube $C$ has a diameter of 12 cm and a length of about 50 cm. In the middle of this tube an aluminium carriage, with a piece of soft iron attached to it, can be moved with the aid of a magnet. In this carriage a metal plate, provided with 6 rectangular shallow hollows is placed. The hollows are filled with the various phosphors so that each phosphor can be brought in succession under the electron beam. The beam is focused by the coil $F$ and deflected by the coils $D$ in two directions. The length of the sweeps could be varied. The bombarded area was always made so small that it was falling within the boundaries of each phosphor layer (2·5 × 2·0 cm$^2$). The metal plate $P$ and the cross tube $C$ are connected and both together were earthed via a microammeter $M$. The primary current was determined by measuring the current of the secondary electrons emitted by the phosphor. The secondary electrons are collected partly by the plate $P$, partly by $C$. At first some secondary electrons will also move to the middle part $B$. This part is, however, not connected to earth so that it will be charged up at a negative potential until no more electrons can be collected. The total secondary current is then measured by $M$.

The metal plate $P$ on which the phosphors are deposited prevents them from charging up except perhaps for a few volts. If the phosphor should be charged up much more than this, the bombarded area would alter its size because of the strong field that would then exist between the surface of the screen and the surrounding parts of the tube which have the anode potential. Such an alteration can be observed for instance when the phosphor is covered with a layer of nitrocellulose.

Several quartz windows are mounted in the cross tube $C$: Two are placed opposite each other ($a$ and $d$) at an angle of 45° to the electron beam. If the carriage is placed parallel to these windows, the light output of phosphor layers upon glass plates can be measured on both sides. The other windows are used to observe the phosphor and to measure spectral distributions.

The absolute efficiencies of some phosphors were determined with the aid of a calibrated thermopile as described in section 3. These phosphors were used as standards. The efficiencies of the other phosphors were determined by comparison with these standard phosphors.

For the relative measurements of blue and U.V. phosphors a quartz Cs-Sb vacuum photocell was used, the sensitivity of which was about constant from 2500 to 5000 Å (fig. 2). For green and red phosphors a Rb-

vacuum photocell in combination with a D.C. amplifier was used. The sensitivity of this cell was about constant from 5000 to 7000 Å. Thus with either a blue or a yellow standard phosphor the efficiency of any arbitrary phosphor can in most cases be determined immediately from the ratios of the photo-currents, at least if no infra-red emission is present. This was checked with the aid of a filter only transmitting radiation with wavelengths $> 7500$ Å.

The measurements of the temperature dependence were also carried out in a demountable tube (fig. 3). The phosphors were deposited on a metal plate $P$ attached to the joint $T$. The plate $P$ could be heated by the element $H$. In this way we were able to heat the phosphor up to about 450 °C. The

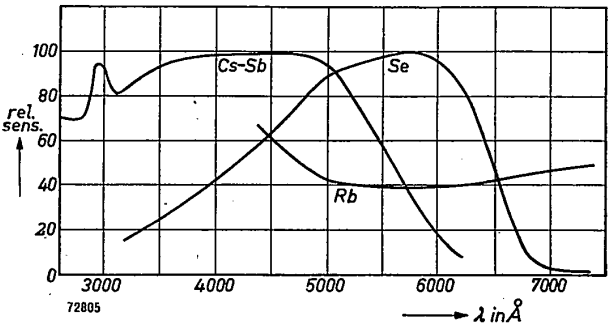



Fig. 2. Relative spectral sensitivity of the photocells used for relative efficiency measurements.

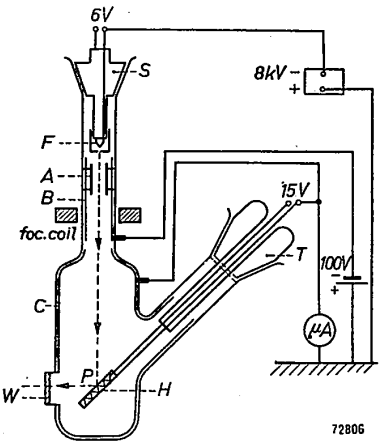



Fig. 3. Schematic drawing of the demountable cathode-ray tube used for the measurements of the temperature dependence.

temperature was measured by a chromel-alumel thermocouple, in contact with $P$.

The accelerating voltage of the tube is variable up to 8 kV. The gun attached to the metal joint $S$ is also demountable. A tungsten filament $F$ is used as cathode. The beam is stationary and focused so that all the primary electrons fall on the phosphor (area of the spot about 2 cm$^2$). The secondary electrons are collected by the plate $P$ and the lower silver layer $C$. Generally a current of about 2 μA is used at a potential of 8 kV.

### 3. Measurement of efficiencies

#### A. *Absolute efficiencies*

The absolute efficiencies of some phosphors were determined with the aid of a calibrated thermopile. The sensitivity of this thermopile was 7·9 μW/cm$^2$ per μV thermovoltage (with an accuracy of about 5%). The thermovoltage was measured by the compensation method, using a sensitive Cambridge galvanometer as indicating instrument.

As was explained in section 2, the phosphors were deposited in a thick layer, so that practically all the light was radiated from the gun side of the phosphor. In order to obtain the intrinsic efficiencies of the phosphors a correction has to be made for the loss of light absorbed in the layer. This can easily be determined by measuring the reflection coefficient $R_\infty$ of an "infinitely thick" layer of the phosphor for its own emission. $R_\infty$ was calculated from measurements of the reflection and emission spectra of the phosphor.

The intrinsic efficiency $\eta_i$ is related to the measured efficiency $\eta_m$ by the formula

$$\eta_i = \frac{2}{1 + R_\infty}\,\eta_m$$

if we assume that the thickness of the layer in which the fluorescence is produced is negligibly small. In that case half of the fluorescence $I$ escapes immediately at the gun side of the phosphor layer while the other half, radiated to the opposite side, is reflected first against the lower layers of the phosphor before it emerges at the gun side. We find thus for the total light output: $\frac{1}{2}I + \frac{1}{2}IR_\infty = \frac{1}{2}(1+R_\infty)I$.

The light output was measured in a direction normal to the phosphor layer through the window $a$, with the electron beam striking the phosphor under an angle of 45°. The voltage was 20 kV, the current density 1 to 2 μA/cm$^2$. To obtain a homogeneously fluorescent area the electron beam was scanned over an area of about 3 cm$^2$. As the cross-section of the electron beam was relatively large with respect to the scanned area,

practically no saturation occurred at the voltage and current density mentioned above. In all cases a distribution of the light according to Lambert's law was assumed.

If the sensitivity of the thermopile is $A$ watts/volt. cm$^2$, $v$ the compensation voltage, $r$ the distance thermopile-phosphor and $t$ the transmission of the window, the total energy of the radiated light per second is

$$I = \pi r^2 v A/t.$$

The power falling on the phosphor is $V_p I_p$ if $V_p$ and $I_p$ are the anode voltage and the beam current respectively. So the measured efficiency is

$$\eta_m = \frac{\pi r^2 v A}{t I_p V_p},$$

and the intrinsic efficiency

$$\eta_i = \frac{2\pi r^2 v A}{t I_p V_p (R_\infty + 1)}.$$

The results are given in Table I.

TABLE I

| phosphor | $I_p$ μA | $v$ μV | $\eta_m$ % | $R_\infty$ % | $\eta_i$ % |
|---|---|---|---|---|---|
| $(CaMg)SiO_3$-0·03 Ti | 7·1 | 1·59 | 7·0 | 91 | 7·4 |
| $Cd_2B_2O_5$-0·002 Mn | 7·1 | 1·24 | 5·3 | 96 | 5·5 |
| ZnS-5.10$^{-4}$Ag-5.10$^{-4}$ Al | 7·1 | 4·84 | 21·0 | 83 | 23·0 |
| | 3·65 | 2·60 | 22·0 | 83 | 24·0 |
| $Zn_2SiO_4$-0·004 Mn | 7·1 | 1·92 | 8·3 | 99 | 8·3 |

To check the measurements, we measured the absolute efficiency of a mixture of $(CaMg)SiO_3$-Ti and $(ZnBe)_2$ $SiO_4$-Mn (used in the projection television tubes) as follows:

An arbitrary selenium photocell was calibrated with the aid of a projection television tube of known light output (measured with the flicker photometer). For this calibration the distance between the screen and the photocell, the dimensions of the fluorescent part of the screen, and the irradiated area of the photocell were made the same as those used in the measurements of the same phosphor mixture in the demountable tube. The identity of the phosphors was verified by measuring the spectral distribution. In this way the projection tube is used as a calibrated light source with the same dimensions and spectral distribution as the exposed phosphor layer in the

demountable tube. For the energy efficiency of the mixture a value of 7·0 % was found, in good agreement with the preceding measurements (see Tables I and II).

B. *Relative efficiencies*

The measured phosphors of the preceding section were now used as standards for relative measurements with the Cs-Sb or Rb vacuum photocells (see section 2 and fig. 2). Table II gives the results for a number of phosphors; the highest efficiencies observed are given. The voltage was 20 kV, the beam current density about 1 $\mu A/cm^2$. The lumen equivalent is the number of lumens that is equivalent to one watt of light energy emitted by the phosphor.

TABLE II

| phosphor | $\eta_m$ % | $R_\infty$ % | $\eta_i$ % | lumen equivalent lm/W | lumen*) output lm/W | colour |
|---|---|---|---|---|---|---|
| ZnS-$10^{-4}$Ag-Cl | 21 | | | 35 | 7 | blue |
| ZnS-$3.10^{-4}$ Ag-$4.10^{-4}$Al | 23 | 83 | 25 | 35 | 8 | blue |
| ZnS-Zn-Cl | 18 | | | 95 | 17 | blue |
| ZnS-Zn-$4.10^{-4}$Al | 22 | | | 125 | 27 | blue |
| (50%Zn-50%Cd)S-$5.10^{-5}$Ag-Cl | 19·5 | | | 500 | 98 | yellow |
| (50%Zn-50%Cd)S-$10^{-4}$Ag-$10^{-4}$Al | 18 | 83 | 19·5 | 500 | 90 | yellow |
| (60%Zn-40%Cd)S-Zn-Cl | 17 | 94 | 17·5 | 400 | 68 | yellow |
| ZnS-$10^{-5}$Cu-$10^{-4}$Al | 23 | 83 | 25 | 400 | 92 | green |
| ZnS-$10^{-4}$Au-$10^{-4}$Al | 15·5 | | | | | green |
| ZnS-0·015 Mn | 4 | 87 | 4·5 | 460 | 18 | yellow |
| ZnO-Zn | 7 | 85 | 7·5 | 340 | 25 | green |
| ZnO | 0·2 | small | | | | u.v. |
| $CaWO_4$ | 3 | 80 | 3·5 | 70 | 2 | blue |
| $MgWO_4$ | 2 | 97 | 2 | 250 | 5 | bluish |
| $Zn_2SiO_4$-0·01 Ti | 8·5 | 94 | 8·5 | 30 | 2·5 | blue |
| $CaO.MgO.SiO_2$-0·03 Ti | 7 | 91 | 7·5 | 70 | 5 | blue |
| $ZrP_2O_7$ | 3·5 | | | | | u.v. |
| $Zn_2SiO_4$-0·004 Mn | 8·5 | 99 | 8·5 | 475 | 40 | green |
| $(0{\cdot}86\ ZnO.0{\cdot}12BeO)_2SiO_2$-0·02Mn | 7 | 97 | 7 | 450 | 31 | yellow |
| $Cd_2B_2O_5$-0·002 Mn | 5·5 | 96 | 5·5 | 240 | 13 | orange |
| $CdSiO_3$-0·01 Mn | 5 | | | 350 | 17 | orange |
| $Cd_5Cl(PO_4)_3$-0·05 Mn | 7 | 95 | 7 | 400 | 28 | orange |
| $\beta Zn_3(PO_4)_2$-0·03 Mn | 8 | 100 | 8 | 170 | 13·5 | red |
| $MgSiO_3$-0·005 Mn | 6 | | | 60 | 3·5 | red |
| $\beta Ca_2P_2O_7$-0·002 Bi | 2 | 100 | 2 | 75 | 1·5 | red |
| $\alpha Ca_2P_2O_7$-0·002 Bi | 2 | | | 35 | 0·7 | red |
| $Sr_3(PO_4)_2$-0·04 Ce | 2·5 | 43 | 3·5 | | | u.v. |
| $NaBaPO_4$-0·002 Ce | 2 | 87 | 2 | | | u.v. |
| $2CaO.Al_2O_3.SiO_2$-0·04 Ce | 4 | 90 | 4 | | | blue |
| NaJ-0·01 Tl | 2 | | | | | u.v. |

*) not corrected for the absorption in the layer

The spectral energy distributions of most of the measured phosphors are well known; we give here only the cathode-luminescence of $ZrP_2O_7$ (unactivated), showing a peak at about 2900 Å (see fig. 4).

The temperature dependence of the efficiency of some phosphors was determined at 8 kV and about 2 $\mu A/cm^2$ (figs 5 and 6). For the blue and ultra-violet emitting phosphors the Cs-Sb vacuum photocell was used, for the other phosphors a selenium barrier-layer cell (fig. 2).

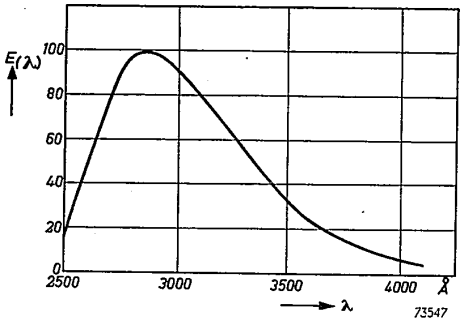



Fig. 4. Spectral energy distribution of the cathode-luminescence of $ZrP_2O_7$ (unactivated).

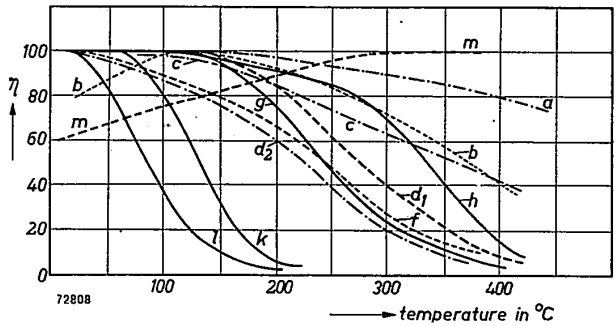



Fig. 5. The relative efficiency $\eta$ as a function of temperature for
(a) $3ZnO.SiO_2.P_2O_5$-0·03Mn
(b) $2CaO.Al_2O_3.SiO_2$-0·04Ce
(c) zincphosphate-Mn
($d_1$) green band of
($d_2$) red band of
} $(ZnBe)_2SiO_4$-Mn
(f) $Cd_5Cl(PO_4)_3$-0·05Mn
(g) $(CaMg)SiO_3$-0·03Ti
(h) $Zn_2SiO_4$-0·004Mn
(k) $Cd_2B_2O_5$-0·002Mn
(l) $MgWO_4$; $CaWO_4$
(m) $Sr_3(PO_4)_2$-0·04Ce

## C. *Voltage dependence of the efficiency*

For a few phosphors the efficiency was measured from 1·5 to 50 kV, namely for $Zn_2SiO_4$-0·004Mn, $(CaMg)SiO_3$-0·03Ti and ZnS-$5.10^{-4}$Ag-$5.10^{-4}$ Al. The results are shown in fig. 7. From low voltages upwards the

efficiencies rise rapidly and reach a more or less constant value at about 20 kV. The small decrease at higher voltages may be due to the increase of self-absorption caused by the increasing penetration depth of the electrons. It is, however, also possible that the number of the elastically scattered electrons is increasing at high voltages [8]).

Our measurements do not confirm the expectations of Coltman [12]) and Longini [2]) that the efficiency for blue sulphide phosphors might be as high as 50% (compare Table II and fig. 7).

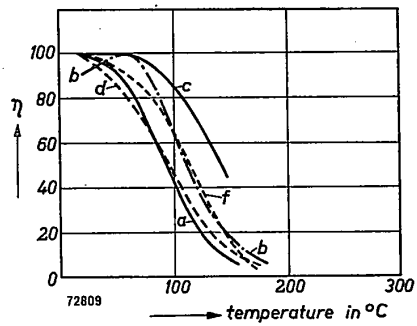



Fig. 6. The relative energy efficiency $\eta$ as a function of temperature for
(a) ZnS-Zn-Cl
(b) ZnS-$10^{-5}$Ag-Cl
(c) ZnS-$10^{-4}$Ag-Cl
(d) (50 Zn-50 Cd)S-Zn-Cl
(f) (50 Zn-50 Cd)S-Ag-Cl

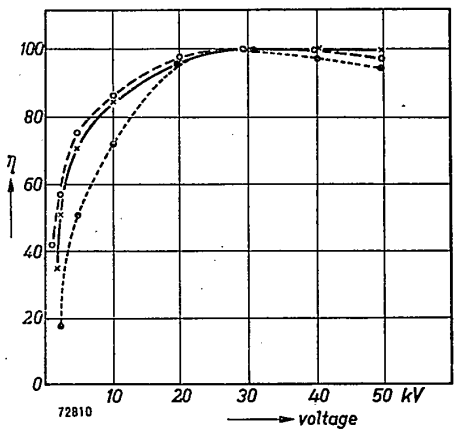



Fig. 7. Relative energy efficiencies up to 50 kV for (CaMg) $SiO_3$-0.03Ti (---o---) $Zn_2SiO_4$-0·004Mn (—×—) and ZnS-$5.10^{-4}$Ag-$5.10^{-4}$Al (---•---).

The efficiencies of the sulphide phoshors at low current densities are nevertheless surprisingly high if one considers the mechanism of the cathode-ray excitation. The primary electrons penetrating into the phosphor lose their energy by collisions with the ions of the lattice. An important part of this energy is dissipated and the rest produces excited states of the base lattice. These states are partly converted into excited states of the luminescence centres and partly return to the ground state by radiationless transitions (the direct excitation of the centres can in many cases be neglected because of the small number of these centres). In the luminescence centres radiationless transitions also may occur.

It is therefore not surprising that the energy efficiency of phosphors under cathode-ray excitation is usually far below that found with ultra-violet excitation where often energy efficiencies of the order of 50% are observed corresponding to quantum efficiences of the order of 100%. For zinc-sulphide phosphors, however, the efficiency under cathode-ray bombardment may be higher than that with ultra-violet excitation (see also Garlick [13]). A few of our measurements are given in Table III.

TABLE III

| phosphor | energy efficiency | | |
|---|---|---|---|
| | cathode rays | 2537 Å | 3650 Å |
| ZnS-5.10$^{-4}$Ag-5.10$^{-4}$Al | 23% | 3% | 21% |
| (ZnCd)S-Zn-Cl | 17% | 2 to 10% | 15 to 35% |
| (ZnCd)S-Ag-Cl | | | |
| ZnS-Cu-Cl | 18% | 10% | 40% |

## 4. Scattering and absorption of light in phosphor layers

The theory of the reflection and absorption of radiation in light-scattering materials was extended for the application to X-ray and cathode-ray screens by Hamaker [14]), Klasens [15]), Coltman e.a. [16]) and Longini [17]).

We use here the formulas of Hamaker, written in the form used by Coltman. We make further the assumption that the fluorescent light is produced in an infinitely thin layer at the gun side of the phosphor layer. This is of course not quite correct, for the penetration depth of a 10-kV electron is about 1 $\mu$. Moreover in a crystal the light undergoes many internal reflections before it escapes from the crystal. The whole crystal must therefore be considered as a source of fluorescence [2]). So the formulas will only be valid for not too thin phosphor layers.

A. *Non-aluminized screens applied on glass plates*

If $I_0$ is the intensity of the fluorescence produced in the layer a part $T$ will radiate at the glass side and a part $R$ at the gun side. $T$ and $R$ are derived from the expressions (24) and (25) in Hamaker's paper [14])

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

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

in which $D$ is the thickness of the layer, $a$ the absorption coefficient, $s$ the scattering coefficient, $\sigma = \sqrt{a(a + 2s)}$ and $r_G$ the reflection coefficient of the glass plate. As a rule $I_0$ is not exactly known; therefore we eliminate it:

$$t = \frac{R}{T} = \frac{1 + r_G}{1 - r_G} \cosh \sigma D + \frac{a + 2s}{\sigma} \sinh \sigma D. \tag{3}$$

Further the reflectivity $R_\infty$ of an "infinitely thick" phosphor layer is known. This gives us the ratio of the absorption and scattering coefficients:

$$\frac{s}{a} = \frac{2R_\infty}{(1 - R_\infty)^2}, \tag{4}$$

while (3) can be written in the form

$$t = \frac{R}{T} = \frac{1 + r_G}{1 - r_G} \cosh \sigma D + \frac{1 + R_\infty}{1 - R_\infty} \sinh \sigma D. \tag{5}$$

If $\sigma D$ is small (e.g., $< 0{\cdot}3$) (5) can be approximated by

$$t = \frac{1 + r_G}{1 - r_G} + (a + 2s) D. \tag{6}$$

If $t$, $R_\infty$, $r_G$ and $D$ are known then from (5) and (4) or (6) and (4) $a$, $s$ and $\sigma$ can be calculated.

Thus by measurement of the light output at both sides of the phosphor, the absorption and reflection coefficients can be determined.

B. *Aluminized screens*

If a phosphor layer is applied on a glass plate with reflection coefficient $r_G$ and covered with an aluminium layer with reflection coefficient $r_A$ we find from equation (25) of Hamaker's paper [14]):

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

$I_0'$ is used instead of $I_0$ to indicate that part of the energy of the electron beam is absorbed in the aluminium layer.

From experiments with non-aluminized screens $a$, $s$ and $\sigma$ can be determined (see section 4A). From the light outputs of a series of metal-backed screens with increasing thickness of the phosphor layer, $r_A$ and $I_0'$ may then be estimated. If the reflection of the glass can be neglected (this is generally allowed) and if $\sigma D$ is small enough to replace cosh $\sigma D$ by 1 and sinh $\sigma D$ by $\sigma D$, (7) is approximated by

$$I = \tfrac{1}{2}I_0' \frac{1+r_A}{\{a + (1-r_A)\,s\}\,D + 1}. \quad (8)$$

If there were no absorption in the powder ($a = 0$) and if the reflection coefficient of the aluminium layer were $= 1$, $I$ would be equal to $I_0'$ for all thicknesses.

### 5. Measurements on thin non-aluminized layers

To correlate the efficiencies measured on thick layers with the light output of thin screens in ordinary non-aluminized cathode-ray tubes and to determine the absorption and scattering coefficients of the phosphors, layers of various thicknesses of some phosphors were settled on glass plates ($2{\cdot}5 \times 3{\cdot}0$ cm$^2$). The metal plate $P$ (fig. 1) in the carriage was replaced by 6 of these glass plates side by side and the light output was measured at both sides of each screen through the windows $a$ and $d$. The voltage was 10 kV, the current density about 1 μA/cm$^2$. The measurements were carried out with the following phosphors: a normal white fluorescing sulphide mixture, willemite, $(CaMg)SiO_3$-Ti, $(ZnBe)_2SiO_4$-Mn and a white fluorescing mixture of the last two phosphors (used in projection tubes). For the first two phosphors the results are plotted in figs 8 and 9; for the other phosphors similar curves were obtained.

The light output at the glass side of the phosphor goes through a maximum and then falls to zero for infinite layer thickness. The maximum is caused by gaps between the grains, so that a proportion of the electrons can pass through the gaps and hit the glass when the layers are thin. This is also shown by the increase of the total light output. If there were no gaps the light output on the glass side would be given by formula (1) of section 4: For $D \to 0$, $T \to \tfrac{1}{2}I_0\,(1-r_G)$. With increasing $D$, $T$ steadily decreases.

The reflection of the glass plate is about 8%, so that the light output at the glass side can never exceed 46% of the total light output. Since the

layer has to be made so thick that the glass is covered completely by the phosphor, the light output is lower.

The light output at the gun side rises at first rapidly and then gradually approaches the value for infinite layer thickness. For the sulphide mixture the maximum light output at the glass side was found to be 30%, 60% being radiated to the gun side and 10% being lost because of electrons not being absorbed or only partly so.

For willemite these percentages were respectively 38%, 52% and 10%, for $(CaMg)SiO_3$-Ti 37%, 52% and 11% and for $(ZnBe)_2SiO_4$-Mn 39%, 51% and 10%.

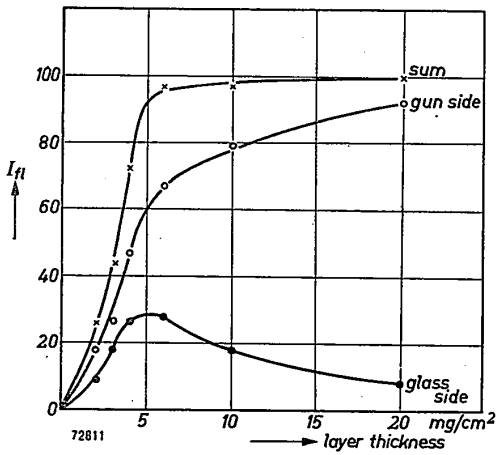



Fig. 8. Light output at the glass and gun sides and the sum of both for white fluorescent sulphide layers at 10 kV (arbitrary units).

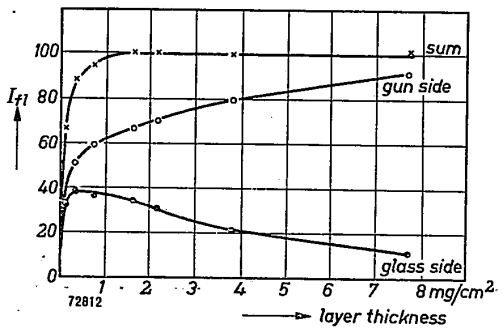



Fig. 9. Light output of willemite layers at 10 kV (arbitrary units).

Actually the light output of a cathode-ray screen is still lower than would follow from these figures because there are still further losses:

(i) The fluorescent powder removed from a normally settled screen shows an efficiency that is about 10% lower than that of the original powder. Perhaps the grains are covered with a thin layer of $SiO_2$ which absorbs part of the electron energy.

(ii) In a 12-inch cathode-ray tube with a sulphide mixture some current saturation usually occurs. Under normal conditions the efficiency loss due to current saturation was found to be approximately 10 to 15%. This was determined by measuring the differences in brightnesses between focused and defocused rasters.

The light output of a non-aluminized sulphide screen with an optimum thickness for 10 kV can thus be calculated as follows:

$$\text{Powder } 50 \text{ lm/W} \xrightarrow[\text{losses}]{\text{settling}} 45 \text{ lm/W} \xrightarrow[\text{saturation}]{\text{current}} 39 \text{ lm/W}$$

$$\xrightarrow[\text{gun side}]{\text{loss to}} 12 \text{ lm/W } (\approx 4 \text{ cd/W}).$$

We return to the measurements on the willemite phosphor, shown in fig. 9. The reflection coefficient of a thick layer of the willemite was 95%, or

$$\frac{s}{a} = \frac{2R_\infty}{(1-R_\infty)^2} = 800. \tag{9}$$

Since the accuracy of the reflection measurement is about 2%, $s/a$ lies between 200 and 2000.

From (6) we find $a + 2s$ and since the ratio $s/a$ is so high that $a$ can be neglected we may write

$$s = \frac{1}{2D}\{t - (1 + r_G)/(1 - r_G)\}.$$

Substituting for $t$ $(=R/T)$ the values of fig. 9 for a few thicknesses we find for $s$ the values in Table IV. The layer thickness has roughly been approximated by the amount of powder in g/cm² divided by the density (3·45 g/cm³). From these results $s$ seems to increase with increasing layer thickness. This is no doubt due to the fact that our theory of scattering is not valid for such thin layers as were used here. Substituting these values of $s$ in (9) we find an absorption coefficient of the order of 1 to 2 $cm^{-1}$. The absorption in willemite for its own emission is apparently very small. For the sulphide mixture (fig. 8) $\sigma D$ is not small enough to apply expression (6). Therefore $\sigma$ was first calculated from (5). Then $s$ was obtained from

IPR2025-00217 EXHIBIT 2010 Page 420 of 485

$$s = \frac{2R_\infty}{1 - R_\infty{}^2}\,\sigma,$$

and $a$ from (4). For this sulphide mixture $R_\infty = 0{\cdot}83$. The results are given in Table V.

The absorption in the sulphide mixture is higher and the scattering smaller than in willemite. The latter is due to the larger grain size of the sulphide phosphors.

For $(CaMg)SiO_3$-Ti, $(ZnBe)_2SiO_4$-Mn and a white fluorescing mixture of these two phosphors the coefficients are given in Table VI.

TABLE IV

| willemite | | | |
|---|---|---|---|
| $D$ $10^{-3}$ g/cm$^2$ | $\sigma$ 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 |

TABLE V

| sulphide mixture | | | |
|---|---|---|---|
| $D$ $10^{-3}$ g/cm$^2$ | $\sigma$ cm$^{-1}$ | $s$ cm$^{-1}$ | $a$ cm$^{-1}$ |
| 6 | 80 | 400 | 7 |
| 9 | 90 | 500 | 8 |
| 12 | 130 | 700 | 12 |
| 20 | 150 | 800 | 14 |

TABLE VI

| | $(CaMg)SiO_3$-Ti | | | $(ZnBe)_2SiO_4$-Mn | | | mixture | | |
|---|---|---|---|---|---|---|---|---|---|
| $D$ $10^{-3}$ g/cm$^2$ | $\sigma$ cm$^{-1}$ | $s$ cm$^{-1}$ | $a$ cm$^{-1}$ | $\sigma$ cm$^{-1}$ | $s$ cm$^{-1}$ | $a$ cm$^{-1}$ | $\sigma$ cm$^{-1}$ | $s$ cm$^{-1}$ | $a$ cm$^{-1}$ |
| 2 | 60 | 600 | 3 | 50 | 600 | 2 | | | |
| 4 | 100 | 1000 | 5 | 75 | 900 | 3 | 60 | 600 | 3 |
| 7 | 120 | 1300 | 6 | 100 | 1200 | 4 | 80 | 800 | 4 |
| 12 | | | | | | | 100 | 1000 | 5 |
| 20 | | | | | | | 100 | 1000 | 5 |

## 6. Measurements on thin aluminized layers

### A. *Absorption of cathode-ray energy in the aluminium layers*

In the preceding section we saw that the light output at the glass side of the screen was 30 to 40% of the total light output, while the light absorption in the layer is negligible. If the phosphor could be backed by an aluminium film with 100% reflection and if no electron energy was lost in the aluminium layer, the light output would be increased by a factor 2·5 to 3.

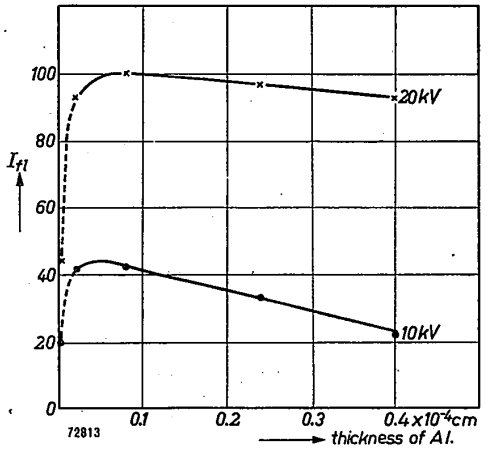



Fig. 10. Light output in arbitrary units of a thin willemite screen as a function of thickness of aluminium backing at 10 and 20 kV.

We determined the loss of energy of the electron beam by measuring the light output of a thin willemite screen as a function of the aluminium thickness. A number of glass plates were covered with a willemite layer (optimal thickness for non-metal-backed layers) and provided in the normal way with various thicknesses of aluminium, namely 0, 0·02, 0·08, 0·24 and 0·4 μ. The light outputs were measured at 10 and 20 kV (fig. 10).

We see that for 20 kV there is only a slight drop in efficiency up to 0·4μ, so that we may conclude that practically no energy is lost in the aluminium layer. At 10 kV the drop in light output for a thickness of 0·4 μ is already considerable. For 0·04 μ, however, the ratio of the efficiencies at 10 and 20 kV is equal to the ratio for the original powder (see fig. 7). Because at 20 kV the energy loss in the aluminium it negligible, it must also be negligible for 10 kV at 0·04 μ. This is in accordance with the calculations of Longini [2]) and others (Longini gives a loss of 3% for 10 kV at an aluminium thickness of 500 Å = 0·05 μ).

To obtain the maximum light output at 10 kV, the thickness of the aluminium backing must lie between about 0·02 to 0·08 $\mu$, while for 20 kV the optimum thickness is about 0·1 $\mu$.

B. *Measurement of the reflection coefficient of the aluminium backing*

A series of thin willemite layers of increasing thickness, settled on glass plates simultaneously with the plates for which the measurements are given in fig. 9, were aluminized. The thickness of the aluminium was 0·3 $\mu$.

The light output was measured at 10, 20, 30, 40 and 50 kV. The results are given in fig. 11; along the vertical axis the relative efficiencies (and not the light outputs) are given. With increasing voltage the maximum shifts to larger thicknesses of the phosphor layer, owing to the increasing penetration depth of the electrons. For the curve of 10 kV the measured points show some spreading, probably due to variation in the thickness of the aluminium backing. The reflection coefficient of the aluminium layer $r_A$ was calculated from the curve for 20 kV (for the higher voltages the equations of section 4 are less accurate as they are based on the assumption of a small penetration depth. We see from the figure that for $D = 7{\cdot}5$ mg/cm$^2$ the efficiency is decreased to two thirds of the maximum efficiency at $D = 1{\cdot}5$ mg/cm$^2$.

Using expression (8) for the light output we can write

$$\frac{I_{1\cdot5}}{I_{7\cdot5}} = \frac{1 + \{a + (1 - r_A)s\}\, D_{7\cdot5}}{1 + \{a + (1 - r_A)s\}\, D_{1\cdot5}} = 1{\cdot}50.$$

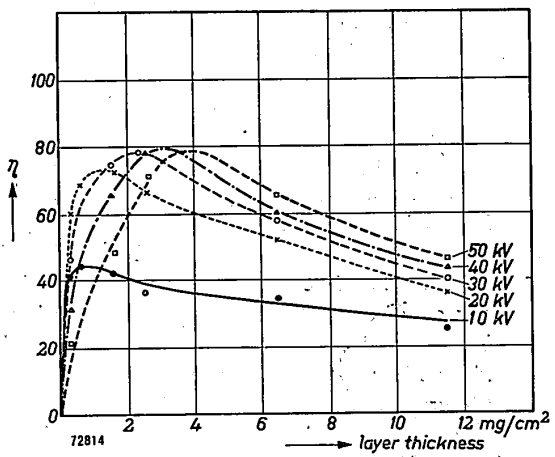



Fig. 11. Relative energy efficiency of thin layers of aluminized willemite (thickness Al 0·3 $\mu$) as a function of the thickness of the willemite layer. Relative efficiency of the original powder in a thick layer at 50,40 and 30 kV is 100, at 20 kV 96 and at 10 kV 85 (see fig. 7).

After inserting the values for $D$ and the corresponding values of $a$ and $s$ from Table IV we find $r_A = 0{\cdot}83$. This is in good agreement with a direct measurement of the reflection of an aluminium film removed from a willemite layer which gave $r_A = 0{\cdot}82$. (The accuracy of the above calculation is, however, smaller than is suggested by the agreement between these two values.)

The same experiment was carried out for a series of glass plates covered with a mixture of $(CaMg)SiO_3$-Ti and $(ZnBe)_2SiO_4$-Mn. From this series we found the reflection coefficient of the aluminium $r_A = 0{\cdot}73$. We may conclude that the reflection coefficient of an aluminium backing is of the order of 80%.

We have already pointed out in section 4 that the efficiency of an aluminized phosphor layer would remain constant with increasing layer thickness, if

(1) the reflection coefficient of the aluminium backing is 100%;
(2) there is no energy loss in the aluminium;
(3) the absorption coefficient $a = 0$.

We have further shown that the absorption of the light in the phosphor is negligible and that, if the thickness of the aluminium is suitably chosen, the energy loss of electrons in the aluminium can also be neglected (voltage not under 10 kV). The decrease in light output at larger screen thicknesses is therefore mainly due to the absorption of the light in the aluminium layer, where the reflection coefficient is only 80%. This agrees with the opinion of Coltman [18]) and is contrary to the opinion of Nicoll [19]) who ascribed the decrease in efficiency to the absorption of the light in the powder. Although Nicoll's experiments were carried out with sulphides whose absorption coefficient is somewhat higher $a$ is still small compared with $(1 - r_A)\, s$ in expression (8).

### C. *Gain of light output with metal backing*

The maximum gain of a factor 2·5 to 3 which can be obtained theoretically by an ideal reflecting layer cannot be reached, mainly because the reflection coefficient of the aluminium layer is only 80%. From the formulas (1) and (8) for the maximum light radiated from the glass side of non-aluminized and aluminized screens we find a gain factor of 2·2 for willemite. We see from fig. 10 that the ratio of the maximum light output to that of the non-aluminized screen (thickness of Al is zero) is experimentally indeed quite near this value.

**Acknowledgement**

We wish to express our thanks to Dr F. A. Kröger, Mr P. Zalm, and Mr J. Broos for making the preparations, to Dr J. Voogd for the absolute

calibration of the thermopile, to Dr Th. P. J. Botden for measuring the spectral distributions and reflection coefficients and to Mr J. C. Gravesteyn for preparing the thin phosphor layers.

*Eindhoven, March 1952*

## REFERENCES

1) P. Lenard, Ann. Phys. **12**, 449, 1903.
W. Kordatzki, A. Schleede and F. Schröter, Physik. Z. **27**, 392, 1926.
H. W. Ernst, Ann. Phys. **82**, 1051, 1927.
W. Rogowski and E. Rühlemann, Arch. Elektrotech. **24**, 691, 1930.
M. Knoll, Z. tech. Phys. **12**, 54, 1931.
A. Güntherschulze and F. Keller, Z. Phys. **77**, 528, 1932.
W. Schnabel, Arch. Elektrotech. **28**, 789. 1934.
M. von Ardenne, Z. tech. Phys. **16**, 61, 1935.
H. W. Leverrenz, J. opt. Soc. Am. **27**, 25, 1937.
T. B. Brown, J. opt. Soc. Am. **27**, 186, 1937.
A. Schleede and B, Bartels, Telefunken Hausmitt. **20**, 104, 1939.
K. Scherer and R. Rübsaat, Arch. Elektrotech. **31**, 821, 1937.
R. T. Orth, P. A. Richards and L. B. Headrick, Proc. Inst. Radio Engrs **23**, 1308, 1935.
L. Levy and D. W. West, J. Inst. El. Eng. London **79**, 11, 1936.
A. W. Moskwin, Bull. Acad. Sci. USSR **9**, 429, 1945.
H. W. Leverenz, An introduction to luminescence of solids, Wiley, New-York—London 1950, pp. 350, 430.
2) R. L. Longini, J. opt. Soc. Am. **39**, 877, 1949.
3) W. B. Nottingham, J. appl. Phys. **10**, 73, 1939.
4) S. T. Martin and L. B. Headrick, J. appl. Phys. **10**, 116, 1939.
5) E. R. Piore and G. A. Morton, J. appl. Phys. **11**, 153, 1940.
6) J. W. Strange and S. T. Henderson, Proc. phys. Soc. London **58**, 383, 1946.
7) U. Fano, Phys. Rev. **58**, 544, 1940.
8) See K. G. McKay, Advances in Electronics I, 65, 1948.
9) See H. Bruining, Die Sekundär-Elektronen-Emission fester Körper, Berlin, 1942.
10) A. Bril, Physica **15**, 361, 1949.
11) A. C. van Dorsten, H. Nieuwdorp and A. Verhoeff, Philips tech. Rev. **12**, 33, 1950.
12) J. W. Coltman, Solid Luminescent Materials, Cornell Symposium, Wiley, N.Y.—London, 1948, p. 417.
13) G. F. J. Garlick, Advances in Electronics II, 164, 1950.
14) H. C. Hamaker, Philips Res. Rep. **2**, 55, 1947.
15) H. A. Klasens, Philips Res. Rep. **2**, 68, 1947.
16) J. W. Coltman, E. G. Ebbighausen and W. Altar, J. appl. Phys. **18**, 530, 1947.
17) R. L. Longini, J. opt. Soc. Am. **39**, 551. 1949.
18) J. W. Coltman, Ref. 12, p. 427.
19) F. H. Nicoll, Ref. 12, p. 418.

Philips Res. Rep. 7, 421-431, 1952

# NEW PHOSPHORS FOR FLYING-SPOT CATHODE-RAY TUBES

by A. BRIL and H. A. KLASENS

535.373.3:621.385.832

**Summary**

The decay rates of a great number of phosphors under cathode-ray excitation have been determined. Ce-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 blackening with 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 $ZrP_2O_7$.

**Résumé**

La constante du temps de persistance d'un grand nombre de phosphors, sous l'excitation à rayons cathodiques, ont été déterminés. Les phosphors activés par Ce valent le mieux pour les tubes à rayons cathodiques „flying spot" utilisés pour la télévision en noir et blanc. De ces phosphors c'est le $2CaO.Al_2O_3.SiO_2$-Ce qui a été choisi étant le meilleur phosphor en combinaison avec des verres que ne noircissent pas les rayons-X douces ni les rayons cathodiques. Des phosphors activés par Bi conviennent comme composant rouge dans des tubes à rayons cathodiques pour la télévision en couleurs. $ZrP_2O_7$ est un phosphor convenable pour U.V. microscopy avec des tubes à rayons cathodiques.

**Zusammenfassung**

Es wurden die bei Kathodenstrahlerregung sich ergebenden Abklingzeiten einer großen Anzahl Phosphore bestimmt. Mit Ce aktivierte Phosphore eignen sich vorzüglich für Braunsche Hochvolt-Abtaströhre in der Schwarz-weiß-Fernsehtechnik. Unter diesen Phosphoren wurde $2\ CaO.Al_2O_3.SiO_2$-Ce als der beste Phosphor ausgewählt in Kombination mit Gläsern, die unter Beschuß mit weichen Röntgenstrahlen oder Kathodenstrahlen sich nicht schwärzen. Mit Bi aktivierte Phosphore eignen sich als rote Komponente in Kathodenstrahlröhren mit wanderndem Lichtfleck beim Farbenfernsehen. $ZrP_2O_7$ ist ein geeigneter Phosphor für U.V. Mikroskopie mittels Abtaströhren.

## 1. Introduction

The flying-spot cathode-ray tube has become an important pickup device, used for television. The screen of the cathode-ray tube, scanned by a beam of constant intensity, is imaged on the object to be transmitted. The transmitted or reflected light falls on the cathode of a phototube whose photocurrent constitutes the signal. The number of picture elements which have to be transmitted to produce a well-defined picture of 625 lines and 25 complete images per second is of the order of $10^7$ per second. The

decay time of the phosphor used in the flying-spot tube should therefore preferably be of the order of $10^{-7}$ second. Too slow decays produce a blurred image. The effect of the afterglow of the phosphor can however be corrected electrically with suitable RC circuits [1] [2] [3]. The highest frequency which can be transmitted is determined by the relation [3]

$$\omega^2 \leqslant \frac{\sigma I_0}{2r^2 e\tau}, \quad \text{if} \quad \omega\tau \gg 1, \tag{1}$$

where $r$ = signal-to-noise ratio, $\omega$ = frequency, $\sigma$ = sensitivity of the photocell, $\tau$ = decay time of the phosphor, and $I_0$ = fluorescence intensity of a screen element at the beginning of the decay.

In fig. 1 the fluorescence intensity of a screen element is shown schematically as a function of time. In the time $\Delta t$ during which the element is excited by the scanning electron beam the intensity rises to the value $I_0$ and then falls to $I_0/e$ in the time $\tau$ if the decay is exponential.

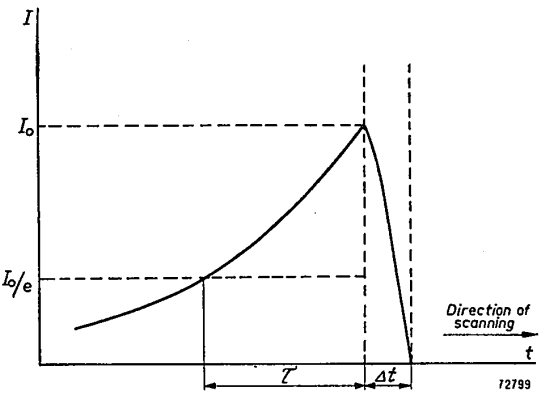



Fig. 1. Schematic representation of the fluorescence intensity of a screen element under scanning conditions.

$I_0$ is a function of $n_0$, the number of excited centres at the end of the excitation, according to the relation

$$I_0 = -\frac{\mathrm{d}n}{\mathrm{d}t} = \frac{n_0}{\tau}. \tag{2}$$

Bril [4] has shown that the efficiency of a phosphor under scanning conditions is proportional to $n_0$ if $\tau > \Delta t$, and that is the case for most phosphors since $\Delta t \approx 10^{-7}$ sec. The quality of the phosphor used in flying-spot cathode-ray tubes is thus determined by the quality factor

$$Q = \sigma\eta/\tau^2. \tag{3}$$

To make $\sigma$ as high as possible the spectral distribution of the light emitted by the phosphor should be adapted to the spectral sensitivity of the photocell. The decay time $\tau$ is determined primarily by the nature of the electronic processes in the phosphor responsible for the emission of light.

Three types of phosphors may be distinguished.

(1) The rate of decay is determined by the rate of recombination between electrons and ionized centres; the decay is not exponential and depends on the density of excitation: ZnO, ZnS.

(2) The electronic process is confined to a group of atoms or ions: tungstates, molybdates, vanadates, zirconates, titanates, uranyl compounds and possibly blue-fluorescent willemite. The decay is exponential.

(3) The electronic process is confined to one ion: phosphors activated with $Mn^{2+}$, $Mn^{4+}$, $Cr^{3+}$, $Sn^{2+}$, $Sb^{3+}$, $Pb^{2+}$, $Tl^{+}$, $In^{2+}$, $Bi^{3+}$ and ions of the rare earth elements. The decay is exponential.

The decay rate of the phosphors belonging to group (3) is influenced to some extent by the nature of the lattice especially when electrons from the outer shells of the activator ions are excited. It is well known that the decay of $ZnF_2$-Mn is a factor 10 slower than the decay of willemite.

The temperature may also be a decisive factor. In many phosphors the decay time decreases approximately proportionally to the efficiency in the quenching range [5-10]). Heating the screen will then have a beneficial effect since according to eq. (3) $Q$ is proportional to $\eta/\tau^2$. Such an effect has indeed been observed by Karolus [11]) who heated the ZnO screen of his flying-spot cathode-ray tube to 150 °C and found that the efficiency was reduced by 50% and the signal-to-noise ratio increased by a factor 2.

It was the purpose of this investigation to compare phosphors of the three groups in order to find which are the most suitable for the flying-spot cathode-ray tube. Most of the measurements were carried out on phosphors prepared by F. A. Kröger or P. Zalm.

## 2. Apparatus

The decay times were measured with the phosphor in a continually pumped demountable tube (fig. 2). The plate $P$ with the phosphor and the electron gun $G$ could both easily be exchanged by means of ground-glass joints. A 0·2-mm oxide-coated Ni-Cr filament was used as a cathode; a tungsten filament would radiate too much light on the photomultiplier. The phosphor is bombarded intermittently by the electrons in a way similar to that used by Nelson, Johnson and Nottingham [12]). A square wave of voltage from the oscillator (*Sq.W.*), the frequency of which is variable, is fed to the grid $g$ of the electron gun of the tube, so that the beam is suppressed periodically.

The fluorescent light falls through a quartz window onto an R.C.A.

photo-multiplying tube, the output of which is fed to the vertical amplifier of the oscillograph (Philips GM 5653); this is a wide-band amplifier which can be used up to 7 Mc/s. The capacity of the connection oscillograph-photomultiplier was about 20 pF. To keep the RC-time small, the input resistance of the oscillograph was reduced by putting a resistance parallel to it. For the fastest decays we used a resistance of 2 kΩ. In this way decay times up to 0·4 μsec could be determined.

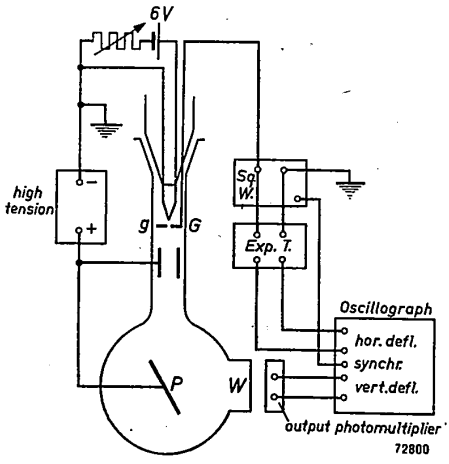



Fig. 2. Circuit diagram.

The linear time base of the oscillograph was replaced by a variable exponential time base (*Exp. T*) according to Tolstoi and Feofilov [13]) [10]). The square-wave pulse on the grid of the cathode-ray tube is also added to the grid of a triode. A condenser is charged via the triode during the period that the phosphor is excited. During the period that the electron beam is suppressed the condenser is discharged across a variable resistance. This voltage supplies the exponential horizontal time base of the oscillograph. If the decay of the phosphor is also exponential, a straight line on the screen may be obtained by changing the values of the capacity and the resistance. This indicates that the decay of the fluorescence follows the same law as the time base. By choosing suitable values for the capacities and resistances the decay time of the phosphor can immediately be read on the time-base apparatus. If the decay is not exponential a mean value is given.

For long decay times, e.g. of phosphors activated with manganese, the

square-wave oscillator was replaced by a mechanical interruptor: The circumference of a disk attached to the shaft of a motor (25 or 1 rev. per sec) was partly provided with a metallic strip and partly insulated. By means of a block brush pressing against the disk a voltage of 100 V was applied periodically to the grid *g*.

The bigh voltage on the tube was in most cases 6 kV, the beam current density about 1 to 10 $\mu A/cm^2$.

## 3. Results

*Group-1 phosphors:* Phosphors of this group, and in particular ZnO, have been used most for flying-spot cathode-ray tubes. Pure ZnO shows an ultraviolet emission with a maximum at 3900 Å [14] [15] [16]. ZnO containing an excess of Zn has a green emission with a peak at 5050 Å. The decay of the U.V. emission is very fast and is of the order of $10^{-7}$ sec [14] [17]. The emitted light is absorbed very strongly by the substance itself. Its efficiency is therefore only about 0·2% [18]. The efficiency of the green-fluorescent ZnO is much higher. The best samples showed an efficiency of 7% [18], but the decay time is ten times longer than that of the U.V.-emitting ZnO. Despite its much lower efficiency U.V.-fluorescent ZnO is preferable to green ZnO for black-and-white television because its quality factor is higher.

For the transmission of coloured pictures, however, a phosphor or mixture of phosphors is required whose emission spectrum covers a wide range from about 4000 to 7000 Å. For this purpose green-fluorescent ZnO is used mostly although it emits very little radiation in the red region (fig. 3).

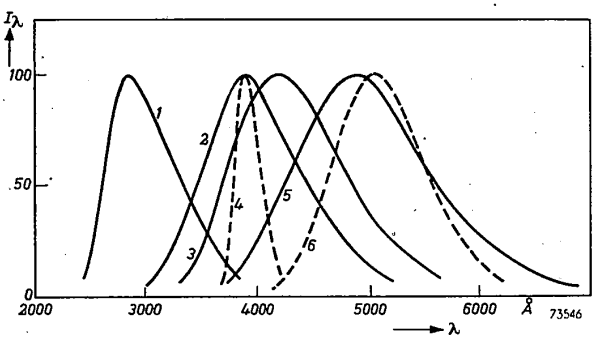



Fig. 3. Spectral distribution curves of various members of group-1 and group-2 phosphors
(1) $ZrP_2O_7$ (2) $Zn_2SiO_4$-0·01 $TiO_2$
(3) CaO . MgO . $2SiO_2$-0·03 $TiO_2$ (4) ZnO(U.V.band)
(5) $MgWO_4$ (6) ZnO (green band),

The decay rates of several sulphide phosphors were determined at 6 kV and 1 $\mu A/cm^2$. Their decay time was of the order of $10^{-4}$ to $10^{-5}$ sec. Although their efficiencies are higher than those of the two ZnO phosphors their quality factor is lower for the decay is much too slow. The selenides are usually somewhat faster than the sulphides.

*Group-2 phosphors.* The decay times of several phosphors are given in Table I together with data from literature. The spectral distribution curves of many phosphors of this group are shown in fig. 3. All the measured phosphors of group 2 have decay times longer than $2.10^{-6}$ sec. The titanium-activated silicates have the highest efficiencies [18]). They are nevertheless inferior to ZnO for black-and-white television because their decay rate is too slow to give a satisfactory quality factor $Q$. Also for colour television they do not offer any advantages above green-fluorescent ZnO.

An interesting member of this group of phosphors, however, is $ZrP_2O_7$ because it has an emission in the short U. V. region with a maximum at 2850 Å combined with a reasonably high efficiency of 3·5% [18]) and a fast decay. It is therefore very suitable for U.V. microscopy with a flying-spot cathode-ray tube. Since the resolution of a microscope is proportional to $1/\lambda$, where $\lambda$ is the wavelength of the light used for illumination, a reduction of $\lambda$ will result in an increased resolving power. In normal microscopy no radiation shorter than 4000 Å can be used because of the sensitivity of the eye. With the flying-spot technique, however, ultraviolet radiation may be used for illumination [30]). The region of 2500 to 3500 Å, emitted by $ZrP_2O_7$, is moreover very important biologically for some constituents of the living cell (such as nucleic acids) have an absorption only in this region. Most of the light emitted by this new phosphor is absorbed by the glass face of the ordinary type of projection cathode-ray tubes. A special tube with a quartz face must therefore be used.

*Group-3 phosphors.* The results are collected in Table II which also includes data from literature.

The decays of the first three activators of Table II, $Mn^{2+}$, $Mn^{4+}$ and $Cr^{3+}$, of which the last two have the same electronic configuration, are slow. The electronic transitions are not accurately known, but forbidden transitions are certainly involved.

The next five ions, $Sn^{2+}$, $Sb^{3+}$, $Tl^{+}$, $Pb^{2+}$ and $Bi^{3+}$, all have the same electronic configuration, viz. that of the mercury atom. The emission corresponds to the 2537-Å Hg line ($^3P_1 \rightarrow {}^1S_0$) and is accordingly fast. The emission of $\beta Ca_2P_2O_7$-Bi is red with a broad band extending from 6000 to 7300 Å and a maximum at 6550 Å [31]); its energy efficiency under cathode-ray bombardment is 2%, about one-third of that of green ZnO. This phosphor may be of importance as an addition to green ZnO for the reproduction of red colours in colour television. Phosphates activated

TABLE I

decay times of phosphors belonging to group 2

| phosphor | decay time (sec) |
|---|---|
| $CaWO_4$ | 2 to $5.10^{-6}$ |
| | $16.10^{-6}$ Schnabel [19]) |
| | $1.10^{-5}$ Briggs [20]) |
| | $\approx 10^{-5}$ Strange and Henderson [21]) |
| | $6.10^{-6}$ Marshall [22]) (X-ray excitation) |
| $CdWO_4$ | $3.10^{-6}$ |
| | $12.10^{-6}$ Briggs [20]) |
| | $\approx 10^{-5}$ Strange and Henderson [21]) |
| $MgWO_4$ | $4.10^{-5}$ |
| | $> 6.10^{-6}$ Marshall [22]) |
| $Zn_2SiO_4$ (blue fluorescent) | $5.10^{-6}$ |
| | $10^{-4}$ to $10^{-5}$ Strange and Henderson [21]) |
| | $\approx 5.10^{-4}$ Vavilov and Levshin [23]) |
| $UO_2$ compounds | 3 to $8.10^{-4}$ Delorme and Perrin [24]) |
| | $10^{-3}$ Kröger a.o. [7]) |
| | $4.10^{-4}$ Tolstoi and Feofilov [13]) |
| $CaO.MgO.2SiO_2$-0·03 $TiO_2$ | $25.10^{-6}$ |
| 2 $ZnO.SiO_2$-0·01 $TiO_2$ | $4.10^{-6}$ |
| $Zr\ P_2O_7$ | $2.10^{-6}$ |

with $Sb^{3+}$ are not excited well with cathode rays. Zalm [35]) tried to determine the decay time of some apatites activated by $Sb^{3+}$ under U.V. ex-

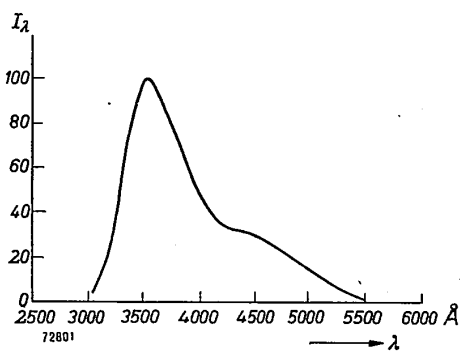



Fig. 4. Spectral distribution curve of $\beta Ca_2P_2O_7$-Sn.Excitation with 2537 Å at room temperature.

TABLE II

decay times of phosphors of group 3 (characteristic emission)

| activator | matrix | decay time (sec) |
|---|---|---|
| $Mn^{2+}$ | Silicates, phosphates, fluorides | $10^{-1}$ to $10^{-3}$ |
| $Mn^{4+}$ | $Mg_2TiO_4$ | $10^{-3}$ (U.V. excitation) Kröger a.o.[7]) |
| $Cr^{3+}$ | $Al_2O_3$ | $16.10^{-3}$ Alyardin a.o. [25]) |
| | | $4.10^{-3}$ Strange & Henderson [21]) |
| $Sn^{2+}$ | $\beta$-$Ca_2P_2O_7$ | $7.10^{-6}$ |
| | $NaCaPO_4$ | $6.10^{-6}$ |
| $Sb^{3+}$ | apatites | $< 10^{-4}$ (U.V. exc.) |
| | | $6.10^{-4}$ (U.V. exc.) Studer and Rosenbaum [26]) |
| | MgS | $10^{-6}$ Leverenz [27]) |
| $Tl^{+}$ | NaJ | $<10^{-6}$ |
| | | 4 to $7.10^{-7}$ Elmore and Hofstadter [28]) |
| | $Ca_3(PO_4)_2$ | $1.10^{-6}$ |
| $Pb^{2+}$ | $CaSiO_3$ | $2.10^{-3}$ (U.V. exc.) Studer and Rosenbaum [26]) |
| | $BaSO_4$ | $1.10^{-6}$ |
| | | $8.10^{-7}$ Marshall [22]) |
| $Bi^{3+}$ | $\beta Ca_2P_2O_7$ | $2.10^{-6}$ |
| $Ce^{3+}$ | phosphates | $<4.10^{-7}$ |
| | silicates | $<4.10^{-7}$ |
| | $2CaO.MgO.2SiO_2$ | $\approx 10^{-6}$ Smith [29]) |
| | fluorides | $\approx 10^{-6}$ |
| | | $10^{-4}$ Kröger and Bakker [30]) |

citation, but he found it to be shorter than $10^{-4}$ sec, the shortest decay time that could be detected with his set-up. His results therefore, do not confirm the figure reported by Studer and Rosenbaum [26]). The same applies to their values for $Pb^{2+}$. $Sn^{2+}$-activated phosphors show a blue or an ultra-violet emission. The spectral distribution of $\beta Ca_2P_2O_7$-Sn is shown in fig. 4. The $Ce^{3+}$-activated phosphors are the fastest of group 3. The Ce-emission corresponds to allowed transitions from a $^2F$-level to the doublet $^2D$ ground state of the ion. The values of $10^{-4}$ quoted by Kröger and Bakker [30]) are erroneous. The efficiency of many Ce-activated phosphors under cathode-ray excitation is quite good; the best ones have an energy efficiency of about 4%. They are therefore superior to U.V.-fluorescent ZnO for black-and-white television.

## 4. The use of Ce-activated phosphors in flying-spot cathode-ray tubes

Many Ce-activated phosphors were made and their efficiency and spectral distribution measured under cathode-ray excitation. Some of the results are shown in figs 5 and 6. Among the phosphates and also among the silicates, phosphors were found with good efficiencies. To select those most suitable for the flying-spot tube one has to take into account absorp-

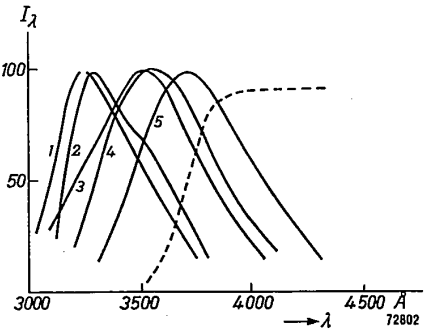



Fig. 5. Spectral distribution curves of various phosphates activated with cerium. Excitation with cathode rays.
(1) $Sr_3(PO_4)_2$-0·015 $Ce_2O_3$; (2) $KBaPO_4$-0·01 $Ce_2O_3$;
(3) $NaBaPO_4$-0·002 $Ce_2O_3$; (4) $KCaPO_4$-0·001 $Ce_2O_3$;
(5) $NaCaPO_4$-0·01 $Ce_2O_3$.
----- Transmission curve of the Ce-containing glass face of projection tubes.

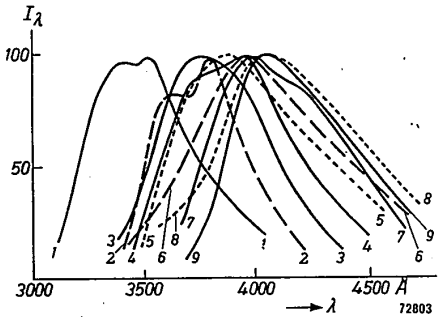



Fig. 6. Spectral distribution of curves of various silicates activated with cerium. Excitation with cathode rays.
(1) $CaO.Al_2O_3.2SiO_2$-0·01 $Ce_2O_3$; (2) $La_2O_3.2SiO_2$-0·08 $Ce_2O_3$
(3) $2CaO.MgO.2SiO_2$-0·03 $Ce_2O_3$; (4) $2\,CaO.ZnO.2SiO_2$-0·03 $Ce_2O_3$
(5) $3Al_2O_3.2SiO_2$-0·3 $Ce_2O_3$; (6) $Na_2O.CaO.2SiO_2$-0·01 $Ce_2O_3$
(7) $2\,CaO.Al_2O_3.SiO_2$-0·015 $Ce_2O_3$; (8) $3\,CaO.2SiO_2$-0·1 $Ce_2O_3$.
(9) $2\,SrO.SiO_2$-0·04 $Ce_2O_3$.

tion losses in the glass face of the tube and in the lenses of the optical system. The tubes used are of the projection type. Ordinary glass windows of such tubes show two kinds of discoloration at the high voltages at which these tubes are run [32]). The soft X-rays produced by the electron bombardment and the electrons penetrating through the screen both darken the glass. The darkening by X-rays can be reduced to a large extent by the incorporation of Ce in the glass. To avoid both types of discoloration a special Ce-containing glass was developed by De Gier and Smelt [34]). The transmission curve of this glass is shown by the dashed line of fig. 5. The Ce-activated phosphates, whose emission is mostly in the short U.V. range, are therefore not suitable in a tube made of this glass. Of the silicates, CaMg silicate developed by Smith [29]) and 2 $CaO.Al_2O_3.SiO_2$ showed the highest efficiencies. Of these two phosphors the latter is to be preferred because its emission is at longer wavelengths, so that less light is absorbed in the face of the cathode-ray tube.

**Acknowledgements**

Our thanks are due to Dr F. A. Kröger and Mr P. Zalm for the preparation of most of the phosphors used for our measurements and to Dr Th. P. J. Botden for the measurements of the spectral distributions.

*Eindhoven, March 1952*

## REFERENCES


[1]) L. H. Bedford and O. S. Puckle, J. Inst. El. Eng. **75**, 63, 1934.
[2]) W. Schnabel, Arch. Elektrotech. **30**, 461, 1936.
[3]) J. J. Müller, Hochfrequenztech. Elektroak. **54**, 111, 1939.
[4]) A. Bril, Physica **15**, 361, 1949.
[5]) G. R. Fonda, J. appl. Phys. **10**, 408, 1939.
[6]) G. F. J. Garlick and M. M. F. Wilkins, Proc. roy. Soc. London A **184**, 408. 1945.
[7]) F. A. Kröger, W. Hoogenstraaten, M. Bottema and Th. P. J. Botden, Physica **14**, 81, 1948.
[8]) F. A. Kröger and W. Hoogenstraaten, Physica **14**, 425, 1948.
[9]) W. Hoogenstraaten and F. A. Kröger, Physica **15**, 541, 1949.
[10]) F. A. Kröger and W. Hoogenstraaten, Physica **15**, 557, 1949.
[11]) A. Karolus, Bull. Schweiz. elektrotech. Ver. **40**, 566, 1949.
[12]) R. B. Nelson, R. P. Johnson and W. B. Nottingham, J. appl. Phys. **10**, 335, 1939.
[13]) N. A. Tolstoi and P. P. Feofilov, Dokl. Ak. Nauk SSSR **59**, 235, 1948.
[14]) R. E. Shrader and H. W. Leverenz, J. opt. Soc. Am. **37**, 939, 1947.
[15]) F. H. Nicoll, J. opt. Soc. Am. **38**, 817, 1948.
[16]) F. I. Vergunas and F. F. Gavrilov, J. exp. theor. Phys. USSR **18**, 224, 1948.
[17]) R. Monnot, Onde électrique **30**, 362, 1950.
[18]) A. Bril and H. A. Klasens, Philips Res. Rep. **7**, 421, 1952.
[19]) W. Schnabel, Arch. Elektrotech. **30**, 461, 1939.
[20]) H. B. Briggs, J. opt. Soc. Am. **31**, 543, 1941.
[21]) J. W. Strange and S. T. Henderson, Proc. phys. Soc. **58**, 369, 1946.
[22]) T. H. Marshall, J. appl. Phys. **18**, 512, 1947.
[23]) S. J. Vavilov and V. L. Levshin, Z. Phys. **48**, 397, 1928.
[24]) R. Delorme and F. Perrin, J. de Phys. **10**, 177, 1929.
[25]) V. N. Alyardin, V. V. Fedorov and V. L. Levshin, C. R. Ac. Sc. URSS **25**, 106, 1939.
[26]) F. J. Studer and A. Rosenbaum, J. opt. Soc. Am. **39**, 685, 1949.
[27]) H. W. Leverenz, Luminescence of solids, John Wiley & Sons, New York 1950, p. 180.
[28]) W. C. Elmore and R. Hofstadter, Phys. Rev. **75**, 203, 796, 1949.
[29]) A. L. Smith, Trans. electrochem. Soc. **96**, 287, 1949.
[30]) F. A. Kröger and J. Bakker, Physica **8**, 628, 1941.
[31]) F. A. Kröger, J. Th. C. Overbeek, J. Goorissen and J. van den Boomgaard, Trans. electrochem. Soc. **96**, 132, 1949.
[32]) J. de Gier, Philips techn. Rev. **10**, 97, 1948.
[33]) F. Eckert, Z. tech. Phys. **7**, 300, 1926
[34]) J. de Gier and J. A. M. Smelt, U.S.A. Patent **2**, 477, 329.
[35]) P. Zalm, private communication.
[36]) F. Roberts and J. Z. Young, paper delivered at the 1952 convention of the Institute of Electrical Engineers, London; Nature **167**, 231, 1951.

# DIMENSIONAL ANALYSIS, UNITS AND RATIONALIZATION

by R. VERMEULEN 537.081.5:621.3.081.5

**Summary**

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.

**Résumé**

Une jonglerie imprudente de formules dimensionelles peut aboutir à des contradictions, par example que 1 sec $= 3.10^{10}$ cm ou que 1 oersted = 1000 A/m. On peut les éviter en définissant la multiplication de quantités physiques - ce que, singulièrement, on n'a jamais fait - et en évitant de confondre des multiplications dont les interprétations physiques sont essentiellement différentes. Les mêmes mesures résolvent les difficultés que comporte le système rationalisé de Giorgi.

**Zusammenfassung**

Ein unterschiedsloser Gebrauch von Maßformeln kann zu Widersprüchen führen, z.B. daß 1 sec $= 3.10^{10}$ cm oder daß 1 Oersted = 1000 A/m. Diese werden beseitigt durch genaue Festlegung der Multiplikation physikalischer Größen, wie sie ganz merkwürdigerweise bisher noch niemals vorgenommen worden ist, und durch Vermeidung der Verwechslung von Multiplikationen, deren physikalische Deutungen wesentlich verschieden sind. Dieselben Maßnahmen lösen Schwierigkeiten, die sich bei der Einführung des rationalisierten Giorgi-Systemes ergeben.

The problem I want to discuss is stated in an admirably concise and clear way by Pohl [1]):

"Gleichsetzungen von Einheiten des Vierersystems mit Vielfachen von Einheiten eines Dreiersystems sind immer unzulässig. Das geht aus der Literatur zur Genüge hervor. Wie oft muß man beispielsweise lesen, es sei die elektrische Elementarladung

$$e = 4{,}8.10^{-10} \text{ el. stat. CGS} = 1{,}6.10^{-20} \text{ el. magn. CGS} = 1{,}6.10^{-19} \text{ Coulomb.} \quad (24$$

Wir ersetzen die nichtssagenden Buchstaben CGS durch die Definitionsgleichungen

$$\begin{aligned} &1 \text{ elektrostat. CGS Ladungseinheit} = 1 \text{ cm}^{3/2}\, g^{1/2}\, \text{sec}^{-1} \\ &1 \text{ elektromagn. CGS Ladungseinheit} = 1 \text{ cm}^{1/2}\, g^{1/2} \end{aligned} \quad (25)$$

und erhalten aus Gl. (24)

$$1 \text{ sec} = 3.10^{10} \text{ cm}.$$

Diese Gleichung macht also die Grundeinheit sec zu einem Vielfachen der Grundeinheit cm ! (A) Ist es nicht an der Zeit, mit derartiger Wirrnis einmal Schluß zu machen? (B) Soll man nicht endlich zugeben, daß man weder die elektrostatischen noch die elektromagnetischen Einheiten der Dreiersysteme durch Gleichheitszeichen mit Einheiten des Vierersystemes, mit Coloumb, Amp. usw. verbinden darf?"

My answer to question (A) is an outright "Yes, the sooner the better", but I do not like the consequences of the proposition expressed in question (B). I am quite convinced that

$$e = 4{\cdot}8.10^{10} \text{ el. stat. CGS, say that } e = p,$$

but I am equally sure that

$$e = 1{\cdot}6.10^{-19} \text{ Coloumb, say that } e = q,$$

and now I am asked to accept the situation that, although $e = p$ and $e = q$, I am not permitted to write down $p = q$. This is a most drastic interference with my habits of logic.

In my opinion the fallacy should not be sought for in equations (24) but in manupilation of (25). These two formulae contain products of powers of physical quantities, and nobody has ever told me what these mathematical operations with physical quantities really stand for. So far as I know, they have no meaning in themselves at all. They are just symbolic expressions of physical relations, and when we do not define properly what is meant by them, we can never be sure that they are not ambiguous. This is even the case when, through lifelong association, we have become so accustomed to them that they have a most homely and trustworthy look; as for instance in the case of the product of two lengths. One should not define the physical quantity "area" as the product of two lengths, but on the contrary, define the product of two lengths as the area of a rectangle with sides equal to these lengths. We have to realize that this definition is only a convention, a choice, though a wise one, out of a multitude of possibilities. For surely there are practical, but not at all fundamental, reasons why we should not define the product of two lengths as, for instance, the area of the ellipse with these lengths as axes [1]).

If now, as Pohl shows so clearly, the unrestrained manipulation of mathematical symbols applied to physical quantities leads to a contradiction, we should first of all examine critically the possibility that these symbols are ambiguous. Where do the formulae (25) originate from? The first one is derived from the equation

$$Q_1 \cdot Q_2 = F \cdot r^2, \tag{1}$$

[1]) The use of the "circular inch" and "circular mil" is based on this conception of the product of two lengths.

expressing the law of Coulomb in the electrostatic system. This is not really a definition of the physical quantity "electric charge", but it is the definition of the product of two electric charges, presupposing that the meaning of the operations with mechanical quantities is already known and agreed upon. The second one results from the equation

$$I_1 \times I_2 = \tfrac{1}{2} F \cdot r/l, \tag{2}$$

relating the force $F$ between two parallel conductors of length $l$ and separated by the distance $r$, to the currents $I_1$ and $I_2$ through them. As the current determines the charge $Q$ that has passed in the time $t$, this equation also defines a product of electric charges:

$$Q_1 \times Q_2 = \tfrac{1}{2} F \cdot t^2 \, r/l. \tag{3}$$

Equations (1) and (3) are two definitions of the product of electric charges that are quite different, and it is really not amazing at all that by not making any distinction between these essentially different symbolic products, we are led to the paradox that a length equals a certain time. If, however, a clear distinction is made (as has been done here by using the different symbols $\cdot$ and $\times$), and if we cling to their physical meaning throughout all manipulations, there can result no contradiction.

Quite an analogous paradox emanates from the introduction of "rationalization", and the proffered solution is not a real one. The exposition of the problem is perhaps most clearly achieved by treating the example given by De Boer [2]). That we arrive at opposite results is caused by a different notion of the concept "unit". In my opinion an excellent definition of "unit" is: "a particular reference-sample of a physical quantity, chosen as a standard". I have, however, a strong doubt that this conception of "unit" is not compatible with the notion used in dimensional analysis, and that from this unconscious difference of premise, the discord and the confusion in this field result. In dimensional analysis the "units" are defined by mathematical formulae expressing all the "derived units" in terms of three or four "fundamental units". Therefore, one is apt to consider these dimensional formulae, like m/sec; g.cm$^2$.sec$^{-2}$; A/m; etc. as the "units". As in the case of the symbolic product, this is taking the symbol for the real thing, and will lead to paradoxes, when the symbols are not unambiguous.

Consider, as De Boer has done, the kinetic energy of a mass $m$ moving with velocity $v$:

$$E = \tfrac{1}{2} m \cdot v^2. \tag{4}$$

From this eqation it can be seen at once that the unit of kinetic energy

(to be indicated by $[E]$) is the kinetic energy of a mass of 2 kg, moving at the velocity of 1 m/sec, and not, as one would possibly presume at first glance, the kinetic energy of a mass of 1 kg moving at the velocity of 1 m/sec Therefore:

$$[E] = \tfrac{1}{2} \cdot 2\,\text{kg} \cdot (1\ \text{m/sec})^2 = \tfrac{1}{2} \cdot 2[\text{m}] \cdot [v]^2, \tag{5}$$

$$[E] = \text{kg} \cdot (\text{m/sec})^2 = [\text{m}] \cdot [v]^2. \tag{6}$$

I want to stress the point that the meaning of equation (6) is that the unit of kinetic energy is the kinetic energy of double the unit mass, moving at unit velocity, although the number 2 does not appear in it. Surely this "unit-equation" is a rather obscure way of expressing this fact and one should not wonder that such formal expressions can easily lead us astray.

Equation (4) contains a factor $\frac{1}{2}$. It is generally agreed upon that this is the most convenient form for the kinetic energy, but there are no fundamental reasons why we may not, should the fancy take us, define a quantity $A$ by the equation

$$A = m \times v^2. \tag{7}$$

Obviously the unit $[A]$ will occur when $m$ = unit mass $[m]$ and $v$ = unit velocity $[v]$. The physical phenomenon in this situation will be only half as intense as in the situation where $[E]$ occurs. It seems reasonable and convenient to express this fact by the equation:

$$[E] = 2 \cdot [A]. \tag{8}$$

However, and this is a conclusion that will upset us, introducing $m = [m] =$ kg and $v = [v] =$ m/sec in (7) gives:

$$[A] = [m] \times [v]^2 = \text{kg} \times (\text{m/sec})^2. \tag{9}$$

There seems to be a flagrant contradiction between (6) and (9) on the one side and (8) on the other. When we forget about (8) or reject this expression, we are led to the conclusions (*i*) that $A = 2E$, and (*ii*) that $[A] = [E] = \text{kg(m/sec)}^2$; in other words that (7) defines a new quantity $A$, and that by doing so the unit remains unchanged. When however, we look at the physical representations, where $[E]$ and $[A]$ are realized, i.e. the "standards" for $[E]$ and $[A]$, we see that the physical phenomena are twice as intense in the one as in the other. The proposition that the "unit" has not changed can therefore be defended only when we are ready to accept a situation where

(a) "unit" means the mathematical formula: kg · (m/sec)²;
(b) "standard" and "unit" are quite different notions;
(c) two standards differing by a factor 2 correspond to one and the same "unit";

(d) only one symbol i.e. kg(m/sec)$^2$ is available for indicating these two different standards;

(e) the symbol "3 kg . (m/sec)$^2$" is meaningless in itself, and has only a physical content when the quantity is mentioned explicitly, e.g. $E = 3$ kg (m/sec)$^2$ corresponds to half the energy that $A = 3$ kg (m/sec)$^2$ stands for.

All this I consider most confusing and misleading and, moreover, these ambiguities are quite unnecessary, as I hope to show. I want to retain the conception of a "unit" as a "sample of the quantity chosen as a standard" and instead of rejecting (8) I most emphatically want to accept it as the true expression of the physical facts. Combining (8) with (6) and (9), we have to conclude

$$[E] = [m] \cdot [v]^2 = 2\,[m] \times [v]^2 = 2\,[A]. \tag{10}$$

This can only be valid when the multiplication signs in (4) and (7) stand for two different kinds of products, having different physical meanings. They must therefore be indicated by different symbols as has been done. Their definitions may be:

$m \times v^2$ stands for the kinetic energy of a mass $m$ moving with velocity $v$ and $m \cdot v^2$ stands for double the same quantity.

From this follows directly that

$$E = \tfrac{1}{2}\, m \cdot v^2 = m \times v^2 = A, \tag{11}$$

so that, in spite of the superficial suggestion that $A = 2E$, we find

$$E = A \tag{12}$$

and

$$\text{kg} \cdot (\text{m/sec})^2 = 2\ \text{kg} \times (\text{m/sec})^2. \tag{13}$$

**By changing the constant $\frac{1}{2}$ in (4) into the constant 1 in (7), we have not introduced any new physical quantity, but a new unit twice as small.**

After the discussion of this example we will treat the problem more generally:

We consider a certain phenomenon $A$, uniquely determined in a given physical constellation by two quantities $P$ and $Q$, and we suppose that we can express the relation between the three quantities $A$, $P$, and $Q$ by an equation.

$$A = c \cdot P \cdot Q. \tag{14}$$

The problem is: Are we free to choose arbitrarily the value of the constant $c$? And if so, what will then be the relation between $A_1$ and $A_2$, satisfying two equations the one with $c = c_1$, the other with $c = c_2$?

First we have to introduce the concept of "coherent units" these being

units so chosen that the equation expressing the relation between the numerical values associated with the physical quantities will have exactly the same form as the equation expressing the relation between the physical quantities themselves. When we indicate the numerical value by $\{A\}$ and the unit by $[A]$, then

$$A = \{A\}\ [A]. \tag{15}$$

The condition that both

$$A = c \cdot P \cdot Q. \tag{16}$$

and

$$\{A\} = c \cdot \{P\} \cdot \{Q\} \tag{17}$$

shall be valid, requires that the coherent units satisfy

$$[A] = [P] \cdot [Q]. \tag{18}$$

This relation seems independent of the value of $c$ and this has given rise to the fallacy that by a change in the constant $c$ the coherent units are not affected, but that a new physical quantity is introduced. That is, from the two equations

$$A_1 = c_1 \cdot P \cdot Q, \tag{19a}$$

$$A_2 = c_2 \cdot P \cdot Q, \tag{19b}$$

it is concluded that

$$A_1 = (c_1/c_2) \cdot A_2, \tag{20a}$$

$$[A_1] = [A_2]. \tag{20b}$$

This, however, is false, as we see when we remember that the unit is defined as "a sample of the quantity, taken as a standard". The unit $[A_1]$ is the phenomenon $A$ that occurs for instance in the constellation defined by $P = (1/c_1) \cdot [P]$ and $Q = [Q]$ while the unit $[A_2]$ is the phenomenon occuring in the constellation where $P = (1/c_2) \cdot [P]$ and $Q = [Q]$. The phenomenon is thus $(c_2/c_1)$ times as intense in the constellation where $[A_1]$ occurs as in the constellation producing $[A_2]$. Such a finding is conveniently expressed by

$$[A_1] = (c_2/c_1) \cdot [A_2]. \tag{21}$$

However, because they are coherent units both have to satisfy

$$[A] = [P]\,[Q] \tag{22}$$

which will be in sheer contradiction with (21), unless the multiplications have different meanings. When this is admitted, the following equations:

$$[A_1] = [P] \cdot [Q], \tag{23a}$$

and

$$[A_2] = [P] \times [Q] \tag{23b}$$

are compatible with (21) provided that

$$c_1 [P] \cdot [Q] = c_2 [P] \times [Q], \tag{24}$$

and therefore

$$c_1 P \cdot Q = c_2 P \times Q. \tag{25}$$

Remembering that

$$A_1 = c_1 P \cdot Q, \tag{26a}$$

$$A_2 = c_2 P \times Q, \tag{26b}$$

we conclude that

$$A_1 = A_2. \tag{27}$$

*Conclusions*

(1) The value of the constant $c$ can be chosen at will in fundamental equations.
(2) It is even admissable to choose for this constant a physical quantity such as $\varepsilon_0$ or $\mu_0$.
(3) A change in this constant implies a change in coherent units.
(4) No change in the physical quantities is involved.
(5) The unrestricted manipulation of undefined mathematical operations with physical quantities can lead to a paradox.
(6) This paradox can be solved by distinguishing between those operations that are essentially different.

To Dr W. de Groot of this Laboratory I am indebted for an instructive example. Imagine two countries: in the one people have the habit of measuring the "thickness" of their trees with callipers, in the other they use a piece of string slung around the tree. Both have the same unit of length: the inch. When they start to trade their trees, a terrible confusion results because their different measuring methods give different values for the thickness of the trees. A solution for their difficulties could be found by abolishing the name "thickness" and introducing two different names for the results of the measurement e.g. "diameter" and "circumference". These could then be measured in one and the same unit: the inch. But another solution can be found by establishing a standard for "thickness": a sample of this property, realized by a standard tree. By comparing any tree with this standard, the same result for its thickness will be found, notwithstanding the use of different measuring methods. This solution involves only a new unit for the quantity "thickness" and no change in the habits of thought nor of the associations connected with the property "thickness". The conversion-factors for expressing the "thickness" in "inches" will be different in both countries. On the other hand, the introduction of two different quantities; "diameter" and "circumference" with one "unit" the

inch, would nevertheless necessitate two standards for thickness: one a tree with a diameter of an inch, one with a circumference of an inch.

Now we have still to apply our conceptions to the problem of the interpretation of rationalization. Here a paradox results for instance when we try to compare the Giorgi-unit for the magnetic field strength $H$ with the non-rationalized electromagnetic unit. The confusion has gone even so far that it has been declared that it is not possible to give any conversion factors between CGS and Giorgi-units. When dimensional analysis leads to contradictions that can only be evaded by negating the possibility of expressing the results of rather simple measurements, surely it should be mended. This can be done by again remembering what the symbols used stand for in reality.

The conflict we will try to settle is that experimentally it can be established that [1])

$$1 \text{ oersted} = (1000/4\pi) \text{ A/m}, \tag{28}$$

while dimensional formulae give us

$$1 \text{ oersted} = 1 \text{ biot/1 cm} = 10 \text{ ampere/0·01 m} = 1000 \text{ A/m}. \tag{29}$$

Now the solution that has been proposed is to introduce two different physical quantities: the rationalized magnetic fieldstrength $'H$ and the non-rationalized magnetic field strength $''H$, and to keep as valid the equation (29) 1 Oe = 1000 A/m. Then we would be forced to enounce the experimental comparison by the complicated statement that "a field where the non-rational $''H = 1$ Oe is just as strong as the field where the rational $'H = (1000/4\pi)$ A/m". The consequence of such a proposal is that for the "*unit*" A/m two different *standards* will exist, one where the rationalized $'H = 1$ A/m, the other where the non-rationalized $''H = 1$ A/m. The intensities of the fields in the two standards will differ by a factor $4\pi$. The result of all this will be that two different standards have to be indicated by one and the same symbol: the "unit" A/m. In my opinion this will result in confusion and misunderstanding.

Another solution, however, is possible when we admit different kinds of operations, rational and non-rational ones, and indicate them with different symbols, when necessary. Defining roughly the product of magnetic field strength and length indicated by $(H \cdot L)$ as the enclosed current, and the product indicated by $(H \times L)$ as $4\pi$ times the enclosed current, the contradiction disappears without having to introduce two different magnetic field strengths. We shall need also different symbols for the inversions

[1]) A/m is the unit for $H$ in the Giorgian, i.e. rationalized system; biot (Bi) is the name proposed for the unit for electric current in the electromagnetic system with four fundamental units; 1 biot = 10 amperes.

of these products; we can for instance distinguish them by one or two accents. The fundamental equations can now be written without any contradiction:

$$\begin{array}{ll} \text{Rationalized} & \text{Non-rationalized} \\ I = \oint H \cdot ds & 4\pi I = \oint H \times ds \end{array} \tag{30}$$

The unit-equations therefore are:

$$[I] = [H] \cdot [L] \qquad [I] = [H] \times [L] \tag{31a}$$

$$\text{ampere} = (\text{A/m}) \cdot \text{m} \qquad \text{biot} = \text{oersted} \times \text{cm} \tag{31b}$$

These can also be written as:

$$(\text{A/m}) = \text{A}\,'/\text{m} \qquad \text{oersted} = \text{Bi}\,''/\text{cm}. \tag{32}$$

The following rules hold:

$$4\pi I = 4\pi H \cdot L \; H \times L$$

$$\begin{array}{ll} (I\,'/L) \cdot L = I & (I\,'/L) \times L = 4\pi I \\ (I\,''/L) \cdot L = (1/4\pi)\, I & (I\,''/L) \times L = I \end{array} \tag{33}$$

From these we get as a result:

$$(\text{A/m}) \cdot \text{m} = (1/4\pi)\,(\text{A/m}) \times \text{m} = \text{ampere} \tag{34}$$

$$\text{oersted} \times \text{cm} = 4\pi \text{ oersted} \cdot \text{cm} = \text{biot}. \tag{35}$$

All relations and conversion factors can now be written down without any ambiguity:

$$\begin{aligned} \text{oersted} &= \text{Bi}\,''/\text{cm} = 10\ \text{A}\,''/\,0{\cdot}01\ \text{m} = 1000\ \text{A}\,''/\,\text{m} = \\ &= (1/4\pi)\ \text{Bi}\,'/\,\text{cm} = (1/4\pi).\ 10\ \text{A}\,'/\,0{\cdot}01\ \text{m} = (1000/4\pi)\ \text{A}\,'/\,\text{m}. \end{aligned}$$

Now that we have available different symbols for the rational standard and the non-rational standard, we can consider the "unit" again as a "sample of the quantity". It will, however, be a wise measure to use the symbol A/m (without any accents) only for the rational unit. To distinguish between rational and non-rational operations is only necessary when using them both in one and the same equation e.g. in calculating conversion factors. In all other cases it has to be evident from the context whether the rational or non-rational system is used.

I surmise that the conception propounded here is essentially the conception all electrical engineers and most of the experimental physicists have been using more or less unconsciously, and that this is the reason why they always wanted to interpret rationalization as a change in units. Now that a sound theoretical foundation has become available, they can

continue to do so without the slightest compunction. I hope that the supporters of dimensional analysis can be persuaded to revise it so as to enable it to express the experimental and physical facts in a way everybody can understand.

*Eindhoven, Juli 1952*

REFERENCES

1) R. W. Pohl, Zur Darstellung der Elektrizitätslehre, Göttingen, 1950, p. 9.
2) J. de Boer, Ned. Tijdschr. Natuurk. **16**, 305, 1950.
3) J. P. Schouten, De Ingenieur **64**, O 103-O108, 1952.

R 203 Philips Res. Rep. 7, 442-466, 1952

# DELTAMODULATION, A METHOD OF P.C.M. TRANSMISSION USING THE 1-UNIT CODE *)

by F. de JAGER

621.396.619.16

**Summary**

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 wellknown 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. A reduction of the quantizing noise is acquired by using a combination of single and double integration in this frequency region. Then the ratio between the r.m.s. values of the information signal and the quantizing noise is proportional to the 5/2 power of the pulse frequency.

**Résumé**

On sait que dans un système de communication l'influence des interférences sur le trajet de transmission peut être réduite d'une manière considérable par codification d'abord du signal de renseignement puis par transmission d'une série d'impulsions de la valeur 0 et 1. Dans les systèmes bien connus de modulation codée on utilise le code binaire à $n$ chiffres. Dans la modulation en delta, cependant, on utilise un "code" ne comportant que 1 chiffre. Dans le cas présent le signal reproduit est obtenu par application d'une série d'impulsions exprimées en quanta à un réseau électrique linéaire. Ce système nous permet d'obtenir une simplification tant des dispositifs de codification que de déchiffrement. La transformation du signal de renseignement en une série d'impulsions de la valeur 0 et 1 est obtenue au moyen d'un circuit de contre-réaction dans lequel la tension fournie au réseau du circuit de contre-réaction est exprimée en quanta tant en amplitude qu'en temps. Le réseau dans le circuit de contre-réaction doit se rapporter au spectre moyen des signaux de renseignements. Pour la parole un réseau intégrateur peut être utilisé comme tel. On a constaté, toutefois, que la caractéristique de fréquence du circuit de contre-réaction pour les fréquences situées entre la fréquence vocale supérieure et la fréquence des impulsions a une influence sur le bruit de quantification du signal reproduit. On obtient une diminution du bruit de quantification par combinaison d'une intégration simple et double dans cette zone de fréquence.

*) Paper presented at the Symposium on "Applications of Communication Theory", held in London, September 1952.

Puis le rapport entre le signal de renseignement et le bruit de quantification est proportionnel à la puissance 5/2 de la fréquence des impulsions.

**Zusammenfassung**

Bekanntlich kann in einen Ubertragungssystem der Einfluss von Störungen auf dem Verbindungsweg dadurch wesentlich herabgesetzt werden, dass man zuerst das Informationssignal codiert und dann ein korrespondierendes Impulsmuster von 0- und Einser-Impulsen sendet. Bei den bekannten Systemen der Impuls-Codemodulation kommt der binäre $n$-Ziffer Schlüssel zur Anwendung. Dagegen wird bei der Deltamodulation ein aus nur einer Ziffer bestehender Schlüssel verwendet. Hier erhält man das wiedergegebene Signal, indem man die quantisierte Impulsreihe einem linearen Netzwerk zuführt. Diese Methode ermöglicht es, die Codierungs- wie auch die Decodierungs-Einrichtungen in ihrem Aufbau zu vereinfachen. Die Umwandlung des Informationssignals in ein Muster von quantisierten Impulsen wird erreicht durch die Benutzung eines Gegenkopplungskreises, worin die dem Rückkopplungsnetz zugeführte Spannung sowohl der Amplitude wie der Zeit nach quantisiert wird. Das Netzwerk innerhalb der Rückkopplungsschleife muss in einem bestimmten Verhältnis zu dem mittleren Spektrum der Informationssignale stehen. Zu diesem Zweck kann bei Sprachübertragung ein integrierendes Netzwerk verwendet werden. Es zeigt sich jedoch, dass der Frequenzverlauf des Rückkopplungsnetzes bei den Frequenzen zwischen der höchsten Sprechfrequenz und der Impulsfrequenz einen merklichen Einfluss auf die Stärke des Quantisierungsgeräusches in dem wiedergegebenen Signal ausübt. Eine Herabsetzung des Quantisierungsgeräusches wird durch die Anwendung einer Kombination von einfacher und doppelter Integration in dem betreffenden Frequenzbereich erreicht. Das Verhältnis zwischen den quadratischen Mittelwerten des Informationssignals und dem Quantisierungsgerausch ist dann proportional einer Potenz 5/2 der Impulsfrequenz.

## 1. Introduction

In deltamodulation the information signal is converted into a sequence of binary pulses in such a manner that a reconstruction of it is obtained by applying the series of binary pulses to a linear network. The output signal is built up from the transients of all foregoing pulses and is represented by a smooth curve in spite of the discontinuous character of the input signal. Each pulse exercises a corrective influence on the output signal, and the latter is caused to approach the wanted signal in much the same manner as a car responds to a succession of corrective motions of the steering-wheel.

In this system of pulse-code modulation the receiver consists only of a linear network and a pulse regenerator for elimination of the transmission noise. The main problem is that of ensuring the right sequence of pulses at the sending end.

For this purpose the method of *quantized feedback*, shown in fig. 1, is applied. Here *P.M.* is a pulse modulator which passes pulses from the pulse generator *P.G.* to the network $F_1$ in one polarity if the voltage $\varepsilon$

is positive, and in the opposite polarity if $\varepsilon$ is negative. The output signal $e_1$ of this network is then compared with the signal $e_0$ to be transmitted. The difference meter $D$ produces the difference signal $\varepsilon$ and applies it to the modulator. If the network $F_1$ is chosen in such a manner that the system is stable, each pulse tries to reverse the polarity of the difference signal $\varepsilon$, with the result that the mean difference between the information signal $e_0$ and the output signal $e_1$ of the feedback network $F_1$ is small. When the pulses which issue from the pulse modulator are applied at the receiving end to an analogous network $F_2$, the output signal $e_2$ is a close approximation of the original information signal.

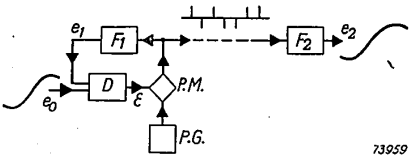



Fig. 1. Basic circuit of a system of deltamodulation. *P.G.* is a pulse generator. *P.M.* is a pulse modulator which passes pulses from the pulse generator *P.G.* in a polarity depending on the polarity of the voltage $\varepsilon$. These pulses are then applied to the network $F_1$. The output signal $e_1$ of this network is compared with the input signal $e_0$, and the difference $\varepsilon$ controls the pulse modulator. When the pulses which issue from the pulse modulator *P.M.* are applied at the receiving end to an analogous network $F_2$, the output signal $e_2$ is a close approximation of the original information signal.

The difference between the original and the reconstructed signals, $e_0$ and $e_2$, gives rise to a "quantizing noise" which can be decreased by increasing the pulse frequency. Contrary to the well-known P.C.M. principle, the information is quantized in a 1-unit code and the "sampling frequency" is made equal to the pulse frequency.

As shown in fig. 2 the system may be seen as developed from an ordinary feedback circuit by inserting in the feedback loop a clipping amplifier $A$ and a periodically operated switch $S$. The characteristics of four types of feedback system, distinguished by the different form of voltage at the input of the network in the feedback loop, are illustrated in fig. 3.

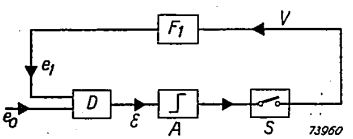



Fig. 2. The method of quantized feedback, illustrated in fig. 1, may be seen as developed from an ordinary feedback circuit by inserting in the feedback loop a clipping amplifier $A$ and a periodically operated switch $S$.

If the amplifier is linear and the switch permanently closed we have the normal feedback system, where the feedback voltage $V$ is continuous and proportional to the difference $e_0$-$e_1$ (fig. 3*a*). If now the switch is closed periodically, the feedback voltage is present only at equidistant time-intervals and zero at other times (fig. 3*b*). In this case the system is working linearly (giving rise to linear difference equations instead of differential equations) and the circuit is used as a "sampling" servo system. We might say that the feedback voltage is now "quantized" in time. If the switch is closed permanently and the amplifier is clipped after infinite amplification, we have the elementary circuit of an "on-off" servomechanism. Here the feedback voltage can have only two different values (fig. 3*c*), depending on the polarity of the difference signal $\varepsilon$, so that the feedback voltage may be said to be quantized in amplitude. The system can now only be described by non-linear differential equations. In the fourth case — which is of most immediate interest — the difference signal is both clipped and pulsed (fig. 3*d*), so that the feedback voltage is quantized in time as well as in amplitude and thus has the same characteristics as the transmitted signal in a pulse-code modulation system. Mathematically the system must be described by non-linear difference equations.

The main characteristics of these four systems are combined in Table I.

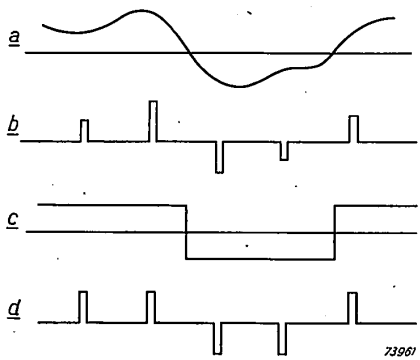



Fig. 3. The form of the voltage $V$ in fig. 2. In *a* the amplifier $A$ is supposed to be linear and the switch $S$ permanently closed, so that $V$ is continuous. In *b* the switch $S$ is closed periodically, so that $V$ is quantized in time. In *c* the amplifier is used as a clipping amplifier - with $S$ permanently closed - so that $V$ is quantized in amplitude. In *d* this voltage is quantized both in time and in amplitude by means of the clipping amplifier $A$ and the periodically operated switch $S$.

TABLE I

| Type of amplifier | Switch closed | Feedback signal | Circuit used in |
|---|---|---|---|
| linear | permanently | continuous | normal feedback |
| linear | periodically | pulsed | "sampling" servo |
| clipping | permanently | clipped | "on-off" servo |
| clipping | periodically | quantized | deltamodulation |

## 2. The method of single integration

The simplest system of deltamodulation is obtained by using an integrating network in the feedback loop. If this network has a large time-constant, the response to an impulse is practically a unit step function

as indicated in fig. 4. (Here the input signal represents an impulse of high amplitude and very short duration). The signal $e_1$ is then built up from positive and negative pulses and has the form of a step curve oscillating around the information signal $e_0$ in the manner indicated in fig. 5. By applying the corresponding series of pulses to another integrating network at the receiving end, the same approximation to the original is obtained, which can be smoothed further by passing it through a low-pass filter. This method has already been described as a method which may be used for the transmission of speech [1]), [2]), [3]), [4]).

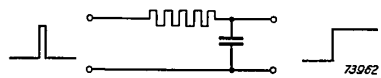



Fig. 4. Each pulse is converted into a step function by means of an *RC* circuit with a large time constant.

As the number of signals which can be transmitted per second is finite, there is some difference between the input signal $e_0$ and the smoothed output signal $e_2$. This is audible as a type of noise which is more or less correlated to the speech. At a pulse frequency of 40 kc/s the intelligibility of the speech is good, but the quantizing noise has an effect on the speech which may be called "sandiness". At higher pulse frequencies this effect is diminished by the better approximation of the signal by the step curve.

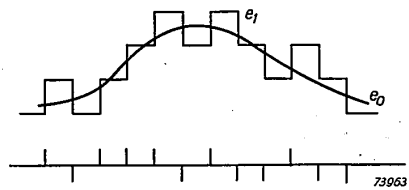



Fig. 5. The approximation signal $e_1$ and the original signal $e_0$ in the system of single integration, together with the pulse series corresponding to $e_1$.

Zero level is transmitted by an alternating pulse series and the approximation signal $e_1$ has the form given in fig. 6. As only high frequencies are present in this signal, the output signal $e_2$ after the low-pass filter will also be zero. From the signal $e_1$ which is alternately higher and lower than the input signal, it is seen, however, that small signals cannot affect this sequence of pulses and thus will be suffering from a threshold effect.

The information contained in the transmitted pulses is mainly correlated to changes of input signal and not to its amplitude. In connection with this,

an overloading of the system begins when the slope of the input signal reaches a certain limit. For a sine wave of frequency $f$ applied to the input, the maximum amplitude which can be transmitted is

$$a = \frac{f_i}{2\pi f}\,\sigma,$$

where $f_i$ is the pulse frequency and $\sigma$ the height of one step in the approximating step function. Thus both the maximum amplitude and the number of distinguishable levels decrease with increase of the frequency to be transmitted. If the information signal were to have the spectrum of white noise this would raise a serious difficulty, but fortunately in the case of speech the higher frequencies contain less energy than the lower ones so that in this respect an integrating network is well adapted to the transmission of speech.

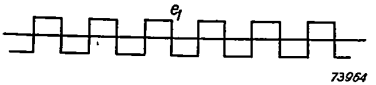



Fig. 6. The approximation signal $e_1$ in the absence of an input signal.

It has been observed as an experimental fact that, as a rule, the system can transmit a speech signal without noticeable overloading if the amplitude of the signal does not exceed the maximum sine-wave amplitude that can be transmitted at a frequency of 800 c/s. It can be shown that, if $A(\omega)$ expresses the mean spectrum of a speech signal, this reference frequency $\overline{f} = \overline{\omega}/2\pi$ would be related to the speech spectrum by the formula

$$\overline{\omega}^2 = \frac{\int_0^{\omega_0} A^2(\omega)\,\omega^2 \mathrm{d}\omega}{\int_0^{\omega_0} A^2(\omega)\,\mathrm{d}\omega},$$

where $\omega_0$ corresponds to the highest frequency to be transmitted *). From the mean intensity-distribution for speech as given by French and Steinberg (ref. 5, fig. 2) we find that for the frequency range 0-3·5 kc/s this reference frequency comes very close to 800 c/s. A comparison of the signal-to-noise ratio in this system with other systems which are limited in amplitude, suggests that it is justifiable to consider the signal-to-noise ratio for a sine wave of 800 c/s.

The negative pulses in the transmission path may be omitted without

*) A derivation of this formula is given in the appendix.

changing the ratio between signal and quantizing noise at the receiving end. This can be seen from the fact that the addition of a periodic series of positive pulses to the original pulse series of fig. 5 has no other effect than a rise of the DC level, as all other frequencies are suppressed by the low-pass filter. On the other hand, this additional pulse series has the effect of doubling the amplitude of the positive pulses and deleting the negative ones. In the same way the pulse modulator at the sending end need not give positive and negative pulses, but may give positive and zero pulses only, depending again on the polarity of the difference signal $\varepsilon$. In this case, with a readjustment of the zero level, the pulse modulator can be simplified considerably. The exact form of the issued pulses is of no importance either, as long as all pulses are of the same form and have vanished before the beginning of a following pulse. For reasons of simplicity however, we shall think in what follows of the approximation signal as being generated by narrow positive and negative pulses.

### 3. The method of double integration

The important question now arises whether the transmission can be improved by using other networks in the feedback loop and at the receiving end respectively, with a view especially to reducing the quantizing noise and the threshold effect.

Formerly, every received pulse had the effect of increasing or decreasing the level of the received signal by a certain amount. Now consider a system where every pulse has the effect of changing the slope instead of the amplitude of the received signal. The reconstructed information signal is then related to the received pulse pattern as shown in fig. 7. The conversion of this pulse series into a signal which is quantized in its first derivative is made possible by using a double-integrating network at the receiving end (fig. 8). If both time-constants are large, a unit pulse at the input provides a unit step function at the first condenser (as in fig. 4) and an output signal of constant slope. In this way the signal of fig. 7 is built

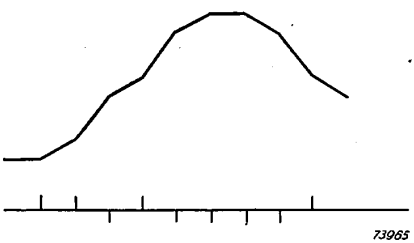



Fig. 7. An approximation signal where each pulse changes the slope instead of the amplitude by a certain amount.

up from a series of pulses where every pulse changes the slope of the output signal by the same amount in the positive or negative direction. When this signal is smoothed by a low-pass filter it may have a closer approximation to the original than when the approximation signal of fig. 5 is smoothed. We will, therefore, investigate how this principle can be used and what improvement can be attained by it.

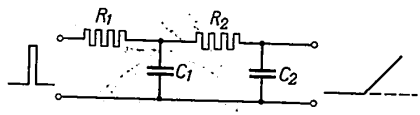



Fig. 8. When this double-integrating network is used, each pulse has the effect of changing the slope in the output signal.

The question of stability arises when this double-integrating network is inserted into the feedback loop. Thus, assuming the input signal $e_0$ to be zero, the signal $e_1$ may cross the zero level with a high value of its first derivative and cause the system to produce negative pulses during a long time-interval. The approximation signal will then be crossing the zero level with a high negative value of its first derivative, so that the oscillation is maintained. These oscillations differ from the oscillations in a normal feedback system in that the oscillatory period can have many different values, depending on the initial conditions. Owing to the quantizing process which makes the feedback voltage related to the polarity of the difference signal and not to its amplitude, there is now a marked relation between the period and the amplitude of these oscillations. Under these circumstances the transmission of an information signal is very poor.

In drawing a picture, however, it is always possible to construct an approximation curve consisting of straight lines, without having difficulty with instability. This is even possible in the case that only specified values of the first derivative are allowed, as in fig. 7. In drawing such an approximation it is usual to make use of the fact that we can see the curve as a whole, and can take into account its further behaviour. We can choose the changes in the derivative of the approximation curve in such a manner as to prevent oscillations. This suggests that, for obtaining an approximation signal of the form of fig. 7, we should make use of some information about the future, by using the principle of *prediction*.

In considering an approximating curve built up from straight lines, the value which will be reached after some time-interval $\tau$ can always be predicted if no further changes in the derivative will occur. Thus use

can be made of an extrapolation of the approximating curve to the dotted points in fig. 9, and these values compared with the values of the original function. If the extrapolated value is lower than the required value, one unit is added to the derivative of the approximating curve, which will tend to make the difference between the two functions smaller. In this way, the approximation can be made stable.

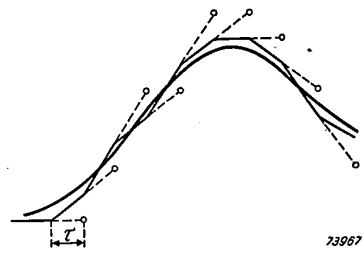



Fig. 9. In this figure the slope of the approximating curve is increased or decreased by one unit according as the extrapolated value of the approximating curve (indicated by the dotted points) is lower or higher than the original curve.

This procedure can now be applied to the electrical system with a view to making the output signal at the second condenser in fig. 8 (which we call $e_2'$) at the time $t + \tau$, equal to the input signal $e_0$ at time $t$. If we can construct a network $F_1$ which generates the extrapolated value of $e_2'$ such that

$$e_1(t) = e_2'(t+\tau),$$

we need only put this network in the feedback loop at the sending end in order to decide whether the slope of the approximation signal should be increased or decreased by one unit, i.e. whether the following pulse should be positive or negative. For this extrapolated voltage

$$e_2'(t+\tau) = e_2'(t) + \tau \frac{de_2'}{dt},$$

where the first term at the right hand side is the voltage on the second condenser of the network in fig. 8.

If $i$ denotes the current through the resistor $R_2$, we have

$$\frac{de_2'}{dt} = \frac{i}{C_2},$$

so that the voltage $e_1(t)$ is found by adding $\tau i/C_2$ to the voltage on $C_2$. This

means that $e_1(t)$ is equal to the output voltage of the network of fig. 10 where

$$\tau = rC_2,$$

and where all foregoing pulses have been applied to the input. The network of fig. 10, to be placed in the feedback loop at the sending end, responds to an impulse in the manner shown in fig. 11.

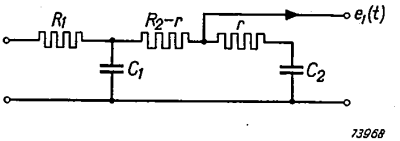



Fig. 10. In this network the voltage $e_1$ corresponds to the voltage that will be observed at the condenser $C_2$ after a short time interval $\tau = rC_2$, provided $\tau$ is small as compared with the time constants $R_1C_1$ and $R_2C_2$.

It is to be supposed that the accuracy of the new approximation curve is greatest for a certain value of $\tau$. For too great a value the points at the ends of the dotted lines in fig. 9 are spread out too widely around the information curve, whereas for too small a value the system tends to be unstable. From elementary considerations it may be expected that the optimal value of $\tau$ is of the order of the time interval between two pulses.

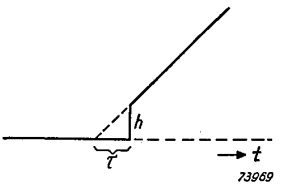



Fig. 11. The response of the network of fig. 10 to an impulse.

An analogy is presented here to the case of driving a car along the road. Then the driver should also look forward over some distance and consider the position of his car in order to know by a mental process of prediction what will be the position of his car at the next moment. Then, by comparing this with the position wanted, he makes the decision whether a corrective motion of the steering-wheel is required in one or the other direction. Thus, driving a car is essentially a feedback system where the information function (the road) has to be approximated by an approximating signal (the car-track), and where an extrapolation over a certain time is necessary. Here also, there is an optimal value for the distance over which one should look at the road ahead; for looking too far ahead would result in cutting off all the windings in the road and thus travelling off the road sometimes,

whereas if one looks only a very short distance ahead the car will start to sway sooner or later as a result of instability of the whole system.

The frequency characteristic $F_1(\omega)$ of the network of fig. 10 has a slope of —12 dB per octave for the lower frequencies and of —6 dB per octave for the higher frequencies. The break-point in the characteristic is found at the frequency $\omega_1$ for which

$$\omega_1 \tau = 1 .$$

The frequency characteristic $F_2(\omega)$ of the network of fig. 8 falls at 12 dB per octave so that both networks $F_1$ and $F_2$ have the same characteristics in the audio frequency band. If these two networks are used in the basic circuit of fig. 1 for a system of speech transmission, a serious difficulty arises in the transmission of the higher frequencies. Though it is true that a signal containing only low frequencies can be approximated to a high degree of accuracy, little advantage can be taken of this fact in practice, as the level of the speech signal has to be lowered so far that the higher frequency components cannot cause over-modulation. Indeed from the point of view of information theory, the network used at the receiving end would be quite unsatisfactory in this case and therefore a further change in the design of these networks is desired.

**4. The connection with information theory**

As much information as possible must be stored in the limited number of pulses which can be transmitted in a given time interval. According to the principles of information theory, this means that a series of pulses with random polarity should already produce a signal with the same frequency spectrum as the mean frequency spectrum of speech, say $A(\omega)$. For in this case the series of pulses required for the transmission of a given signal has a minimum correlation and thus contains a maximum amount of information. Since the spectrum of a series of pulses with random polarity and a relatively high frequency is that of white noise, at least for the low frequency region, it may be inferred that the receiving network should have the same frequency characteristic as the mean spectrum of the speech signals to be transmitted; namely

$$|F_2(\omega)| = A(\omega) .$$

As the spectrum of speech in the first approximation agrees fairly well with the frequency characteristic of an integrating network, we may return to the single-integrating network of fig. 4 at the receiving end.

It is now desirable to change one of the time constants in the feedback network of fig. 10 in such a manner that, in the audio frequency band, the frequency characteristic of the network $F_1(\omega)$ has also the charac-

teristic of a single-integrating network. Then the response function to an impulse commences as shown in fig. 11, but after some time the slope decreases and the response finally becomes constant. This does not interfere with the improvement in approximation gained by using the prediction time $\tau$, for this improvement is related to the beginning of the response function only, or, to put it otherwise, to the fact that the frequency characteristic of $F_1(\omega)$ is falling 12 dB per octave in the higher frequency region up to the frequency $\omega_1 = 1/\tau$ after which it has a negative slope of 6 dB per octave.

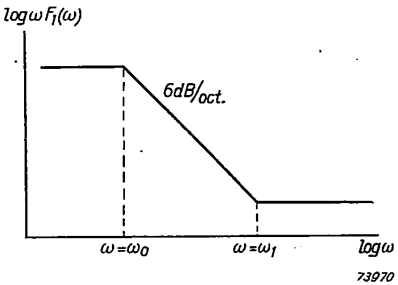



Fig. 12. The frequency characteristic of the feedback network in the system of double integration. Here $\omega_0 = 2\pi f_0$, where $f_0$ is the highest frequency to be transmitted, and $\omega_1 \tau = 1$, where $\tau$ is about equal to the interval between two pulses.

If $\omega F_1(\omega)$ in dB is plotted against $\omega$ on a logarithmic scale, it may be concluded that this characteristic should have the form drawn schematically in fig. 12. Here $\omega_0$ is the highest frequency to be transmitted and $\omega_1 = 1/\tau$, where $\tau$ is about equal to the time interval between two pulses. In order to keep $\omega F_1(\omega)$ constant up to $\omega_0$, an additional branch in the feedback network is used, consisting of a resonance circuit parallel to the condenser $C_2$. The feedback circuit thus has the form as shown in fig. 13. For the case in which a pulse frequency of 100 kc/s is used to

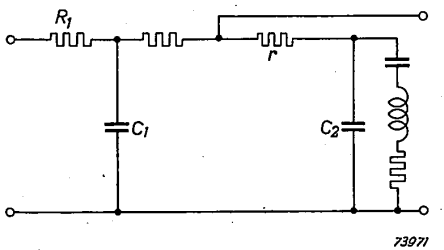



Fig. 13. The network used for realizing the frequency characteristic of fig. 12. The resonance circuit parallel to $C_2$ is used in order to keep $\omega F_1(\omega)$ constant up to $\omega_0$, as in fig. 12.

transmit a speech signal in the band of 200-3800 c/s, the response $u(t)$ to an impulse is given in fig. 14. (The marked time intervals indicate the moments of pulse repetition.) The receiving network consists of an integrating RC circuit with cut-off frequency 200 c/s, followed by a low-pass filter with cut-off frequency 3800 c/s.

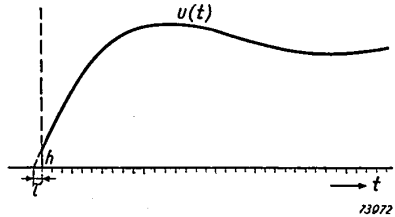



Fig. 14. The response of the network of fig. 13 to an impulse.

The quantizing noise at the output of the receiver can now be measured (in a manner which will be described later) as a function of the time constant $\tau = r\ C_2$ in the feedback circuit of fig. 13, which can be changed by a variation of $r$. The result is given in fig. 15, where $E$ is the r.m.s. value of the quantizing noise in the mentioned audio band and $\sigma$ is, as before, the height of a unit step generated by one impulse. Here $\tau$ is expressed in units of the time interval $T$ between two pulses. It is found that for a value of $\tau$ which is a little smaller than $T$, the quantizing noise reaches a minimum of about 5% of the height of one unit step.

For this value of $\tau$ it was observed that in the absence of an input signal the zero level could be transmitted in two different ways, namely by

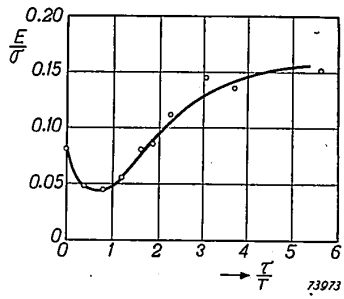



Fig. 15. Measurements of the quantizing noise as a function of $\tau$. Here $E$ is the r.m.s. voltage of the quantizing noise in the band from 200-3800 c/s and for a pulse frequency of 100 kc/s.; $\sigma$ is the height of a unit step generated by one impulse; $\tau = rC_2$ in fig. 13; $T$ is the time interval between two pulses.

means of the pulse patterns

$$+ - + - + - + -$$

and

$$+ + - - + + - - .$$

For greater values of $\tau$ only the first series of pulses was transmitted in the absence of a signal, whereas for smaller values pulse series of longer periods were also observed. Thus when the quantizing noise reaches a minimum, the approximation signal is oscillating to a slight degree around the information signal. The pulses sent from the pulse modulator in this case were + and 0 pulses, the positive ones having a pulse width of 0·5 $T$.

## 5. The quantizing noise in the system of single integration

We are interested now to know the effect of changing the pulse frequency $f_i$. Generally speaking the quality of transmission is improved in two ways by increasing it. First the number of steps can be increased, so that signals of greater amplitude can be transmitted, and secondly a corresponding change in the frequency spectrum of the noise will cause a further reduction of the quantizing noise in the low-frequency band. These two effects are different for the two systems of single and double integration which will therefore be discussed separately.

In the single integration system the noise is produced by the difference signal $\varepsilon = e_1 - e_0$, which, at the receiving end, is smoothed by means of a low-pass filter *). This difference signal has an irregular form as shown in fig. 16. Usually it is non-periodic and has therefore a continuous frequency spectrum.

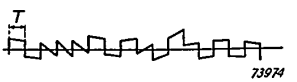



Fig. 16. The characteristic form of the difference signal $\varepsilon$ in the system of single integration.

In general the signal of fig. 16 will show very little correlation for those time intervals which are large compared with $T$. From the well-known relation between the correlation function and the power spectrum of a signal it follows that in this case the spectrum will be uniform for frequencies which are small with respect to $f_i$. The power of the quantizing noise received after a low-pass filter with cut-off frequency $f_0 = \omega_0/2\pi$ will thus be proportional to $f_0$.

*) Strictly, we should take $\varepsilon' = e_1(t) - e_0\,(t - T/2)$, as the approximation signal $e_1$ is, in the mean, retarded over a time interval $T/2$ with respect to the original $e_0$.

If the pulse frequency is doubled, using the same value $\sigma$ for the height of one step, the new difference signal $\varepsilon$ will have the same irregular character shown in fig. 16, with the same amplitude distribution but a different time-scale. This means that the total power in this signal is the same but that the power in the frequency band 0-$f_0$ is only half that of the former case. So the noise power after the low-pass filter at the receiving end is inversely proportional to the pulse frequency $f_i$. Calling $N$ the r.m.s. value of the quantizing noise we thus have

$$N^2 :: (f_0/f_i)\,\sigma^2.$$

The maximum r.m.s. value $S$ of a sine wave of frequency $f$ is proportional to $f_i/f$, as seen previously, so that

$$S :: (f_i/f)\,\sigma.$$

The signal-to-noise ratio $S/N$ can thus be expressed by the formula

$$\frac{S}{N} = c_1 \frac{f_i^{3/2}}{f \cdot f_0^{1/2}}. \qquad (1)$$

## 6. The quantizing noise in the system of double integration

In considering the system of double integration we again start by investigating the difference signal $\varepsilon = e_1 - e_0$. If the response function of the feedback network is as given in fig. 11, $\varepsilon$ has the irregular form of fig. 17, as long as no overmodulation occurs. It turns out that the characteristic form of this signal is practically independent of the manner in which the response function of the feedback network varies for large values of time, so long as the variation is smooth. The difference signals found on applying the response function of fig. 11 and fig. 14 respectively have practically the same characteristic form, so long as the values of $h$ and $\tau$ are the same for both cases. Here again, there is little correlation for time intervals large compared with $T$, so that the spectrum of $\varepsilon$ is uniform in the low frequency band.

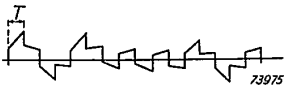



Fig. 17. The characteristic form of the difference signal $\varepsilon$ in the system of double integration.

At the receiving end use may be made either of the network $F_1$ followed by a low-pass filter with cut-off frequency $\omega_0$, or a single integrating network followed by a low-pass filter. This can be seen from the frequency

characteristic of fig. 12. So the quantizing noise at the receiving end can be found by passing the difference signal $\varepsilon$ through a low-pass filter. If the pulse frequency $f_i$ is changed, the time constant $\tau$ must change also, since $\tau$ is related to the time interval between pulses for reasons of stability. With $h$ constant the same irregular form of the difference signal is observed on a different time scale, but again with the same amplitude distribution. Thus, as in the case of single integration, the received noise power is inversely proportional to $f_i$ and, by reason of the spectrum being uniform, proportional to $f_0$. The r.m.s. value of the noise is now related to $h$ instead of $\sigma$ so that

$$N^2 :: (f_0/f_i)\, h^2 .$$

The relation between $h$ and $\sigma$ can be found at once from the frequency characteristic in fig. 12. For, if $u(t)$ is the response of the network $F_1(\omega)$ to an impulse, it is known from transient analysis in electrical networks that

$$u(0) : u(\infty) = \lim_{\omega \to \infty} \omega F_1(\omega) : \lim_{\omega \to 0} \omega F_1(\omega) .$$

Here $h = u(o)$ while $\sigma$ was defined as the height of a unit step at the integrating network at the receiving end. By reason of the corresponding frequency characteristics of $F_1(\omega)$ and $F_2(\omega)$ for the low-frequency band up to $\omega_0$, $\sigma$ is equal to $u(\infty)$. From the characteristic of fig. 12 and the relation $\omega_1 \tau = 1$ we thus find that

$$\frac{h}{\sigma} = \frac{\omega_0}{\omega_1} = \omega_0 \tau .$$

For the higher pulse frequencies, $\tau$ should be varied in proportion to the pulse interval, in order to prevent oscillations, so that

$$\frac{h}{\sigma} :: \frac{f_0}{f_i} .$$

The maximum r.m.s. value $S$ of the sine wave of frequency $f$ that can be generated at the single integrating network at the receiving end is given by

$$S :: \frac{f_i}{f}\, \sigma ,$$

so that for the system of double integration the signal-to-noise ratio $S/N$ is

$$\frac{S}{N} = c_2 \frac{f_i^{3/2}}{f \cdot f_0^{3/2}} . \qquad (2)$$

**7. Numerical values for the constants $c_1$ and $c_2$.**

The improvement in signal-to-noise ratio thus varies with the 3/2 or 5/2 power of the pulse frequency for the systems of single and double integration respectively (whereas it goes exponentially in using the digits of a binary code). The numerical value of the constants $c_1$ and $c_2$ of formulae (1) and (2), however, are not easily found, on account of the high degree of non-linearity in these systems. For evaluating the first constant, $c_1$, we can make use of:

(a) the Fourier spectrum of the quantized signal;

(b) the correlation function of the signal $\varepsilon$;

(c) numerical methods.

The first method is quite exact but very complicated because of the many frequency components which lie in the low-frequency band up to $\omega_0$. The amplitudes of these components all vary considerably with small changes of the input signal. With the second method it is possible to use a first approximation of the correlation function and to evaluate the spectrum of the signal $\varepsilon$ as a continuous spectrum instead of as a line spectrum. This has the advantage that the mean noise power is more easily found than from the Fourier spectrum directly, but on the other hand the result obtained is only approximate. The third method can be used in a straightforward manner by constructing the difference-signal $\varepsilon$, corresponding to a given input signal, over a large number of pulse intervals and then applying a smoothing procedure to the signal in order to determine its noise content in the lower frequency band. In order to avoid a certain periodicity of the quantized signal in this case it is desirable to take two or more sine waves of different frequencies for the input signal.

These methods, when used, gave somewhat different results, though not differing by more than 30% between the greatest and smallest value, and it was found that, approximately,

$$c_1 = 0{\cdot}20\,.$$

In the system of double integration it is difficult to say anything about the numerical value of the Fourier components or the correlation function of the difference signal, because of the combination of a high degree of non-linearity with the response function of a rather complicated network. So the best we can do here is to use the numerical method by constructing the difference signal $\varepsilon$ for a given input signal over a longer period and considering the result of applying this signal to a low-pass filter. As the approximation curve $e_1(t)$ is the result of a number of transients of known form produced by the foregoing pulses, the polarity of

the following pulse, and thus the signal $\varepsilon$, can be obtained by an iterative process, once the signal $e_0(t)$ has been given. By constructing the difference signal in this way over a period of 80 pulse intervals it was found from the r.m.s. value at the output of a mathematical low-pass filter that the constant in equation (2) should have, approximately, the value

$$c_2 = 0{\cdot}026\,.$$

## 8. Experimental results

We must now consider how to measure the signal-to-noise ratio in practice. The maximum voltage $S$ is easily measured at the receiving end, but measurement of the quantizing noise requires removal of the transmitted information signal. A convenient way to do this is by using at the output a band-pass filter covering a large proportion of the audio-frequency band, as drawn in fig. 18. If we now apply to the system a signal I of a somewhat lower frequency that the cut-off frequency $\omega_c$ (but so that its second harmonic falls between $\omega_c$ and $\omega_0$) we can measure at once the amount of quantizing noise within the pass-band $\omega_c - \omega_0$..

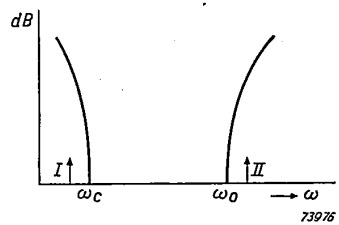



Fig. 18. The attenuation curve of a band-pass filter used for measuring the quantizing noise. I and II are the frequencies of the applied information signals.

In the same way it is possible to apply a signal II of a somewhat higher frequency than $\omega_0$ to the input at the sending end and again measure the quantizing noise in the band $\omega_c$—$\omega_0$. A value for the quantizing noise in the full frequency band up to $\omega_0$ can then be deduced by multiplying the observed noise power by a factor $\omega_0/(\omega_0-\omega_c)$, being based upon the fact that, as a rule, the noise will have a uniform frequency spectrum for those frequencies which are small in relation to the pulse frequency. As for the determination of the maximum amplitude of a signal, it is easiest to apply first an input signal of such an amplitude that the system is fully overloaded, resulting in a triangular output signal at the receiving end, and then diminish this measured voltage by about 3 dB, since at this level a sine wave can just be transmitted with negligible distortion.

The results of measurements of $S/N$ for different pulse frequencies are shown in fig. 19, for a reference signal of 800 c/s and a frequency band from 200 to 3800 c/s. (The filter used for measuring the quantizing noise had a pass-band from 1250 to 3800 c/s so that the measured noise power had to be multiplied by the factor 1·41). The lower curves correspond to single integration, and the higher to double integration. The dotted curves 1 and 2 give the theoretical values according to eq. (1) and (2), and the curves I and II in each case represent the values measured according to the method I and II respectively.

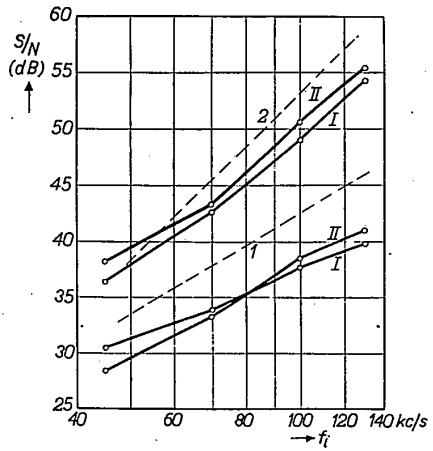



Fig. 19. Measurements of the signal to noise ratio $S/N$, for a sine wave of 800 c/s and for an audio frequency band from 200-3800 c/s, as a function of the frequency $f_i$. The lower curves correspond to single integration and the higher to double integration. The dotted lines 1 and 2 represent the numerical values according to the formulae (1) and (2). The curves I and II correspond to the measurements carried out with an input signal of low- and high frequency respectively. The improvement in signal-to-noise ratio is proportional to the 3/2 and 5/2 power of the pulse frequency, for the system of single and double integration respectively.

From the slope of these curves it can be seen that the improvement in signal-to-noise ratio, at least for the higher pulse frequencies, is in fair agreement with the 5/2 and 3/2 power of the pulse frequency. But the measured values are about 3 and 5 dB below the theoretical values for the system of double and single integration respectively — which, perhaps, may be due to some imperfections in the experimental apparatus.

In measuring these values for the system of single integration a small signal of low frequency was added, as otherwise it was difficult to determine the mean value of the quantizing noise, which varies in a haphazard

manner depending on the amount of periodicity in the quantized sine wave. By adding a low-frequency component the quantising effect was made more irregular, which had a smoothing effect on the deviations of the quantizing noise.

Apart from the lower amount of quantizing noise in the system of double integration, the noise in it has a different character from that with single integration, especially in the reproduction of low-level signals. The noise produced with single integration has more or less the character of a line spectrum with many sharp interference tones, while that in the double system is more evenly smoothed in a continuous spectrum. This can best be observed by gradually changing the frequency of a small input signal. As the frequency is steadily increased, the frequencies of the interference tones in the first system are moving upward and downward in a rapid way, whereas in the latter they are far less pronounced against a background of white noise. This is connected with the lowering of the threshold effect acquired by using smaller values at the beginning of the response function of fig. 14.

It is seen from fig. 19 that an improvement of about 10 dB can be obtained by using the method of double integration, at a pulse frequency of 100 kc/s. As the receiver is the same in both the single and double integration systems this means that the reduction of the noise power by a factor 10 has been achieved by a better grouping of the pulses used for transmission of the information. The manner in which this is possible, notwithstanding the same height of the steps in the integrating network at the receiving end, is illustrated in fig. 20. The information signal is represented as a continuous curve.

The method of single integration gives the approximation *1*, with a low-frequency component as indicated by the dotted curve. With double integration the approximation *2* is obtained, which, in spite of the fact that it is oscillating to a higher degree around the information curve, has a low-frequency component that will more closely resemble the original than the dotted curve in the other figure. In the first case the pulses only try to make the step function agree with the instantaneous level of the information function; in the latter, however, they also try to balance the total amount of overshoot on both sides of the curve. This is made possible by choosing a special shape of the frequency characteristic in the feedback network for those frequencies between the highest frequency to be transmitted and the pulse frequency. This feedback network, which controls the grouping of the pulses, has two functions to perform in this respect. First it should store the necessary information about the errors that have been made in the past, due to the process of quantizing the information signal into a step curve of discrete levels

only, so that the feedback network has to serve as a *memory*. And secondly, it has to take into account the behaviour of the resulting low-frequency component in the near future, in order to determine how these errors should be compensated by choice of the next pulses, thus making use of *prediction* over a short time-interval.

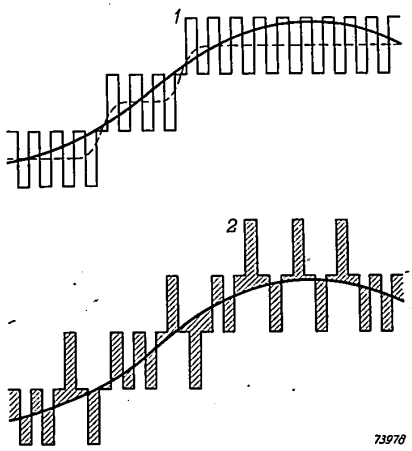



Fig. 20. These figures illustrate schematically the improvement in the accuracy of the approximation signal that is gained by a better grouping of pulses in the transmission path. Although the same amplitude levels are used in obtaining the approximations *1* and *2*, the low frequency component of the latter has a smaller deviation from the original signal than the first. (Because of its small deviation, the low-frequency component of the approximation *2* has not been drawn.) The approximation signal *2* is obtained in the case of double integration by using such a feedback network at the sending end that, to some extent, the total amount of overshoot on both sides of the curve is balanced.

In this manner, according to fig. 15, it is possible to arrive at an approximation signal which, after the smoothing process in the low-pass filter, differs from the information signal required by only a small percentage of one step. As this is essentially the result of the compensating effect of a number of positive and negative pulses and the use of a low-pass filter, some time will elapse before the required signal can be observed at the output. The whole process of quantization and reconstruction of the transmitted information is thus inherent in a certain time-delay.

## 9. Some general remarks

With a pulse frequency of 100 kc/s a very good reproduction of speech is possible. In this case a bandwidth of 50 kc/s per channel is necessary

in the transmission path, which is of the same order as usually applied in pulse-position modulation. As the interfering signals in the path of transmission are now, however, allowed to rise to about one half of the amplitude of the pulses, this system can be used under rather unfavourable conditions.*)

Compared with pulse-code modulation we arrive at the quality of an 8-unit code. For, taking 800 c/s as reference frequency for the reasons given previously, a signal-to-noise ratio of 50 dB is observed, while according to a formula given by Bennett [6]) the signal-to-noise ratio in dB for a full-load test tone, in using a code of $n$ digits, is

$$D = 6\,n + 3\,,$$

which gives $n = 8$ for $D = 50$ dB.

In this case the bandwidth needed for transmission with deltamodulation is about 50% greater than by using pulse-code modulation. For, taking a sampling frequency of 8 kc/s as is usual for the transmission of speech, the 8-unit code would require a pulse frequency of only 64 kc/s. The somewhat higher bandwidth needed for deltamodulation does not give much difficulty, however, if the system is applied to a short-wave radio link. A video bandwidth of 5 Mc/s in this case will already allow the transmission of 100 speech-channels.

It is comparatively easy to improve the quality of transmission by applying a higher pulse frequency. (One only has to make a corresponding change of the time constant $\tau$). For experimental purposes a system was set up for high-quality musical reproduction. Here the audio-bandwidth was extended to 12 kc/s and the signal-to-noise ratio (for a signal of 800 c/s again) raised to 67 dB, by using a pulse frequency of 500 kc/s. As a rule the reproduction of music was very good, but sometimes a difficulty arose when the frequency spectrum of the music contained much energy at higher frequencies (especially if a piano was played forte at the higher tones). In this case the level had to be adjusted to achieve a compromise between the quantizing noise and the effect of overmodulation, which means that the use of a single integrating network at the receiving end is not well-adapted to music of this kind.

In many respects the apparatus used in deltamodulation need not meet the high requirements often necessary in using a binary code with a high number of digits. Primarily this is due to the fact that the coding device here consists of an electrical network which is already linear in itself. If the output voltage $e_1$ of this network, is added to the input

*) The requirements for the signal-to-noise ratio in the transmission path are the same as for pulse-code modulation, if the comparison is made for the same pulse frequency.

voltage $e_0$ (which again may be done by means of a linear network) we are only interested in the *polarity* of the resulting voltage, and so may use an amplifier with a non-linear characteristic for producing the voltage which controls the pulse modulator. The only linear element in the quantizing apparatus being a linear network, it is even possible that in the future the whole system can be constructed in a rather inexpensive way by making use of transistors.

**Acknowledgements**

The author wishes to express his thanks to J. F. Schouten for his stimulating interest which has led to the investigations in this field. Special acknowledgement should be made of J. A. Greefkes whose help in the design and construction of the experimental apparatus was invaluable. Finally, our thanks are expressed to Prof. Willis Jackson from Imperial College, London, for reading the manuscript and for many valuable comments on the English text.

## APPENDIX

**Comparison of two systems which are limited in amplitude and in slope of the input signal respectively**

Figure 21 represents a transmission system in which the amplitude is limited at the voltage $N_1$. Fig. 22*a* represents a system where the slope of the input signal is limited at a certain value $S_1$. For the behaviour of this system we can consider the system of fig. 22*b* where the input signal is first applied to a differentiating network $\varphi(\omega)$ and then to a transmission system that is limited in amplitude, say at the voltage $N_2$.

It will be supposed that the input signal has a continuous spectrum

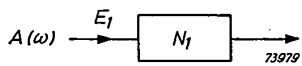



Fig. 21. A transmission system in which the amplitude is limited at the voltage $N_1$.

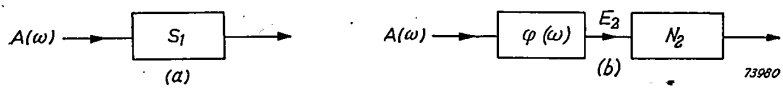



Fig. 22. In (*a*) the slope of the input signal is limited at the value $S_1$. In (*b*) the input signal is first applied to a differentiating network and then to a system which limits the amplitude at the value $N_2$. For a definite relation between $S_1$ and $N_2$ both signals applied to the systems (*a*) and (*b*) will be limited in slope at always the same time.

$A(\omega)$ in which phase relations between the different frequencies may be considered as being at random. (A signal of this type can be generated by a source of white noise that is followed by a network with the frequency characteristic $A(\omega)$. In this case the amplitude distribution is a Gaussian one and this property is maintained after passing this signal through an arbitrary linear network).

If $E_1$ is the r.m.s. voltage at the input of the system that is limited at the voltage $N_1$ in fig. 21 and, equally, $E_2$ the r.m.s. voltage in the system limited at the voltage $N_2$ in fig. 22, then both systems will give rise to distortion during an equal number of time intervals if

$$\frac{E_1}{N_1} = \frac{E_2}{N_2}.$$

If $\omega_0$ is the highest frequency of the input signal we have

$$E_1^2 = \int_{-\omega_0}^{\omega_0} A^2(\omega)\ \mathrm{d}\omega = 2\int_0^{\omega_0} A^2(\omega)\ \mathrm{d}\omega;$$

and if $\varphi(\omega)$ is the frequency characteristic of the differentiating network, the spectrum of the signal at the input of $N_2$ is $A(\omega)\ \varphi\ (\omega)$, so that

$$E_2^2 = 2\int_0^{\omega_0} |A(\omega)\ \varphi(\omega)|^2\ \mathrm{d}\omega.$$

Now consider the case of applying a sine wave of frequency $\overline{\omega}$ and of voltage $U$ such that limitation just occurs in both systems of fig. 21 and 22. We will define $U_1$ and $U_2$ as the r.m.s. voltages at the input of the limiters $N_1$ and $N_2$ and thus have $U_1 = U$, $U_2 = U|\varphi(\omega)|$.

Limitation of signals in both systems occurs at the same time when

$$\frac{U_1}{N_1} = \frac{U_2}{N_2}.$$

From these equations, together with the relation $\varphi(\omega) = kj\omega$ for a differentiating network, we find for the reference frequency

$$\overline{\omega}^2 = \frac{\int_0^{\omega_0} A^2(\omega)\omega^2\ \mathrm{d}\omega}{\int_0^{\omega_0} A^2(\omega)\ \mathrm{d}\omega}.$$

Thus, if both systems can transmit a sine wave of this frequency with

the same voltage, they can also transmit an input signal of the defined frequency spectrum with the same voltage, and if overloading occurs, it will take place for an equal number of time intervals in both systems.

*Eindhoven, November 1952*

## REFERENCES


1) French patent specification Nr. 987.238 dd. 22 May 1948.
2) E. G. Beard, Some notes on time division telephony and methods of pulse modulation. Proc. J. Instn. Radio Engrs. Aust. **11**, 14-16, Jan. 1950; J. Brit. Instn. Radio Engrs. **10**, 242-243, July 1950.
3) L. J. Libois, Un nouveau procédé de modulation codée: la modulation en $\Delta$. Echo des Recherches, avril 1951. Onde élect. **32**, nr. 298, 26-31, Jan. 1952.
4) J. F. Schouten, F. de Jager and J. A. Greefkes, Deltamodulation, a new modulation system for telecommunication. Philips tech. Tijdschr. **13**, 249-258, Sept. 1951 (in Dutch). Philips tech. Rev. **13**, 237-245, March 1952.
5) N. R. French and J. C. Steinberg, Factors governing the intelligibility of speech sounds. J. acoust. Soc. Amer. **19**, 90-119, Jan. 1947.
6) W. R. Bennett, Spectra of quantized signals. Bell. Syst. tech. J. **27**, 446-471, 1948.

# MEASUREMENT OF GROUP-DELAY TIME IN NETWORKS

by A. van WEEL 621.317.342:621.392.5

**Summary**

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.

**Resumé**

Un principe bien connu de mesure de l'écart de temps de propagation de groupe est basé sur la mesure des rapports des phases d'une modulation à basse fréquence d'une onde porteuse de haute fréquence, avant et après le passage par une réseau inconnu. L'article décrit un appareil très sensible de mesure des phases dans lequel l'angle de déphasage à mesure est introduit dans un circuit oscillant, influençant ainsi la fréquence d'oscillation. Dans certaines conditions les variations de phases sont proportionnelles aux variations de fréquences qui en résultent. Des variations de phase de 0,01 degré sont mesurées sans difficulté, et permettent de mesurer des variations du temps de propagation de groupe de $10^{-9}$ sec, en employant une fréquence de modulation de 30 kc/s.

**Zusammenfassung**

Eine bekannte Methode für die Messung von Gruppen-Laufzeiten besteht im Messen der Phasenbeziehungen an einem Nf-Modulationssignal, das einem Hochfrequenzträger überlagert ist, vor und nach dem Durchgang des Signals durch ein unbekanntes Netzwerk. In dem Aufsatz wird eine äusserst empfindliche Phasen-Messeinrichtung beschrieben, bei der der zu messende Phasenwinkel in einen Schwingungskreis introduziert und so die Frequenz der Schwingung beeinflusst wird. Unter bestimmten Verhältnissen sind die Phasen-schwankungen proportional den resultierenden Schwankungen der Frequenz. Phasenschwankungen von 0,01 Grad sind ohne Schwierigkeiten messbar; mithin ist es möglich, mit einem Modulationsfrequenz vor 30 kHz Veränderungen von $10^{-9}$ sec in der Gruppen-Laufzeit zu messen.

## 1. Introduction

The undistorted transmission of information over a communication channel is only possible if this channel is free from non-linear distortion, linear distortion, and phase distortion. Only with audio signals is the last kind of distortion relatively unimportant, at least as measured in the ultimately recovered audio-frequency band. With certain modulation systems

(frequency modulation, pulse modulation, etc.) the phase characteristics of the high-frequency parts of the communication channel are of great importance even for the transmission of audio signals, as a non-linear phase characteristic introduces non-linear distortion in the modulation.

Telegraphy and television transmission are very sensitive to phase distortion, and it is well known that for high-quality reproduction of the original signal, severe demands should be put on the linearity of the phase-frequency characteristic of the transmission channel. When relaying a signal over a large number of successive repeater stations, the distortion in the successive repeaters add in the ultimate signal, and therefore tolerances in the separate repeaters should be kept within narrow limits.

For this reason it is important to have measuring devices enabling the determination of the phase-frequency characteristics of networks with a high degree of accuracy. As the deviation of linearity of the phase-frequency characteristic is the most important property of this characteristic, the phase delay time ($\varphi/\omega$) or the group delay time ($\partial\varphi/\partial\omega$), measured as a function of the frequency, gives a good indication of the amount of deviation from the ideal linear characteristic.

With high- or intermediate-frequency circuits only the group delay time is a well-defined property of the network itself, because with high-frequency networks a phase delay time can only be defined for a certain modulating signal and in reference to a certain carrier frequency. The phase properties of such circuits are therefore best defined by a characteristic of the group delay time versus frequency.

Video- or audio-frequency circuits on the other hand could better be characterized by the phase delay as a function of frequency. However, the group delay time is in many cases easier to measure; from a group delay characteristic, the phase delay can in a simple manner be calculated by graphical integration.

The accurate measurement of group delay times has been the subject of a number of recent investigations [1], [2], [3], [4]. This article will describe a principle for measuring group delay times with a high degree of sensitivity with a very simple apparatus.

## 2. Principle of group-delay measurement

All group-delay measurements are based on the measurement of the phase variation impressed on a low-frequency modulation of a high-frequency carrier when passing the unknown network [5]. Provided the two side bands have the same but opposite phase differences with respect to the carrier frequency, the phase variation of the modulation equals the phase difference between each side band and the carrier frequency.

For a sufficiently low modulating frequency $(p/2\pi)$ this phase difference follows from

$$\varphi = \frac{\partial \varphi}{\partial \omega} p = \tau p,$$

$\tau$ being the group delay time of the network.

As only *variations* of the group delay $\tau$ are of interest, the measurement is usually confined to the determination of the phase *variation* of the detected $p$-signal after passing the unknown network as modulation on a carrier frequency, when this carrier frequency is varied through the frequency band in question.

The tolerances in delay time in television repeaters are about $10^{-9}$ sec [1]). With a modulating frequency of, for instance, 1 mc/s, the phase-measuring device should be sensitive to phase variations of $2\pi.10^6.10^{-9} = 6.10^{-3}$ radian $\approx$ 0·4 degree. However, using an modulating frequency of 1 Mc/s means averaging the actual group delay over a frequency band of 2 Mc/s; so for accurate measurements a much lower modulating frequency should be used. The inherent difficulty with smaller modulation frequencies is of course the necessity of a more sensitive phase-measuring device. The main problem in devising an accurate group-delay-time-measuring system therefore is the realization of a very sensitive phase meter.

In the group delay meter to be described, a modulating frequency of 30 kc/s is employed, from which follows a phase-meter sensitivity of 0·0002 radian $\approx$ 0·01 degree.

### 3. Phase-measuring device

It is well known that in a stable oscillating system the total phase variation when going round the oscillating loop, should equal 0 or $n \times 2\pi$ radians. Any phase variation occurring in the loop is promptly counteracted by a variation of the oscillating frequency causing an equal phase variation of opposite sign. This effect can be used to convert a given phase variation into a frequency variation, and this frequency variation can easily be measured with a high degree of accuracy.

The block diagram of fig. 1 shows how this principle is used to measure group delay times of a network.

The oscillating loop consists of a selective amplifier, tuned to a low frequency $(p)$, the output of which modulates the high frequency $(\omega)$ signal of a standard-signal generator. The modulated high-frequency signal is subsequently passed on through the unknown network and detected. The so recovered $p$-signal is fed to the input of the $p$-amplifier, thus closing the loop for the $p$-frequency signal.

By suitable dimensioning amplification and phase relations the system

IPR2025-00217 EXHIBIT 2010 Page 474 of 485

oscillates in the $p$-frequency. The exact value of this frequency depends on the phase properties of both the $p$-amplifier and the network.

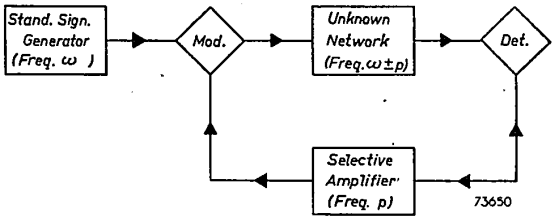



Fig. 1. Block diagram of the delay-time meter.

Now, suppose the standard-signal generator to be tuned to a new frequency $\omega'$. At this frequency $\omega'$ the group delay $\tau'$ of the network has changed from the value $\tau$ at the frequency $\omega$ to $\tau' = \tau + \Delta\tau$. Therefore, the phase of the modulating $p$-signal when passing the network changes with an amount $\Delta\varphi = p\Delta\tau$. This phase variation is counteracted in the oscillating circuit by a frequency variation $\Delta p$ such that

$$\Delta\varphi + \left(\frac{\partial\varphi}{\partial p}\right)\Delta p = p\Delta\tau + t'\Delta p = 0.$$

The group delay time $t'$ in this equation is the group delay time of the complete $p$-loop consisting of the $p$-amplifier and the network. So $t' = t + \tau$, $t$ being the delay time of the $p$-amplifier.

By choosing $t$ large with respect to $\tau$ a very simple and linear relation between the delay variation to be measured and the resulting frequency variation is realized:

$$\frac{\Delta\tau}{t} = -\frac{\Delta p}{p}.$$

In this equation the value of $t$ and $p$ are system constants; therefore the scale of the frequency-measuring device, used to measure the frequency variation $\Delta p$, can directely read in milli-micro seconds. The sensitivity of this system can be estimated from the system values as used in delay meters which were realized following this principle: $t = 70$ μsec; $p = 2\pi.30000 = 2.10^5$ rad/sec. Thus to a delay-time variation of $10^{-9}$ second, a frequency variation of 0·5 c/s corresponds.

**4. Choice of system constants.**

The choice of the system constants $t$ and $p$ depends on the sensitivity required and on the maximum delay variation to be measured. For a selective

circuit consisting of an inductance $L$, a condenser $C$, and a resistance $R$ in parallel, the group delay time in the neighbourhood of resonance is given by

$$t = \frac{2RC}{1 + (\Delta p\, 2RC)^2} = \frac{t_0}{1 + (\Delta p t_0)^2}; \qquad t_0 = 2RC.$$

Thus the linear relation between $\Delta\tau$ and $\Delta p$ holds only for limited values of $\Delta p$. Selective networks with a more linear phase-frequency characteristic than a simple parallel circuit can be realized. But even with linear phase networks the delay time $t = \partial\varphi/\partial\omega$ depends in most cases on the value of $\Delta p t_0$; so for any special selective network, a certain maximum tolerable value of $\Delta p t_0$ follows from the network properties.

As $\Delta p t_0$ equals $p\Delta\tau_{max}$ (not considering the sign), this same condition determines the maximum value of $p\Delta\tau_{max}$; thus the maximum value of the frequency $p$ follows from the maximum delay time variation to be measured.

The delay time $t_0$ should be chosen large with respect to the delay time of the networks to be measured. On the other hand, the smaller the value of $t_0$, the larger the frequency variation $\Delta p$ and therefore the more sensitive the apparatus. Values of $f_p = 30$ kc/s and $t_0 \approx 70$ $\mu$ sec proved useful for measuring delay time variations from $1.10^{-9}$ to $1000.10^{-9}$ sec. The frequency variations in this system were measured in a conventional frequency discriminator circuit.

Thus the complete delay-measuring device consists in essence of a tuned amplifier with an amplitude-limiting device (as in any normal oscillator), a modulator stage, and a frequency discriminator, the output of which is measured by a valve voltmeter. A detector is usually included in the unknown network and is therefore omitted in the measuring device. The scale of the valve voltmeter reads in milli-micro seconds and has several ranges, for instance 0 to 30, 0 to 100, 0 to 300, and 0 to 1000 m$\mu$sec.

### 5. Sensitivity and accuracy

The sensitivity of this system for measuring delay times can, if necessary, easily and substantially be improved over the value of $10^{-9}$ sec as realised in the apparatus as constructed up to now. By choosing a smaller value of the delay time $t$, the sensitivity increases accordingly.

However, difficulties arise in maintaining sufficient accuracy, because the oscillating loop reacts on any phase variation in the same way and does not discriminate between "wanted" phase variations, introduced in the unknown network, and any "unwanted" phase variation occurring somewhere in the circuit as a consequence of a disturbance of some kind. Therefore the more sensitive the device should be, the more stable the oscillating circuit should be constructed, whereas this stability cannot depend on highly selective circuits as in normal stable oscillators.

For applications where long delay times have to be determined and thus large phase variations should be measured, a linear relationship between these phase variations and the corresponding frequency variations can be maintained by choosing a selective network in the $p$-amplifier with a sufficiently linear phase-frequency characteristic. The larger the phase variation to be measured, the more tuned circuits will be necessary in this linear phase network.

The system as described only measures the delay time *variations* in networks, the absolute value of the delay time being unimportant in most cases. Absolute values of delay time can be measured by closing the oscillating loop successively over the unknown network and a known network and measuring the delay-time variation. The known network can be replaced by a direct connection.

## 6. Video-frequency measurements

The low value of the modulating frequency (30 kc/s) makes it possible to use the same device for measuring delay-time characteristics of video-frequency circuits. However, a substantial difficulty is encountered as a consequence of the fact that any output of the modulator of the $p$-frequency itself is passed on by the video network to be measured without any attenuation. As a consequence the $p$-frequency component at the output of the detector consists of two components: the wanted component which passed the network as modulation on a carrier frequency, and an unwanted component which passed the network as $p$-frequency. As the components may show a phase difference, an error in the measurements is likely to be introduced in this way,

To estimate the value of this error let us consider the phase-frequency characteristic of the video network (fig. 2). The phase variation of wanted and unwanted components are respectively

$$\varphi' = \left(\frac{\partial\varphi}{\partial\omega}\right)_{\omega} p = \tau_{\omega} p$$

and

$$\varphi'' = \left(\frac{\partial\varphi}{\partial\omega}\right)_{0} p = \tau_0 p\,.$$

Thus the phase difference between $\varphi'$ and $\varphi''$ follows from

$$\varphi' - \varphi'' = (\tau_{\omega} - \tau_0) p\,.$$

The error $\varepsilon$ in the phase of the resulting voltage $V$ follows from fig. 3 and is given by

$$\varepsilon \approx (\varphi' - \varphi'')\,\frac{V''}{V'},$$

and the error in the delay-time measurements by

$$\delta\tau = \frac{\varepsilon}{p} = \frac{\varphi' - \varphi''}{p} \frac{V''}{V'} = (\tau' - \tau'') \frac{V''}{V'}.$$

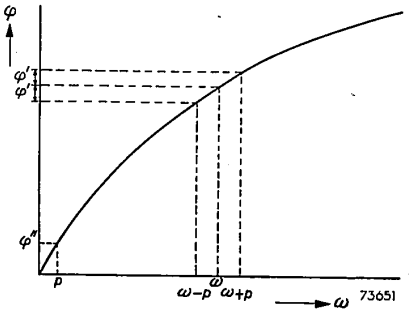



Fig. 2. Phase-frequency characteristic of a video-frequency network.

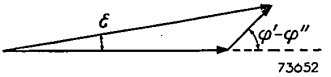



Fig. 3. Resultant error $\varepsilon$ originating from the difference $\varphi' - \varphi''$.

As only delay time *variations* are to be measured, the relative error $\delta\tau/(\tau' - \tau'')$ equals $V''/V'$, from which follows the maximum tolerable voltage of the "direct" $p$-frequency component. By using a balanced modulator and detector the level of the unwanted component $V''$ can easily be kept 40 dB below the wanted component $V'$, which is sufficient for an accuracy of 1%, provided the modulator and detector do not introduce appreciable delay variation between the $p$-frequency and the rest of the video-frequency band.

An apparatus as realized along these lines measures delay characteristics from 100 kc upwards. No maximum value for the carrier frequency follows from the principle underlying this system. However, practical reasons prescribe different modulator circuits for different frequency ranges.

*Eindhoven, June 1952*

REFERENCES

[1]) D. H. Ring, Bell Syst. tech. J. **27**, 247-264, 1948.
[2]) J. Selz and J. Iltis, Câbles et Transm. **5**, 337-346, 1951.
[3]) L. E. Hunt and W. J. Albersheim, Proc. Instn. Radio Engrs. **40**, 454-459, 1952.
[4]) J. R. Clayton, D. C. Espley, G. W. S. Griffith and J. M. C. Pinkham, Proc. Instn.elect. Engrs. **98**, 204-223, 1951.
[5]) H. Nyquist and S. Brand, Bell. Syst. tech. J. **9**, 522-549, 1930.

R 205 Philips Res. Rep. 7, 474-477, 1952

# ON PRE-EXPONENTIAL FACTORS IN FORMULAE FOR IONIC CONDUCTIVITY IN SOLIDS

by Y. HAVEN and J. H. van SANTEN

537.311.32:541.135.4

**Summary**

It is pointed out that, if in formulae for equilibrium constants and rate constants the energy is linearly dependent on temperature, $E = E_0 - \alpha T$, the pre-exponential factor is not increased by a factor $\exp(\alpha/k)$.

**Résumé**

On montre que si dans les formules pour les constantes d'équilibre et pour les vitesses de réaction l'énergie dépend linéairement de la température, $E = E_0 - \alpha T$, le facteur pré-exponentiel n'est pas augmenté d'un facteur $\exp(\alpha/k)$.

**Zusammenfassung**

Es wird darauf hingewiesen, daß, wenn die Energie in den Formeln für die Gleichgewichtskonstante und für die Reaktionsgeschwindigkeit linear von der Temperatur abhängt, $E = E_0 - \alpha T$, der Faktor vor $\exp(E/kT)$ nicht um einen Faktor $\exp(\alpha/k)$ erhöht wird.

The ionic conductivity in crystals can be explained in terms of lattice defects and their mobilities [1-3]). E.g. the concentration of lattice defects in alkali halides can be represented by

$$x_1 \cdot x_2 = A \exp(-E/kT) \tag{1}$$

and their mobilities by

$$\mu = B \exp(-U/kT), \tag{2}$$

where $A$, $B$, $E$ and $U$ are constants, $x_1$ and $x_2$ = fractions of unoccupied positive and negative lattice sites resp., $\mu$ = mobility of the ions, $E$ = energy of formation of a pair of lattice defects, $U$ = energy of activation for moving, $k$ = Boltzmann's constant, and $T$ = temperature in °K.

According to simple theory the constant $A$ in eq. (1) should be unity. The experimental values for $A$, however, are found to be larger by factors $10^3$-$10^7$. Likewise in eq. (2) the experimental value for the constant $B$ is greater than the theoretical one, though the discrepancy (about a factor 10) in general is less than in eq. (1); cf. [4]).

One of the attempts to reduce the gap between theoretical and experimental data is that of Jost [5]) and of Mott and Gurney [1]) who supposed that the energies $E$ and $U$ might depend linearly on temperature. Writing

$$E = E_0 - \alpha T, \tag{3}$$

where $\alpha$ is a positive constant, and introducing (3) in (1) one obtains

$$x_1 \cdot x_2 = A \exp(\alpha/k) \cdot \exp(-E_0/kT), \tag{4}$$

thus reducing the discrepancy by a factor $\exp(\alpha/k)$, whilst the experimental energy is no longer $E$ but $E_0$.

The idea of a temperature dependency of $E$ and $U$ is often found in the literature. But a closer inspection reveals that a temperature dependency of $E$ does not influence the pre-exponential factor in the way as concluded from eq. (4). It appears that the temperature dependency of $E$ involves automatically a temperature dependency of an entropy term $S$ that has an opposite effect.

Let us consider the equilibrium of vacant lattice sites in a monatomic crystal. In a restricted temperature region we can represent the concentration by the experimental formula

$$x = A_{\text{exp}} \exp(-E_{\text{exp}}/kT), \tag{5}$$

where $x$ is fraction of vacant lattice sites and $A_{\text{exp}}$ and $E_{\text{exp}}$ are constants. Following the argument of e.g. ref. [2]) we shall investigate the meaning of $A_{\text{exp}}$ and $E_{\text{exp}}$. Let $n$ be the number of vacant sites and $N$ the number of occupied sites; $x = n/(n + N)$. The equilibrium condition at constant volume then is given by

$$\partial F_2/\partial n = 0, \tag{6}$$

and at constant pressure by

$$\partial G_2/\partial n = 0, \tag{7}$$

where $F_2 = E_2 - TS_2$ = Helmholtz free energy, $E_2$ = energy content, $S_2$ = entropy, $G_2 = F_2 + pV_2$ = Gibbs free energy, $p$ = pressure, $V$ = volume. The index 2 refers to the state with lattice defects and index 1 will refer to the state without lattice defects. We shall consider the equilibrium at constant volume. We assume that the concentration-dependent parts in $F$ are given by

$$E_2 - E_1 = n(E_2' - E_1'), \tag{8}$$

$$S_2 - S_1 = n(S_2' - S_1') + k \ln\{(n + N)!/n!N!\}, \tag{9}$$

where $E'$ and $S'$ are standard values that are independent of concentration, but may depend on temperature. With these assumptions the equilibrium condition (6) yields

$$x = \exp\{-(F_2' - F_1')/kT\} = \exp\{-(E_2' - E_1')/kT\} \cdot \exp\{(S_2' - S_1')/k\}. \tag{10}$$

From (5) and (10) the experimental energy $E_{\text{exp}}$ is found to be

$$E_{\text{exp}} = kT^2(\partial \ln x/\partial T)_v = F_2' - F_1' - T(\partial (F_2' - F_1')/\partial T)_v \tag{11}$$

or since
$$(\partial F/\partial T)_v = -S, \tag{12}$$
$$E_{\text{exp}} = E_2' - E_1'. \tag{13}$$

On the other hand, the pre-exponential factor, using (5), (10) and (13), is found to be
$$A_{\text{exp}} = \exp\{(S_2' - S_1')/k\}. \tag{14}$$

Therefore, the procedure indicated by formulae (3) and (4) does not work; the pre-exponential factor always being given by eq. (14). It may even be shown that assumption (3) has a tendency to reduce the factor $A$, thus increasing the discrepancy between theory and experiment.
Denoting the specific heat at constant volume $(\partial E'/\partial T)_v$ by $c_v$ we find for the entropy
$$S' = S_0' + \int_1^T \frac{c_v}{T}\, dT, \tag{15}$$

where $S_0'$ = entropy at 1 °$K$.

From (3) it follows that
$$c_{v2} - c_{v1} = -a \tag{16}$$

and substituting (15) and (16) in (14) we find
$$A_{\text{exp}} = T^{-a/k} \exp\{(S_{02}' - S_{01}')/k\}.$$

Thus under the conditions of a linear dependency of $E$ upon $T$ the pre-exponential factor $A$ can be divided in a term $\exp\{S_{02}' - S_{01}')/k\}$, $(S_{02}' - S_{01}')$ being an entropy difference at 1 °K, and an additional factor $T^{-a/k}$. As $a > 0$, this factor is smaller than unity.

Let us consider as an example LiF. In LiF the concentration of vacant lattice sites at about 1000-1100 °K is given by [4])
$$x_1 . x_2 = 2{\cdot}5.10^5 \exp(-31200/T).$$

In order to give an estimation of the temperature dependency of $E$ we might assume that $E$ is proportional to $1/a$, $a$ being the lattice parameter. Putting the coefficient of linear expansion $4.10^{-5}$/°K [6]) we can write
$$E = 64300 - 2{\cdot}5\ T \text{ cal/mole}.$$

This gives for the extra factor in $A$
$$T^{-a/k} = T^{-1{\cdot}25} = 0{\cdot}0002 \text{ at } 1000\ ^\circ\text{K}.$$

It should be remembered that in the expression for conductivity the factor $A$ enters in the form $A^{1/2}$.

At constant pressure starting from (7) and using everywhere the enthalpy $H = E + pV$ instead of $E$ and in (15) $c_p$ instead of $c_v$, the argument runs completely parallel, yielding the same conclusions.

The conclusions pertain to any equilibrium constant and should hold for rate processes as well (e.g. ionic mobility in conduction or diffusion processes). According to the transition state theory of Eyring c.s. [7]) the rate constant can be analyzed into an equilibrium factor characterizing the equilibrium between the initial state and the activated complex (in our case the top of the potential hill), and a rate factor representing the thermal velocity of the particles at the potential hill. For the equilibrium factor, which contains the energy and entropy of activation, the conclusions derived above for equilibrium constants are not affected.

*Eindhoven, July 1952*

REFERENCES

1) N. F. Mott and R. W. Gurney, Electronic processes in ionic crystals, Clarendon Press, Oxford, 1940.
2) W. Jost, Diffusion in solids, liquids and gases, Academic Press Inc., New York, 1952.
3) Y. Haven, Rec. Trav. chim. Pays Bas, **69**, 1259-1274, 1950.
4) Y. Haven, Rec. Trav. chim. Pays Bas, **69**, 1471-1489, 1950.
5) W. Jost, Z. phys. Chem. A **169**, 129-134, 1934
6) A. Eucken and W. Dannöhl, Z. Elektrochem. **40**, 814-821, 1934.
7) S. Glasstone, K. J. Laidler and H. Eyring, The theory of rate processes, McGraw Hill, New York, 1941.

## NOTE TO B. D. H. TELLEGEN, A GENERAL NETWORK THEOREM, WITH APPLICATIONS, PHILIPS RES. REP. 7, 259, 1952

Mr K. Posthumus kindley drew my attention to the fact that the theorem was also derived and used in the article of K. Posthumus and Tj. Douma, "Some considerations on the frequency stability of ultra shortwave driving-oscillators", Philips Transmitting News **3**, nr. 4, December 1936.

B. D. H. TELLEGEN

# AUTHOR INDEX


| | Page | |
|---|---|---|
| Bakker, J. | 344, | Kinetics of the emulsion polymerization of styrene |
| Beurs, P. H. de, see Jonker J. L. H. | | |
| Boer, J. B. de, V. Onate and A. Oostrijck | 54, | Methods for measuring and calculating luminance of road surfaces |
| Boerdijk, A. H. | 303, | Close-packing of equal spheres |
| Botden, Th. P. J. | 197, | Transfer and transport of energy in luminescent solids |
| Bril, A., and H. A. Klasens | 401, | Phosphors under cathode-ray excitation |
| Bril, A., and H. A. Klasens | 421, | Phosphors for flying-spot cathode-ray tubes |
| Bril, A., see Kröger, F. A. | | |
| Bruining, H., see Schagen, P. | | |
| Diemer, G., and H. Dijkgraaf | 45, | Langmuir's $\xi$, $\eta$ tables |
| Diemer, G., and K. Rodenhuis | 36, | Optimum geometry of microwave amplifier valves |
| Diemer, G., see Knol, K. S. | | |
| Dijkgraaf, H., see Diemer, G. | | |
| Dikhoff, J. A. M., see Kröger, F. A. | | |
| Francken, J. C., see Schagen, P. | | |
| Haringx, J. A. | 112, | Instability of thin-walled cylinders subjected to internal pressure |
| Haringx, J. A. | 189, | Instability of bellows subjected to internal pressure |
| Haven, Y., and J. H. van Santen | 474, | Formulae for ionic conductivity in solids |
| Jager, F. de | 442, | Deltamodulation, a P.C.M. transmission with 1-unit code |

| | Page | |
|---|---|---|
| Jonker, J. L. H. | 1, | Theory of secondary electron emission. |
| Jonker, J. L. H., A. J. W. M. van Overbeek and P. H. de Beurs | 81, | Decade counter valve for high counting rates. |
| Klasens, H. A., see Bril, A. | | |
| Knol, K. S., and G. Diemer | 251, | H.F. diode admittance with retarding D.C. field |
| Kröger, F. A., A. Bril and J. A. M. Dikhoff | 241, | Luminescent screen for television tubes |
| Lindt, W. J. van de | 28, | Application of multi-hole coupling to waveguide attenuators |
| Onate, V., see Boer, J. B. de | | |
| Oostrijck, A., see Boer, J. B. de | | |
| Overbeek, A. J. W. M., see Jonker J. L. H. | | |
| Pannenborg, A. E. | 131, 169, 270, | On the scattering matrix of symmetrical waveguide junctions |
| Rodenhuis. K., see Diemer, G. | | |
| Santen, J. H. van, see Haven, Y. | | |
| Schagen, P., H. Bruining and J. C. Francken | 119, | A simple electrostatic electron-optical system with only one voltage |
| Stevels, J. M. | 161, | Power factor of glasses. |
| Tellegen, B. D. H. | 295, 477, | A general network theorem, with applications |
| Vermeulen, R. | 432, | Dimensional analysis, units, and rationalization |
| Volger, J. | 21, | Specific heat of barium titanate |
| Weel, A. van | 467, | Group-delay times in networks |
| Wise, M. E. | 321, | Dense random packing of unequal spheres |

# INDEX ACCORDING TO THE UNIVERSAL DECIMAL CLASSIFICATION SYSTEM


| UDC-number | Report |
|---|---|
| 513.468 | *R* 197, 198 |
| 535.241.44 | *R* 184 |
| .373.1 | *R* 193 |
| .373.3 | *R* 201 |
| .376 | *R* 200 |
| 536.631 | *R* 180 |
| 537.081.5 | *R* 202 |
| .311.32 | *R* 205 |
| .37 | *R* 192 |
| .533.331 | *R* 187 |
| .533.8 | *R* 179 |
| .543(083.5) | *R* 183 |
| 539.2 | *R* 192 |
| .371 | *R* 186, 191 |
| 541.127 | *R* 199 |
| .135.4 | *R* 205 |
| 542.952:623 | *R* 199 |
| 547.538.141 | *R* 199 |
| 548.0.536 | *R* 180 |
| 621.3.081.5 | *R* 202 |
| .315.612.6 | *R* 189 |
| .317.3 | *R* 188, 190, 196 |
| .317.342 | *R* 204 |
| .317.374 | *R* 189 |
| .318.572 | *R* 185 |
| .385.1.029.64 | *R* 182 |
| .385.2.011.21 | *R* 194 |
| .385.832 | *R* 185, 200, 201 |
| .386.832 | *R* 193 |
| .392.1 | *R* 195 |
| .392.26 | *R* 181, 188. 190, 196 |
| .392.5 | *R* 195, 204 |
| .396.619.16 | *R* 203 |
| 628.971 | *R* 184 |