COLD CATHODE GLOW DISCHARGE TUBES

WESTON

Cold Cathode Glow Discharge Tubes

Cold Cathode Glow Discharge Tubes

G. F. Weston, M.SC., A. INST. P.

London ILIFFE Books Ltd

ILIFFE Books Ltd
42, Russell Square
London, W.C.1

© G. F. WESTON 1968

First Published in 1968

Filmset by Photoprint Plates Ltd.,
Wickford
Essex

Printed in England by
J W Arrowsmith Ltd.,
Bristol

Contents

Preface 7
1 Townsend's Ionisation Coefficients 1
2 Breakdown at Low Pressures 41
3 The Glow Discharge 67
4 Cathodic Sputtering 115
5 The Temporal Growth and Decay of a Discharge 151
6 Stabilisers and Reference Tubes 191
7 Switching Diodes and Trigger Tubes 225
8 Stepping Tubes 281
9 Indicators and Display Tubes 315
Index 369

CHAPTER 1

Townsend's Ionisation Coefficients

1.1. Voltage-Current Characteristics of a Discharge

If two plane-parallel metal electrodes are considered distance, d, apart in a gas at reduced pressure, then providing no ions or electrons are present the gas will behave as an insulator. In practice, however, there is always some ionisation of the gas due to cosmic radiation, and if a small potential is applied between the electrodes some of these ions and electrons will be collected, resulting in a small current of the order of 10^{-15} A passing through the gas. As the potential is increased the percentage of ions and electrons collected increases until saturation is reached when all the electrons and ions formed in the gas are collected at the electrodes. This is shown as region A on the voltage-current characteristic plotted in Fig. 1.1. The saturation current, i_o, depends on the incident radiation and may well fluctuate with time. At higher potential gradients two effects come into play; the electrons gain sufficient energy to ionise the gas by electron collision and electrons are ejected from the cathode by either positive ions, metastable atoms,* or photons produced in the gas. As a result of these processes the current increases rapidly with increasing potential until breakdown is said to occur, and the discharge becomes self-sustaining, i.e. independent of i_o, region B, Fig. 1.1. The maximum potential reached, V_s, is known as the breakdown potential. Attempt to raise the potential above V_s results in space charge being set up by the increasing current, and an actual reduction in the potential drop across the tube until a minimum is reached, region C, Fig. 1.1. At the minimum, the potential remains substantially constant over a wide current range and the physical characteristics of a glow discharge are

*A metastable atom is an atom excited to an energy level from which it cannot spontaneously decay back to the ground state.

observed. The potential in this region, V_m, is known as the maintaining potential.

The fact that the maintaining potential of a discharge is lower than the potential required to initiate the discharge is fundamental to all glow discharge devices, and the variation of V_s and V_m with geometry, gas filling and electrode material is of prime importance.

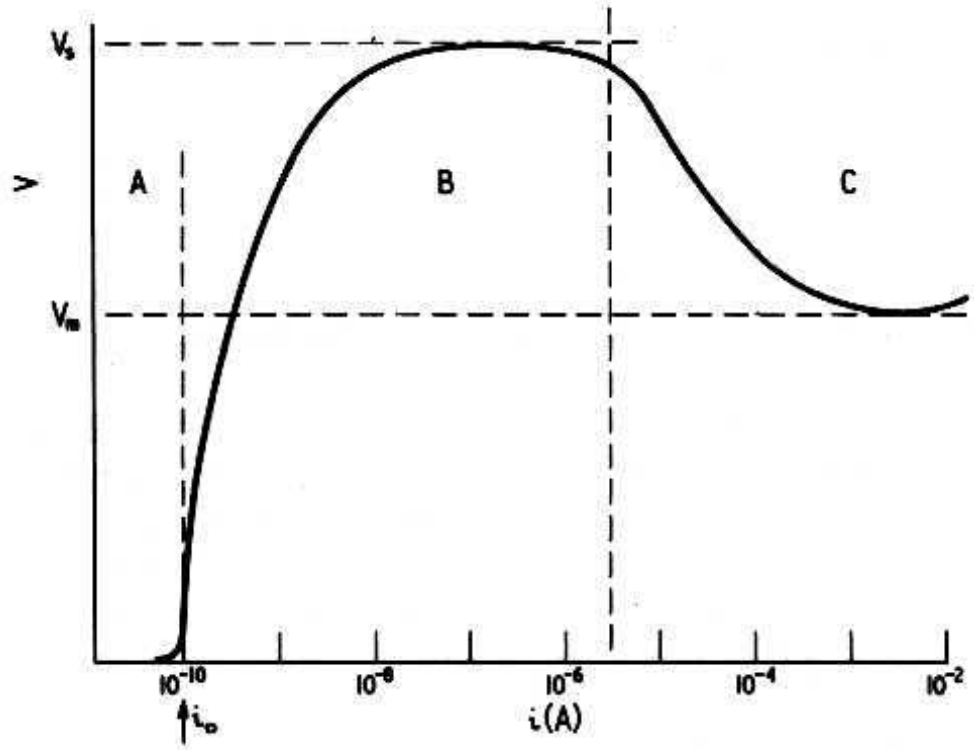


Fig. 1.1. V-i characteristic of a gas discharge diode

1.2. Townsend's First Ionisation Coefficient

The first detailed study of the *V-i* characteristics of a gas-filled diode was made by Townsend[1] in 1900. In his experiments the cathode was irradiated with ultra-violet light so that the saturation current i_o represented photoelectron emission from the cathode. He showed that for a given electric field, the initial increase of current from i_o (region B, Fig. 1.1) was exponentially dependent on d, the inter-electrode spacing. He introduced a coefficient α to define the number of electrons produced by a single electron travelling unit distance in the direction of the field, and derived the equation for the current as

$$i = i_o \exp \alpha d \tag{1.1}$$

This equation assumes that the number of free paths in the inter-electrode gap is sufficiently large and distributed to treat the

ionisation as occurring continuously throughout the gap. Because the electrons from the cathode require to travel a certain distance, d_o, to gain their equilibrium energy a correction factor must be introduced. Equation 1.1 then becomes:

$$i = i_o \exp \alpha(d - d_o) \tag{1.2}$$

This equation can be expressed in terms of the potential difference, V, across the inter-electrode gap if a coefficient η, introduced by some investigators,[2] is used in place of α. The coefficient η is defined as the number of electrons formed in a gas by a single electron dropping through a potential of 1 V. Thus $\eta = \alpha/E$ where E is the field in the gas, and Equation 1.2 becomes

$$i = i_o \exp \eta(V - V_o) \tag{1.3}$$

where V_o is the correction constant and will be of the order of the ionisation potential of the gas.

Since the early investigations of Townsend a considerable amount of work has been carried out on the determination of α or η, normally referred to as Townsend's first coefficient, in various gases under different conditions. The importance of such determinations in predicting the discharge characteristics will be apparent in later chapters. Experimental techniques and results of measurement covering early work have been fully discussed by Loeb,[3] and the physicist interested in the more fundamental work in this field is referred to this excellent book. Only a brief review of important features will be given here. Since the inert gases are employed in the majority of cold cathode tubes emphasis will be placed on values for these gases, in particular results obtained since 1955.

1.2.1. THE DEPENDENCE OF α ON THE FIELD AND PRESSURE

Empirically α was found to depend on both the field E and the pressure p. Since the ratio of field to pressure E/p dictates the energy reached by an electron between collisions it might be expected that α would depend on this ratio. In fact for most cases α/p can be expressed as a unique function of E/p, and therefore values of α/p_o are usually given in the literature rather than α, where p_o is the pressure corrected to 0°C. Townsend[1] first deduced a relation between α/p and E/p on a semi-empirical basis at a time when the mechanism of ionisation was not completely understood. However, the relation gave the general shape of the experimental curve, and has since been corroborated by von Engel and

Steenbeck[4] and Druyvesteyn and Penning[5] using more rigorous analysis. The relation is given by the equation

$$\frac{\alpha}{p} = \mathrm{A} \exp\left(-\frac{\mathrm{B}p}{E}\right) \tag{1.4}$$

where A and B are constants depending on the gas.

In deriving Equation 1.4 a number of simplifying assumptions are made which are known to be inaccurate, for example the probability of ionisation by an electron with energy above the ionisation energy is taken as 1. However by suitably choosing the constants the equation can be made to fit experimental data over a limited E/p range. Values of A and B given by von Engel and Steenbeck[4] and by Takeishi[6] together with the E/p range for which Equation 1.4 can be applied are listed in Table 1.1.

Table 1.1
Values of the constants in Equation 1.4
and the $\frac{E}{p}$ range for which they apply

Gas	A (cm^{-1} $torr^{-1}$)	B (V cm^{-1} $torr^{-1}$)	*E/p range* (V cm^{-1} $torr^{-1}$)
Air[4]	14·6	365	150 – 160
Argon[4]	13·6	235	100 – 600
Carbon dioxide[4]	20·0	466	500 – 1,000
Hydrogen[4]	5·0	130	150 – 400
Helium[6]	1·74	45	25 – 95
Neon[6]	4·8	116	100 – 400
Xenon[6]	18·6	234	100 – 400

Expressing Equation 1.4 in terms of η,

$$\eta = \mathrm{A}\frac{p}{E} \exp\left(-\mathrm{B}\frac{p}{E}\right) \tag{1.5}$$

which gives the curve shown in Fig. 1.2 when plotted on a log scale.

A development of the theory has been given by Emeleus, Lunt and Meek[7] which allows the insertion of the appropriate ionisation probability P_i, and energy distribution function F(W). If F(W)dW is the fraction of electrons per unit volume with energies between W and $W+\mathrm{d}W$ then

$$\frac{\alpha}{p} = \frac{1}{v_d} \int_{E_i}^{\infty} vP_i(W)\mathrm{F}(W)\mathrm{d}W \tag{1.6}$$

where v is the velocity of an electron of energy W, v_d the mean drift

velocity of the electrons at unit pressure, and E_i the ionisation energy of the molecule.

Using existing data on electron drift velocity, and ionisation probabilities according to Tate and Smith[8] they found reasonable agreement with experimental values of α/p for hydrogen, oxygen and nitrogen when a Maxwellian energy distribution was assumed. The values calculated for argon, however, were 3 – 4 orders too high, showing that the number of electrons with energy above the

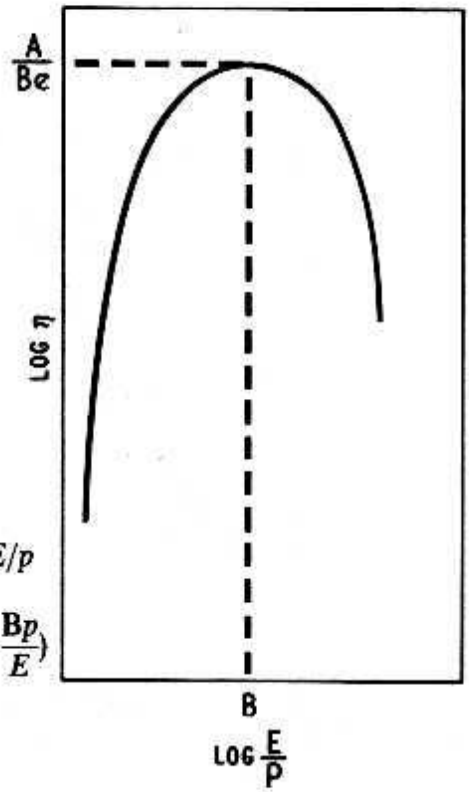


Fig. 1.2. Characteristic curve for log η against log E/p

The curve is drawn for the relation $\eta = \frac{Ap}{E} \exp\left(-\frac{Bp}{E}\right)$

ionisation energy must be considerably less than given by a Maxwellian distribution. In a subsequent paper by Deas and Emeleus,[9] calculations were made assuming a distribution in which inelastic collisions were taken into account. Although the values obtained were lower, they were still too high for atomic gases such as argon and neon.

Druyvesteyn[10] and later Druyvesteyn and Penning[5] have calculated η for neon and argon by an energy balance equation which they used in preference to an equation such as 1.6. Their results are shown graphically in Fig. 1.3. In the 'rigorous' calculation an energy distribution which takes into account elastic, excitation and ionising collisions is used. The simplified calculation assumes a linear dependence of the ionisation and excitation probability on electron energy, and an approximate distribution function. The simplified equation is of similar form to the original Townsend Equation 1.4.

$$\eta = \frac{a_i}{a_h V_i} \exp\left(-\frac{Bp}{E}\right)$$

where a_i and a_h are constants relating the field to the ionisation and excitation probabilities and V_i is the ionisation potential. Although Druyvesteyn and Penning claim good agreement, the curves are plotted on a logarithmic scale and in fact are out by a factor of 2 at some points.

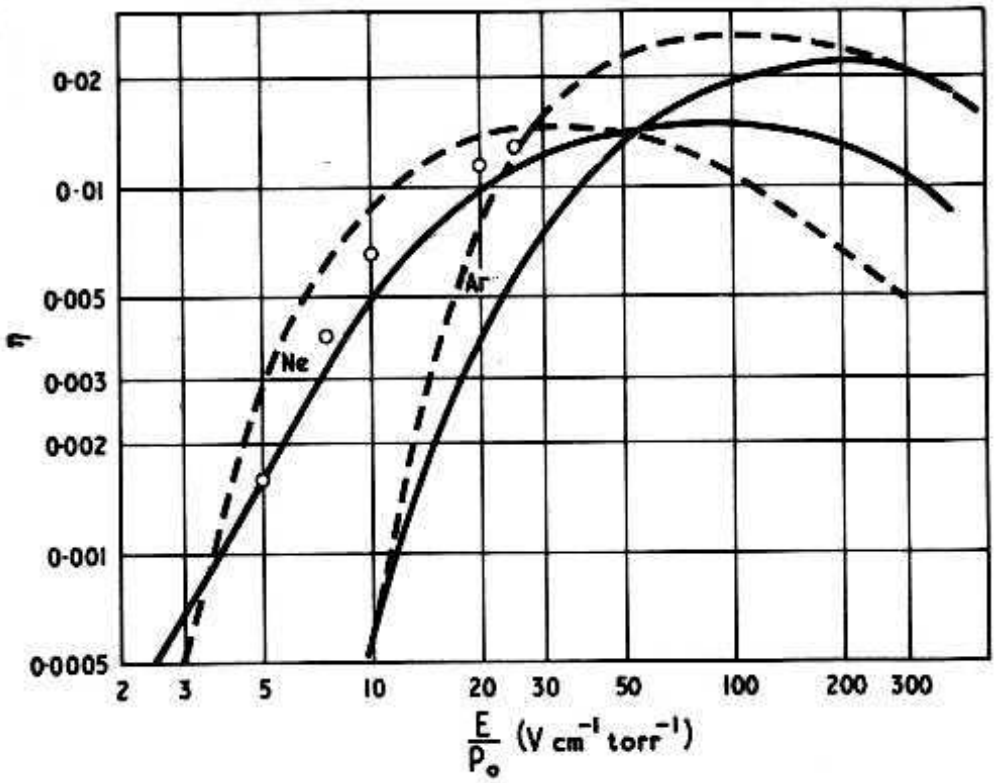


Fig. 1.3 η as a function of E/p_o in Ne *and* Ar *according to Druyvesteyn and Penning*[5]. *Experimental (——), simplified calculation (– – –), and more rigorous calculations for* Ne *(circles)*

A closer fit with experimental data for neon and argon has been obtained by Ward[11] over the range of E/p between 10 and 400 $\mathrm{Vcm^{-1}\,torr^{-1}}$ using the empirical formula:

$$\alpha/p = \mathrm{C}\exp(-\mathrm{D}p^{\frac{1}{2}}/E^{\frac{1}{2}}) \tag{1.7}$$

where C and D are constants depending on the gas.

Table 1.2
Values of C and D in Equation 1.7 for the inert gases according to Ward[11]

Gas	*Ionisation Potential* (V)	C ($cm^{-1}\ torr^{-1}$)	D ($[\mathrm{V}\ cm^{-1}\ torr^{-1}]^{\frac{1}{2}}$)
He	24·58	4·4	14·0
Ne	21·56	8·2	17·0
Ar	15·75	29·22	26·64
Kr	14·00	35·69	28·21
Xe	12·13	65·30	36·08

The agreement with the experimental values obtained by Kruithof can be seen in Fig. 1.4. C and D values for the inert gases are given in Table 1.2.

It would thus appear that for molecular gases over a limited E/p range log α/p is approximately proportional to E/p but that for atomic gases a half power law is more appropriate.

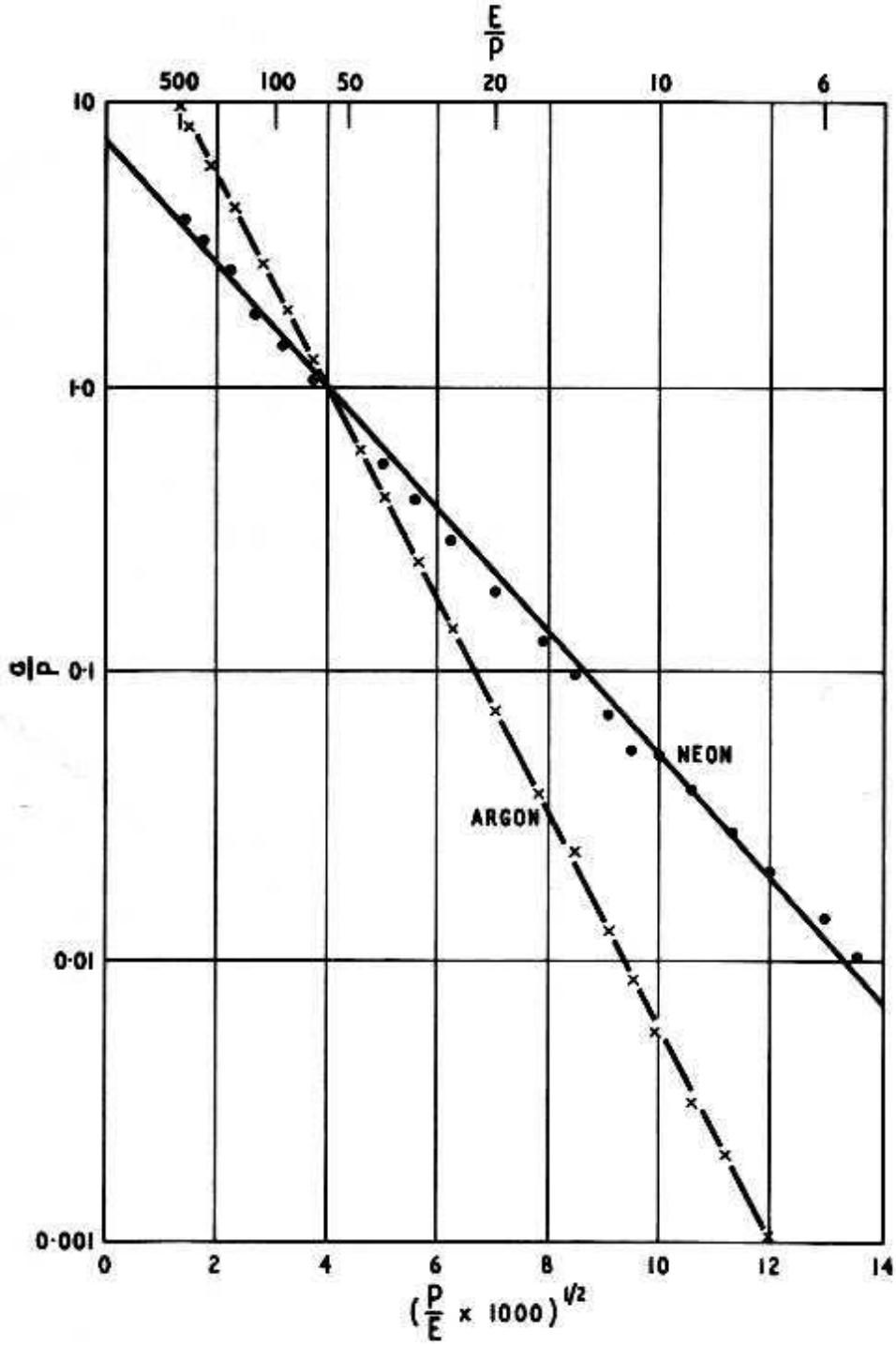


Fig. 1.4. Variation of Townsend's first coefficient α with field for Ne *and* Ar *according to Ward*[11]*. The values of α are taken from Kruithof*[25]

1.2.2. MEASURED VALUES OF THE FIRST TOWNSEND COEFFICIENT IN PURE GASES

Early determinations of the Townsend coefficient α, prior to 1926, were carried out in vacuum systems which would be considered poor compared with modern techniques. Consequently, although this early work established the basic theory of the non-selfsustaining gas discharge, little reliance can be placed on the values of α obtained. In fact it was not until 1931, when studies of very pure inert gases and mixtures were undertaken by Penning and his colleagues in Eindhoven[12] that the importance of small amounts of impurity was recognised (see Section 1.2.3). Therefore although care was taken to clean the gas before admitting it to the tube it is likely that the values obtained in the 1920's refer to gases contaminated with an equilibrium pressure of mercury.

The first detailed measurements on gases of fairly high purity were made by Ayres[13] in 1923, using silver electrodes. By avoiding grease taps he was able to obtain a considerable improvement on earlier vacuum systems. In spite of this, his measurements showed variations over a period of an hour. His measurements were made using hydrogen, nitrogen and argon, but the values of α/p he obtained are higher than those attained by later workers, and since no liquid air trap was used the gases were presumably contaminated with mercury vapour.

Table 1.3

Reported values of α/p for neon as a function of E/p

E/p (V $cm^{-1}torr^{-1}$)	*Townsend*[14] *and McCallum* α/p (cm^{-1} $torr^{-1}$)	*Kruithof*[24] *and Penning* α/p_o (cm^{-1} $torr^{-1}$)	*Chanin*[30] *and and Rork* α/p_o (cm^{-1} $torr^{-1}$)	*Fletcher*[33] *and Davies* α/p *taken from curve* (cm^{-1} $torr^{-1}$)
2	—	0·00065	0·0005	—
3	—	0·00192	0·00154	—
5	—	0·0080	0·00726	—
7	—	0·0204	0·0196	—
10	0·058	0·0501	0·0462	—
15	0·15	0·116	0·133	—
20	0·24	0·1924	0·198	—
50	0·8	0·715	0·680	0·5
70	—	1·037	1·03	—
100	1·57	1·493	1·53	1·0
150	—	2.110	2.16	1·4
200	2·53	2·596	2·62	1·7
300	—	3·22	3·60	2·4
400	—	3·56	4·12	2·9

In 1928 Townsend and McCallum[14] published their results on neon gas with nickel electrodes. Being aware of the need for purifying the gas thoroughly, they constructed their system in quartz which could be baked at high temperature, and cleaned the gas by passing a subsidiary discharge for a considerable period before taking readings. Their values of α/p, which do not appear to be corrected for temperature, for various E/p values are shown in Table 1.3. Later, in 1934,[15] the same authors presented their results on helium gas using the same apparatus. The results are given in Table 1.4.

Table 1.4

Reported values of α/p as a function E/p for helium

E/p (V cm^{-1} $torr^{-1}$)	†*Townsend[15] and McCallum (cm^{-1} $torr^{-1}$)	*Davies, Llewellyn Jones[34] and Morgan (cm^{-1} $torr^{-1}$)	Chanin[35] and Rork (cm^{-1} $torr^{-1}$)	*Fletcher[33] and Davies (cm^{-1} $torr^{-1}$)
3	—	—	0·0008	—
7	—	0·017	0·0161	—
10	—	0·049	0·0513	—
12	0·095	—	—	—
15	—	0·095	0·113	—
20	—	0·14	0·189	—
25	0·29	0·21	—	—
30	—	0·26	0·342	—
32·5	0·413	—	—	—
40	—	0·43	—	—
50	0·675	0·58	0·630	0·5
60	0·76	0·68	—	—
70	—	0·81	0·917	—
80	—	1·03	—	—
90	—	1·2	—	—
97	1·11	—	—	—
100	—	1·3	1·31	1·06
150	—	—	1·83	—
200	—	—	2·08	2·04
300	—	—	2·43	2·58

†No correction appears to be made for temperature.
*Taken from curve.

These experiments overlapped the studies carried out in Aachen by Paavola 1929[16] and Masch 1932[17] and also the work of Sanders and Posin in 1932,[18] 1933,[19] and 1936.[20] Their results will not be given here, since their experiments were carried out in nitrogen, air and oxygen, which are not of direct interest in glow discharge

devices. Also as it was demonstrated later by Bowles[21] the gases used were probably contaminated with mercury vapour. However, their contribution to gas discharges was important in establishing the general shape of the $\alpha/p - E/p$ curve and demonstrating that no simple single function, such as expressed in Equation 1.4, could represent the whole curve.

Of more interest were the experiments pursued in Eindhoven by Penning and his co-workers with very pure inert gases during the period 1934–1940. The work stemmed out of breakdown measurements where they established the effect of small admixtures of argon in neon gas (Penning and Addink[22] 1934). The effect due to an increase in α as a result of collisions of the second kind between neon metastable atoms and argon atoms is discussed in Section 1.2.3. This was followed up by direct measurement of α/p_o for argon from the V–i characteristics in a paper by Kruithof and Penning[23]

Table 1.5

Reported values of α/p for argon as a function of E/p

E/p (V cm^{-1} $torr^{-1}$)	α/p_o *Kruithof*[23, 25] *and Penning* (cm^{-1} $torr^{-1}$)	α/p *Huxford*[26] *taken from curve*		α/p *Huxford*[27] *and Engstrom taken from curve* (cm^{-1} $torr^{-1}$)	α/p *Davies*[28] *and Milne* (cm^{-1} $torr^{-1}$)
		CsAgO *cathode* (cm^{-1} $torr^{-1}$)	NiO *cathode* (cm^{-1} $torr^{-1}$)		
5	$2{\cdot}4 \times 10^{-4}$	—	—	—	—
6	$5{\cdot}6 \times 10^{-4}$	—	—	—	—
7	0·00113	—	—	—	—
10	0·00576	—	—	—	—
12	0·0129	—	—	—	—
15	—	—	—	—	0·0327
16	0·0384	—	—	—	—
20	0·0766	—	—	—	—
25	0·1405	—	—	—	—
30	—	—	—	—	0·182
40	0·433	—	—	0·53	—
50	0·677	—	—	0·85	—
60	0·935	—	—	—	0·834
100	1·986	—	—	2·3	1·99
140	3.019	—	—	—	—
150	—	—	—	—	3·12
200	4·466	4·0	—	4·1→5	—
240	5·3	—	—	—	—
250	—	—	—	—	5·48
400	7·87	—	—	6→9	7·96
1,000	12·3	8·1	7·89	~ 15·0	14·3
1,500	—	—	—	—	17·32

in 1936. To ensure clean conditions the system was thoroughly baked first in hydrogen, and then in vacuum before admitting the gas. The gas was cleaned by a high frequency discharge and also by evaporating barium getter films on to the walls of subsidiary bulbs connected to the main discharge tube. Their results are given in Table 1.5. In the next year the same authors[24] published values of α/p_o for pure neon and also for a range of neon-argon mixtures showing the marked effect of small admixtures of argon in neon. The values of α/p_o for neon are given in Table 1.3 and for neon-argon mixtures in Table 1.8. The latter are expressed in terms of η and the value of α/p can be deduced by multiplying η by the appropriate E/p. In 1940 Kruithof[25] using the same apparatus extended the measurements to include the inert gases krypton and xenon (α/p_o values for these gases are given in Table 1.6) and also extended the E/p range for argon.

Table 1.6

Values of α/p as a function of E/p for Krypton and Xenon according to Kruithof[25]

E/p (V cm^{-1} $torr^{-1}$)	α/p_o	
	Krypton (cm^{-1} $torr^{-1}$)	*Xenon* (cm^{-1} $torr^{-1}$)
5	0·00015	—
10	0·00414	0·00064
16	0·0317	0·0076
20	0·069	0·0218
32	0·250	0·126
40	0·418	0·248
50	0·660	0·441
60	0·927	0·666
80	1·470	1·198
100	2·035	1·804
200	4·834	4·940
500	10·45	12·52
1 000	14·67	19·77
2,000	16·78	25·34

Because of the care taken both in the experimental measurements and in the treatment of the results, Kruithof and Penning values of α/p_o for the inert gases are recognised as being reliable, and later measurements have been in general agreement. Most of these later measurements have been made over the past three or four years, the exception being in the case of argon, where Huxford published some values in 1939,[26] and with Engstrom in 1940,[27]

as part of a study of argon filled photo-cells. Because of fatigue of the photo-emissive cathode the results showed considerable spread. However values of α/p for argon taken from the graphs in these two papers have been included in Table 1.4 for comparison. More recent measurements in argon have been published by Davies and Milne[28] who have measured η for hydrogen, argon and neon using thin copper films deposited on a glass substrate as the electrodes. They claim cleaner surfaces by this method, and also less likelihood of gas impurities being evolved from the components since bulk metal is avoided. Their system used fixed electrode spacing, and E/p was varied by altering the pressure. It is not clear whether the gas handling system used with the inert gases could be considered as ultra-clean, such as that associated with bakeable taps etc., and in the case of neon their values indicate contamination, possibly by argon. However their results for argon given in Table 1.4 are in fair agreement with those of Kruithof and Penning. The only other data on α/p for argon was given in 1961 by Golden and Fisher[29] also employing fixed-gap tubes. At E/p below $9\text{Vcm}^{-1}\,\text{torr}^{-1}$ they found that α/p was not a unique function of E/p but dependent on the distance between electrodes, d. Their results are shown in Fig. 1.5. This departure from Paschen's law was attributed to the production of highly excited atoms by resonance radiation at some distance from the position where the electrons lose their energy: these excited atoms then produce molecular ions in collision with ground state argon atoms. Such a reaction will not obey the similarity principle. Disregarding this variation with d, their α/p values were lower than those of Kruithof and Penning particularly below $E/p = 7\text{V cm}^{-1}\,\text{torr}^{-1}$.

Reliable values for pure neon have recently been obtained by Chanin and Rork[30] using ultra-high vacuum techniques with bakeable metal valves and capacitance manometer. They used a metal envelope tube with nickel electrodes and went to considerable lengths to ensure the purity of the gas, using cataphoretic cleaning[31] for at least 24 hr to eliminate inert gas impurities. Their results are given in Table 1.3 and are in good agreement with Kruithof and Penning's values. Similar values (as η) have also been obtained by Smith[32] over a limited E/p range in fixed gap tubes with molybdenum electrodes which were thoroughly cleaned by extensive sputtering in a glow discharge. On the other hand, Fletcher and Davies[33] have also measured α/p for neon, but obtain values lower than previous workers, particularly at high E/p. They employed a gold film on a glass substrate as the electrodes in tubes similar to those used by Davies and Milne. Although they claim that their lower values are due to a higher purity of the gas, this

would appear doubtful since if this were so it would be expected to be more marked at the lower E/p values rather than at the higher values.

In the case of helium, for many years the only available values of α/p have been those of Townsend and McCallum obtained in 1934.

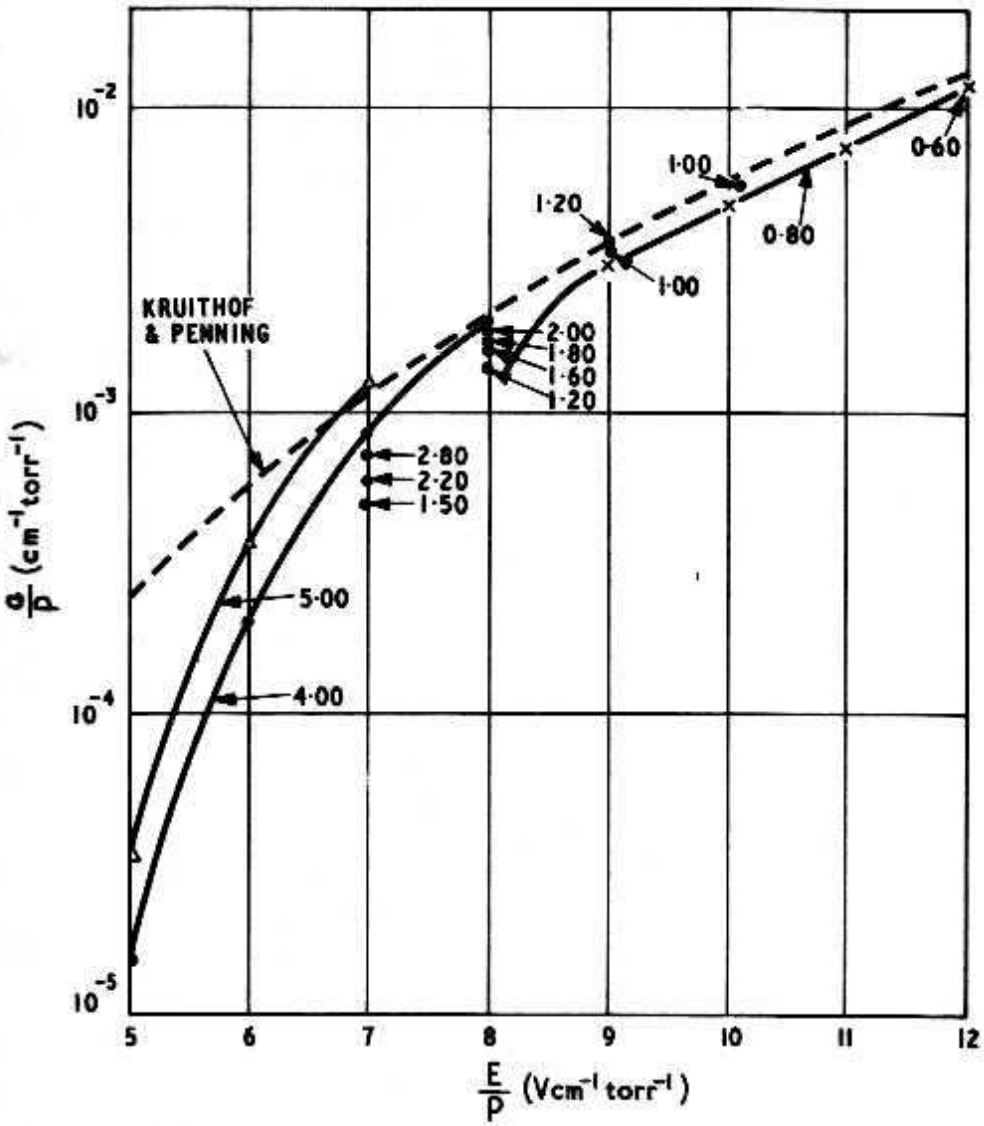


Fig. 1.5. Variation of α/p with E/p for argon for different electrode spacing according to Golden and Fisher.[29] The numbers associated with the experimental points indicate the value of d at which evaluations were made. Curves are drawn for d = 5·0, 4·0 and 0·8 mm. Dashed curve shows Kruithof and Penning's measurements

In 1962, however, Davies, Llewellyn Jones and Morgan [34] published values over an E/p range 6 –100 Vcm^{-1} torr^{-1} and two years later Chanin and Rork[35] gave values over a wider E/p range. Davies, Llewellyn Jones and Morgan used silver electrodes which had been carefully polished and employed high vacuum techniques. Their values of α/p are given in Table 1.5. They were rather lower than Townsend and McCallum values at low E/p ($\sim$ 10 V cm^{-1} torr^{-1}) but agreed in the region of 100 V cm^{-1} torr^{-1}. The agreement in fact is remarkable in view of the improvement in vacuum

technique which has taken place over the past 30 years. Chanin and Rork used nickel electrodes in a stainless steel envelope and went to rather greater lengths to ensure the purity of the gas. Their results, also in Table 1.5, are close to those of Davies, Llewellyn Jones and Morgan, which suggests that they represent values for uncontaminated helium. However, Fletcher and Davies[33] using evaporated film electrodes quote lower values particularly at $50 < E/p < 200\ \mathrm{V\ cm^{-1}\ torr^{-1}}$, and furthermore found that the values decreased with increased purification of the gas by cataphoresis. Approximate values taken from their curve are also given in Table 1.5. As with neon the explanation for these low values is difficult to understand.

A gas often used in cold cathode tubes where high switching speeds are required is hydrogen, and results have therefore been included on this gas. Hydrogen has been studied more extensively than the inert gases, but possibly because the gas is chemically active and can affect the cathode surface agreement between published values above $E/p = 100\ \mathrm{V\ cm^{-1}\ torr^{-1}}$ has not been as good as with the inert gases. A summary of the values published in the literature is given in Table 1.7. The first values free from mercury contamination are probably those of Hale[36] published in 1938. The values, however, are high compared with later measurements. In 1955 a number of values were reported over a limited range of low E/p showing good agreement. These are summarised by Rose[37] who made extensive measurements over the range $15 < E/p < 1{,}000\ \mathrm{V\ cm^{-1}\ torr^{-1}}$. In determining α from the log V-i plot, Rose assumes that at low field the second coefficient γ (see Section 1.3) could be neglected, but it was shown later by Jones and Llewellyn Jones[38] that this leads to considerable error above $E/p = 130\ \mathrm{V\ cm^{-1}\ torr^{-1}}$ and that lower values are obtained if γ is taken into account. Jones and Llewellyn Jones do not claim their gas and system to be ultra pure, although it is interesting that good agreement is obtained at the lower E/p values with later measurements (see Table 1.7). The fact that contamination in hydrogen at low E/p is not serious is borne out by experiments by Haydon and Robertson[39] in 1961 who measured α/p in systems which were not baked, but sealed with greased joints. A recent extensive study of hydrogen with thin film electrodes evaporated on to glass substrates has been carried out by Davies and his co-workers at Keele university. Their method, aimed at producing the purest conditions of gas and cathode surface have, however, shown rather large variations so that although initially[28] their values were higher than Jones and Llewellyn Jones their more recent publication[33] quotes values which are considerably lower.

Table 1.7

Reported values of α/p as a function of E/p for hydrogen

E/p (V cm^{-1} $torr^{-1}$)	α/p_o							
	Hale[36] *quoted in ref.* 3 (cm^{-1} $torr^{-1}$)	*Rose*[37] (cm^{-1} $torr^{-1}$)	*Jones*[38] *and Llewellyn Jones* (cm^{-1} $torr^{-1}$)	*Davies*[28] *and Milne* (cm^{-1} $torr^{-1}$)	*Haydon*[39] *and Robertson* (cm^{-1} $torr^{-1}$)	**Chanin*[30] *and Rork* (cm^{-1} $torr^{-1}$)	*Golden*[40] *Nakano and Fisher* (cm^{-1} $torr^{-1}$)	*Fletcher*[33] *and Davies* (cm^{-1} $torr^{-1}$)
20	—	0·0046	0·0046	—	—	—	—	—
30	0·123	0·0495	0·0495	—	—	—	—	—
50	—	0·315	0·315	0·317	0·315	0·310	0·33	0·39
100	0·785	1·26	1·25	1·26	1·25	1·29	1·34	1·23
200	—	2·55	2·12	2·38	2·38	2·64	2·50	2·0
300	3·200	3·21	2·40	2·97	2·55	3·57	3·34	2·18
400	4·29	3·6	—	3·22	2·55	3·7	3·60	2·23
500	—	3·9	—	3·46	2·55	3·7	4·00	—
600	5·810	4·1	—	—	—	—	3·80	—

*Values deduced from curve.

Probably one of the best experimental studies of hydrogen is that of Chanin and Rork in 1963[30] using stainless steel envelopes with nickel electrodes. They took elaborate precautions to outgas the tube thoroughly and clean the hydrogen, and their system employed modern ultra-high vacuum techniques. Their results are in close agreement with those of Rose, although γ was taken into account. At high E/p values they found that α/p depended on the portion of the $V-i$ curve used for calculations. They showed in fact that at these values of E/p, α/p was a strong function of the number of collisions, and thus the electrons were not in equilibrium with the applied electric field. They suggest, as a result, that at $E/p > 100\ \mathrm{V\ cm^{-1}\ torr^{-1}}$ measured values of α/p will tend to lose their significance, and that discrepancies between published values in this region can be accounted for.

However, Golden, Nakano and Fisher[40] pointed out that the variations in α/p could be explained by the method of analysis which was employed by Chanin and Rork to derive α/p from their experimental curves. Golden *et al* used a computer, and analysed their results by deriving the values of i_o, η, γ, and V_o which best fitted all the experimental data at each E/p_o. Their values given in Table 1.7 also agree with those of Rose. It is, therefore, concluded that above $E/p_o = 100\ \mathrm{V\ cm^{-1}\ torr^{-1}}$ the higher values of α/p reported in the literature are the most reliable.

1.2.3. THE EFFECT OF GAS MIXTURES ON THE FIRST TOWNSEND COEFFICIENT

Collisions resulting in ionisation of a gas other than those directly between electrons and atoms have been known for some time and are known as collisions of the second kind. In particular it is possible for an excited atom to ionise atoms of lower ionisation potential in a collision of this type. The importance of this reaction on the value of α was first observed by Penning and his co-workers in their study of neon-argon mixtures. In this gas an argon atom with an ionisation potential of 15·8V is ionised by collision with the metastable neon atom with an energy level of 16·58 eV:

$$\mathrm{Ne^* + Ar \rightarrow Ne + Ar^+ + e}$$

This reaction is in addition to the normal ionisation process, and thus the effect of adding Ar to Ne is to enhance the measured ionisation coefficient α, over that of either constituent. This phenomenon is known as the 'Penning effect' in honour of its discoverer. Very low concentrations of argon, $10^{-5}\%$, have only a

Kruithoff and Penning's Values of $\eta = \alpha/E$ as a $f(E/p)$ in Ne-Ar Mixtures, Reduced to 0°C, for Various Quantities of Ar

(Mean Values of $100\eta = 100\alpha/E$)[23]

	η (V^{-1})									
% Ar / E/p_o ($Vcm^{-1}torr^{-1}$)	0	10^{-4}	3.10^{-4}	10^{-3}	3.10^{-3}	0.01	0.1	1	10	100
0·5	—	0·008	0·021	0·050	0·095	0·146	0·048	—	—	—
0·6	—	0·026	0·065	0·158	0·298	0·461	0·309	0·021	—	—
0·7	—	0·051	0·125	0·302	0·569	0·879	0·75	0·068	—	—
0·8	—	0·076	0·185	0·444	0·835	1·28	1·26	0·148	—	—
0·9	—	0·097	0·237	0·568	1·07	1·65	1·68	0·263	—	—
1·0	—	0·116	0·281	0·670	1·26	1·94	1·99	0·408	—	—
1·2	—	0·146	0·347	0·827	1·55	2·38	2·54	0·76	0·0113	—
1·5	—	0·175	0·410	0·969	1·81	2·79	2·97	1·24	0·0385	—
2·0	0·033	0·203	0·461	1·07	2·00	3·07	3·43	1·88	0·105	—
2·5	0·046	0·221	0·484	1·11	2·06	3·15	3·67	2·27	0·201	—
3·0	0·064	0·238	0·498	1·12	2·06	3·14	3·71	2·58	0·319	—
4·0	0·106	0·272	0·522	1·12	2·01	3·05	3·70	2·93	0·590	—
5·0	0·159	0·313	0·546	1·10	1·94	2·91	3·64	3·11	0·852	—
6·0	0·221	0·358	0·575	1·09	1·87	2·78	3·56	3·22	1·12	0·0095
8·0	0·362	0·476	0·664	1·11	1·77	2·56	3·33	3·23	1·50	0·0259
10·0	0·501	0·607	0·768	1·15	1·73	2·40	3·22	3·20	1·77	0·0576
12·0	0·623	0·728	0·851	1·18	1·69	2·27	2·98	3·14	1·94	0·108
15·0	0·772	0·866	0·963	1·24	1·65	2·13	2·79	3·07	2·18	0·203
20·0	0·962	1·06	1·11	1·32	1·58	2·01	2·58	2·97	2·37	0·383
25·0	1·103	1·17	1·23	1·38	1·57	1·93	2·43	2·87	2·50	0·56
30·0	1·210	1·27	1·31	1·43	1·56	1·88	2·30	2·80	2·58	0·74
40·0	1·351	1·37	1·41	1·48	1·58	1·81	2·15	2·67	2·63	1·08
50·0	1·430	1·44	1·46	1·51	1·57	1·75	2·06	2·58	2·63	1·36
70·0	1·482	1·48	1·49	1·51	1·54	1·66	1·93	2·32	2·54	1·71
100·0	1·493	1·49	1·49	1·49	1·49	1·53	1·79	2·08	2·38	1·99
150·0	1·406	1·41	1·41	1·41	1·43	1·46	1·60	1·77	2·08	2·18
200·0	1·298	1·30	1·30	1·30	1·31	1·34	1·42	1·55	1·79	2·23
250·0	1·182	1·18	1·18	1·18	1·20	1·22	1·27	1·36	—	2·19
300·0	1·073	1·08	1·08	1·08	1·09	1·11	1·15	1·22	—	2·13
400·0	0·890	0·89	0·89	0·89	0·89	0·89	0·90	0·93	—	1·97

small effect on the value of α. As the concentration is increased, α increases until a maximum is reached. Because the energy levels involved are close, the probability of the reaction is high and the maximum in α occurs with quite small admixtures of argon (0·1 %). At concentrations above this value direct ionisation and excitation of the argon takes place, reducing the average energy of the electrons and resulting in fewer neon metastables. The increase in α due to the Penning effect can be seen from Kruithof and Penning's values of η[24] for various admixtures of argon in neon given in Table 1.8 and expressed graphically in Fig. 1.6. The maximum value of η for the neon-argon mixtures occurs at an E/p value approximately corresponding to the electron energy for optimum production of neon metastables. This is at a lower E/p than the maximum for pure gases which occurs at optimum electron energy for ionisation. The two maxima can be seen clearly for most of the mixtures.

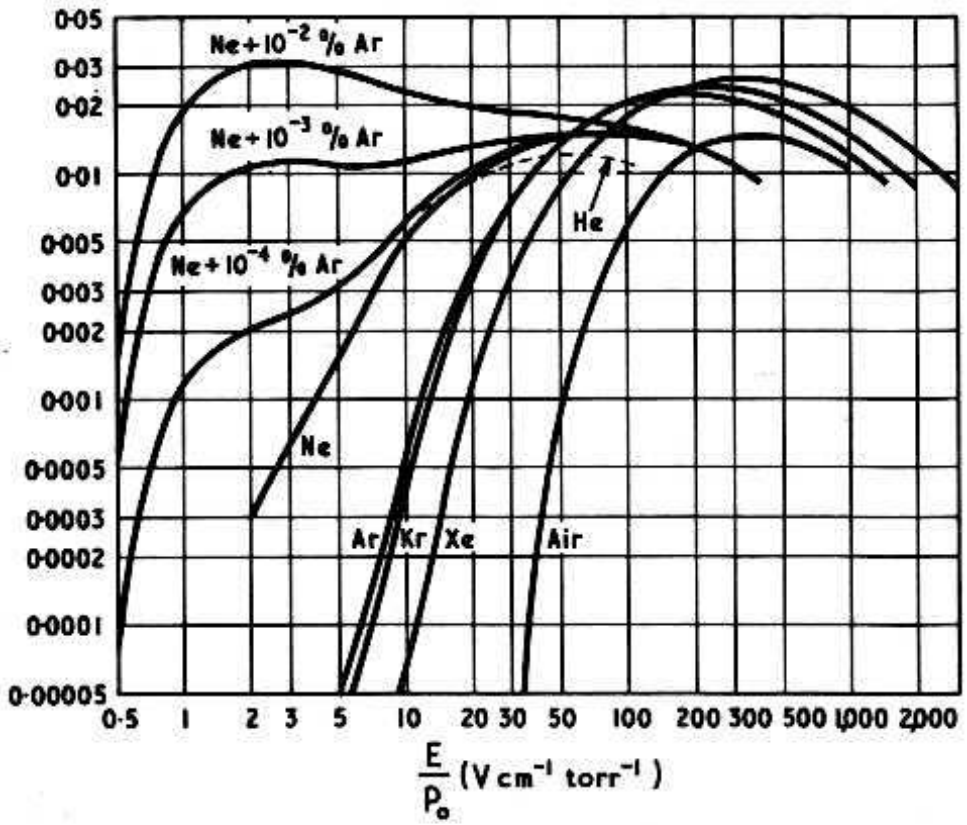


Fig. 1.6 Values of η for the rare gases, some Ne-Ar *mixtures and air, as a function of E/p (curves taken from Druyvesteyn and Penning*[5]*)*

Since all inert gases have metastable states, other inert gas mixtures will exhibit the Penning effect. As example He-Ar, He or Ne with Kr or Xe, and mercury vapour in He, Ne, or Ar. Further, most of the molecular gases evolved from surfaces used in gas discharge tubes such as CO_2, O_2, H_2O, have ionisation potentials below the metastable levels of He and Ne, and thus measure-

ments in pure gases are very sensitive to impurities. The effect on α will depend on how close the ionisation potential is to the metastable level of the parent gas, since the probability of the reaction is a function of the difference in energy, showing a resonant effect when the two are equal. Although the consequences of the 'Penning

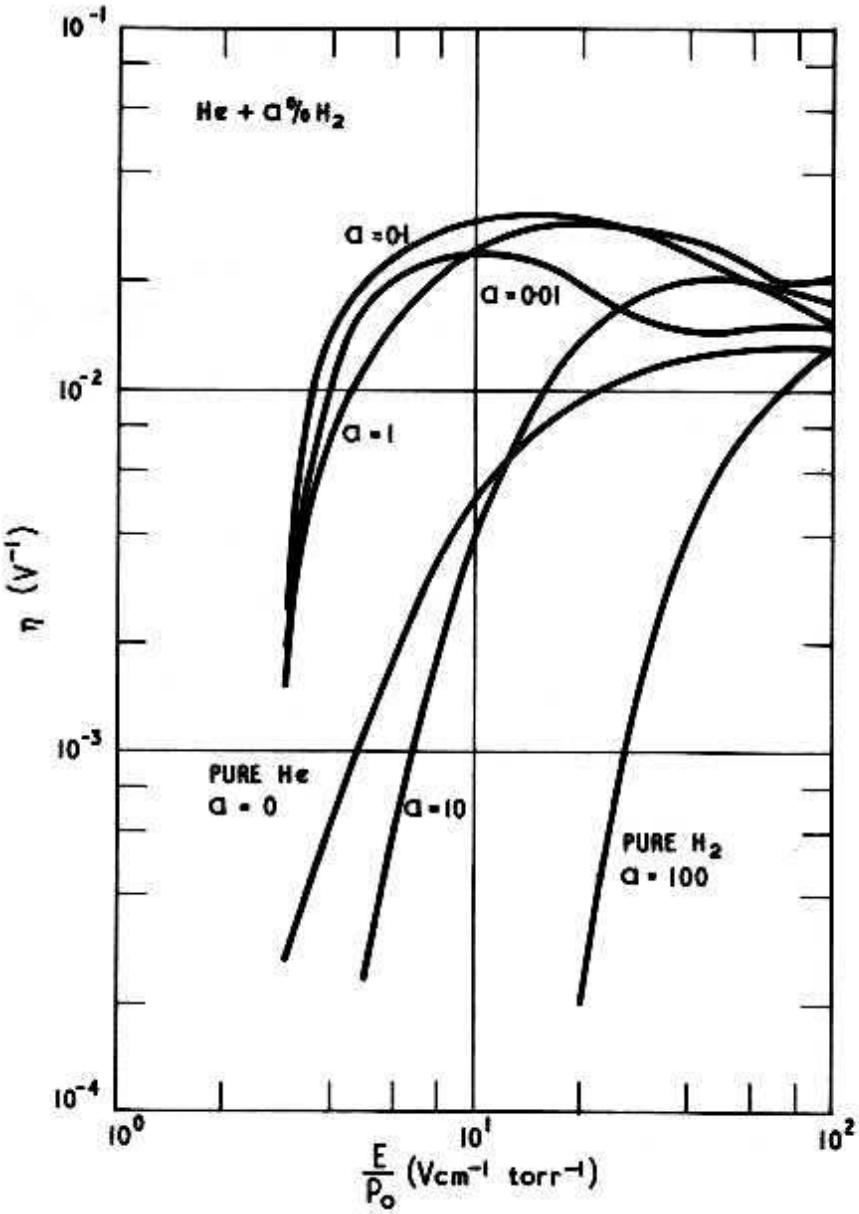


Fig. 1.7. The variation of η with E/p_o in He-H *mixtures according to Chanin and Rork*[41]

effect' in terms of the breakdown and maintaining potentials have been reported for gas mixtures other than Ne-Ar only one other published paper has been presented on the direct measurement of α/p. This is a recent paper by Chanin and Rork[41] using He-H and Ne-H gas mixtures. Their results are shown in Figs. 1.7 and 1.8. As with Ne-Ar both mixtures exhibited the maximum effect with concentrations of the lower ionisation potential gas of the order

0·1 %. At this concentration Ne-H gave a maximum η of 0·0285 at E/p of 3V cm^{-1} torr^{-1} whereas He-H gave a value of 0·0302 at $E/p = 15$V cm^{-1} torr^{-1}. Since the maximum value of η occurs at the E/p value for optimum production of metastables of the parent gas, the maximum for Ne-H is at the same E/p as for Ne-Ar, and lower than for hydrogen in helium, which has a higher metastable level.

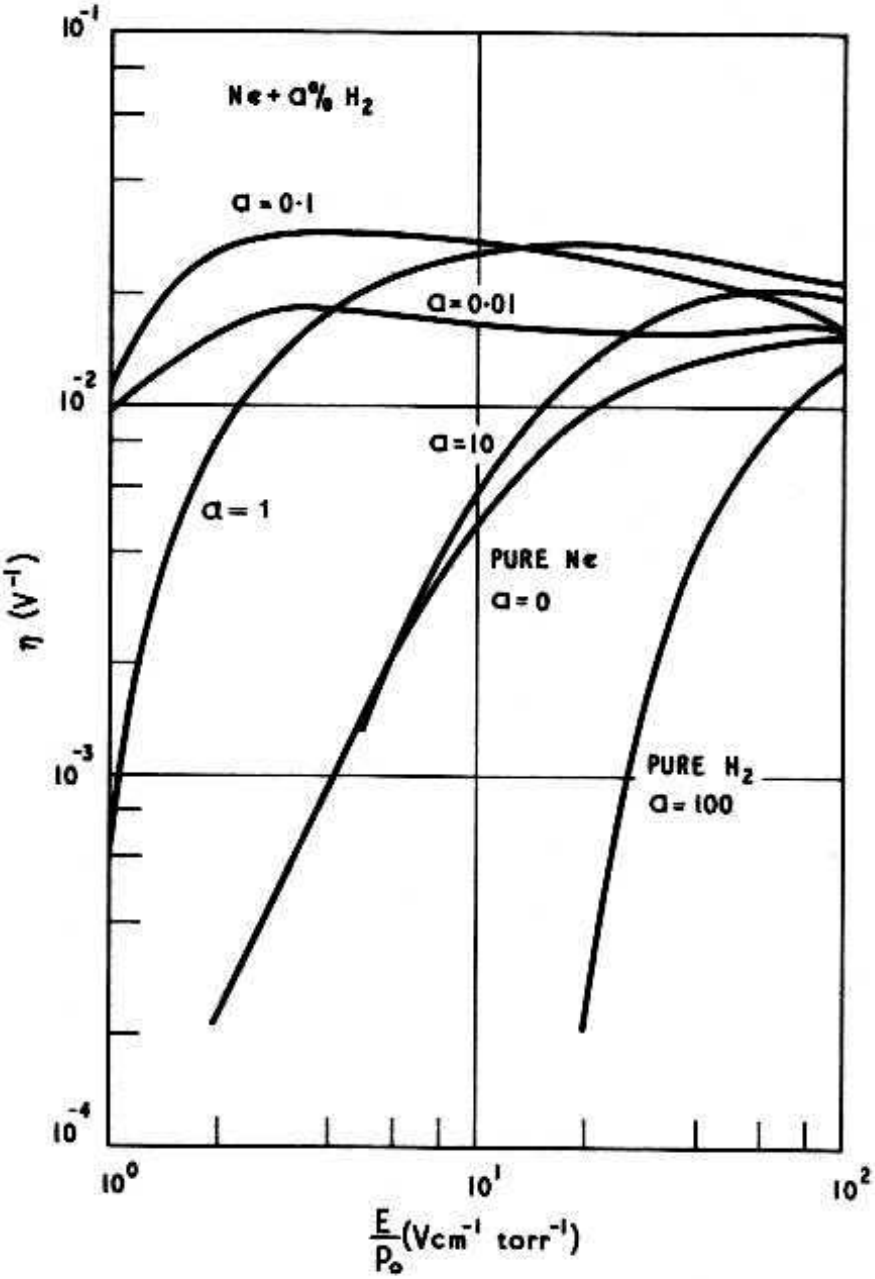


Fig. 1.8. The variation of η with E/p_o in Ne-H *mixtures according to Chanin and Rork*[41]

Apart from the Penning effect, actions in the gas involving photons are possible which enhance the value of α. An example of this was reported by Hale[36] using hydrogen with sodium-coated platinum electrodes. He observed an enhancement of α corresponding to a decrease in the secondary electrons produced by photons at the cathode, and ascribed the effect to photoionisation

of some volatile compound of hydrogen and sodium. In general, however, photoionisation will only occur at higher pressures where $1/\mu$ (μ = the coefficient of absorption for photoionisation radiation) is less than the distance between the electrodes. The effect is unimportant in normal glow discharge tubes working in the pressure range of 10–50 torr.

1.3. Townsend's Second Coefficient

Although the initial increase in current above i_o in Fig. 1.1. is exponentially dependent on the applied voltage or the gap distance for a given E/p, at higher values of potential the current increases more rapidly, and Equations 1.2 and 1.3 no longer apply. Further, since the equations imply that the current remains proportional to i_o, they cannot explain the onset of a self-maintained discharge. To explain this increase in current, Townsend introduced the idea of a secondary process in the gas giving rise to further electrons. Apart from ionisation collisions between electrons and gas atoms or molecules, other inelastic collisions will take place in which the atoms are excited either to metastable levels or to levels from which they will decay with the emission of photons. Thus in the gas there are ions, metastable atoms, and photons available as agents in the production of secondary electrons either in the gas or at the cathode.

Townsend assumed that the secondary effect was due to ionisation collisions between the positive ions and the gas atoms, and expressing the number of electrons produced by one positive ion travelling unit distance as β he derived the formula:

$$i = i_o \frac{(\alpha-\beta)\exp(\alpha-\beta)d}{\alpha-\beta\exp(\alpha-\beta)d} \tag{1.8}$$

It is now known that the ionisation of molecules by positive ions is highly unlikely under normal breakdown conditions, and that the secondary electrons are either ejected from the cathode by ions, metastable atoms, or photons incident on the surface, or formed by photoionisation in the gas. At low pressures, below atmospheric, the influence of photoionisation is negligible, and therefore in most glow discharge tubes the secondary process will take place at the cathode. Whatever secondary mechanism is present singly or jointly, the corresponding equations for the anode current are similar to the original Townsend equation, and it can be shown[3] that to a first approximation all the secondary processes

can be expressed by a single coefficient γ in the equation:

$$i = i_o \frac{\exp \alpha(d-d_o)}{1-\gamma[\exp \alpha(d-d_o)-1]} \tag{1.9}$$

Where γ is the total number of electrons produced per positive ion formed in the discharge, and can be expressed as a linear sum representing separately the action of photons, γ_p; positive ions, γ_i; and metastables, γ_m at the cathode, i.e.

$$\gamma = \gamma_p + \gamma_i + \gamma_m$$

From Equation 1.9 i/i_o becomes infinite when the denominator becomes zero. Townsend set this condition as the threshold for a spark breakdown. Thus at breakdown potential V_s:

$$\gamma \exp \alpha(d-d_o) - \gamma = 1 \tag{1.10}$$

or using η,

$$\gamma \exp \eta(V_s - V_o) - \gamma = 1 \tag{1.11}$$

These equations give the relation between breakdown potential and Townsend coefficients α and γ, and for gases at low pressure good agreement is obtained when the coefficients and the breakdown potential are determined separately. Clearly γ depends on the gas and the cathode material; it will also depend on the pressure and on the field.

Measurements of γ as a function of E/p have been made by a number of investigators. Since, however, γ includes γ_i, γ_p and γ_m it is not surprising that the relationship between γ and E/p is very dependent on the type of cathode employed. As an example Kruithof[25] measured γ for copper electrodes with the inert gases Ne, Ar, Kr and Xe. He found that at high E/p (i.e. $p \to 0$) γ decreased as E/p decreased. This was attributed to the loss of electrons by diffusion back to the cathode as the pressure is increased. However in the case of Kr and Xe, γ increased again at lower E/p values, and this he attributed to electrons released by photons, whereas at higher E/p positive ions played the major role. However, both the shape of the curves and the average values depended on the surface condition of the cathode.

The dependence of γ on the surface state of the cathode is of prime importance, since it affects breakdown and maintaining potential characteristics of a discharge. Experience with cold cathode tubes has shown that high stability of these characteristics is only obtainable with a limited number of metals and then only after extensive cathodic sputtering to clean the surface. γ values obtained from the $V-i$ characteristics below breakdown

(Townsend method) using tubes with large electrode structures, or values obtained from high pressure measurements where cathode cleaning cannot be assured, are therefore of little value to the cold cathode tube designer. For this reason, although curves of γ against E/p obtained this way are to be found in the literature (Loeb[3] summarises most of the work up to 1955) they are not given here, and the valve engineer or physicist should use them with caution.

The relative importance of the components of γ due to ions, metastables and photons is, however, of interest, and consideration is given to the separation of the effects and their measurement in the next two sections.

1.3.1. DYNAMIC STUDIES OF THE SECONDARY MECHANISM

Normal static measurements cannot segregate the various γ mechanisms which may or may not be present simultaneously. However, each mechanism will be associated with a characteristic time interval resulting from the transit time of the active particle (for example an electron will take some 10^{-7} sec to cross a 1 cm gap whereas ions take about 10^{-5} sec and metastables 10^{-4}–10^{-3} sec depending on pressure and gas composition) and a study of the temporal growth or decay of a discharge can be used to distinguish the various processes.

The first attempt to use a dynamic method to separate the γ mechanisms was that of Engstrom and Huxford[27] in 1940, arising from the study of current delays in gas photocells. The delays suggested a metastable action, and the experiments were based on a time scale in which the ejection of electrons by photons and ions were considered as instantaneous. Their method involved applying light pulses, by means of a rotating sector in front of a filament lamp, to the cathode of a parallel-plate discharge tube similar to that used by Kruithof[25]. The current was measured by an electrometer at a time t after the light had either been turned on or extinguished. By varying t they obtained curves similar to that in Fig. 1.9. Argon gas was used, and Townsend's first coefficient was measured from the static characteristics. The total γ was taken from the time-lag curves, since the total current amplification, A, is given by:

$$A = \frac{i}{i_o} = \frac{\exp \alpha(d-d_o)}{1-\gamma[\exp \alpha(d-d_o)-1]} \tag{1.12}$$

where i is the height of the pulse.

The effect of photons and/or positive ions (γ_p and γ_i) can be obtained by measuring the instantaneous current fraction i_a from which

$$i_a A = \frac{\exp \alpha(d-d_o)}{1-(\gamma_p+\gamma_i)\left[\exp \alpha(d-d_o)-1\right]} \tag{1.13}$$

γ_m due to the metastables will be given by the difference between the secondary coefficients deduced from Equations 1.12 and 1.13, i.e.

$$\gamma_m = \gamma-(\gamma_p+\gamma_i) \tag{1.14}$$

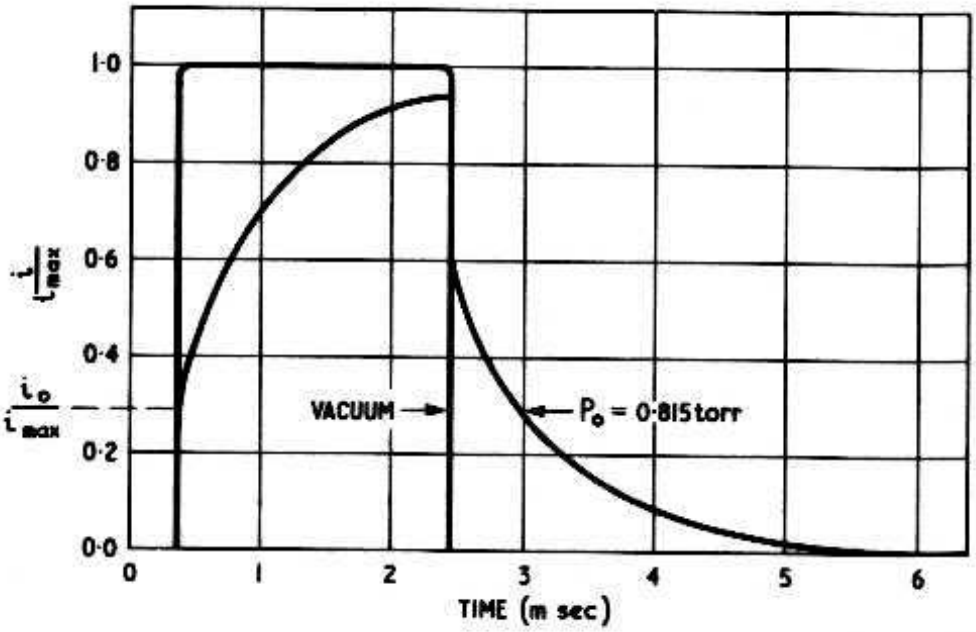


Fig. 1.9. Samples of the time-lag curves as measured by Engstrom and Huxford[27] for vacuum and for argon at p_o = 0·815 torr (d = 0·399 cm and V = 57 V)

Analysis of the curved portion of the time-delay curve gave information on the time constant from which the coefficient of diffusion and also the efficiency of excitation to the metastable level in argon could be deduced. Values of total γ and γ_i, which include the contribution due to photons, as a function of E/p are given in Figs. 1.10 and 1.11 respectively. These are for silver-caesium-oxide cathodes with argon gas. As can be seen at low E/p photons or ions play a small role in the secondary process and the contribution due to metastables is important. As E/p increases the contribution due to ions or photons becomes more pronounced and metastables play very little part.

By increasing the time resolution with oscilloscopes the dynamic method was extended by Gugelberg[42], Molnar[43] and Hornbeck[44] to resolve the effect of photons, positive ions, and metastables at the cathode. This necessitated complete analysis of the build up and decay curve of the discharge in terms of all three mechanisms.

Gugelberg used an analysis involving emission by photons and positive ions, and omitted the effect of metastables, although he recognised that in some of his experiments metastable atoms were involved. In the work of Molnar and Hornbeck, both at the Bell Laboratories, Hornbeck studied very short time delays involving

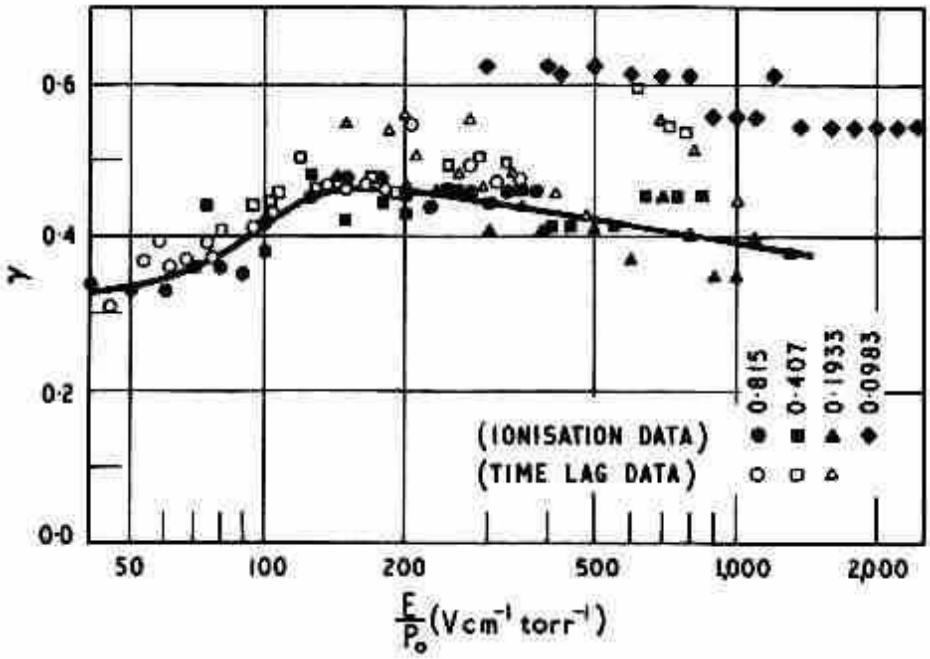


Fig. 1.10. Variation of the second Townsend coefficient γ with E/p_0 for argon with an activated caesium cathode according to Engstrom and Huxford.[27] Solid points are from ionisation data, and open points from time-lag data

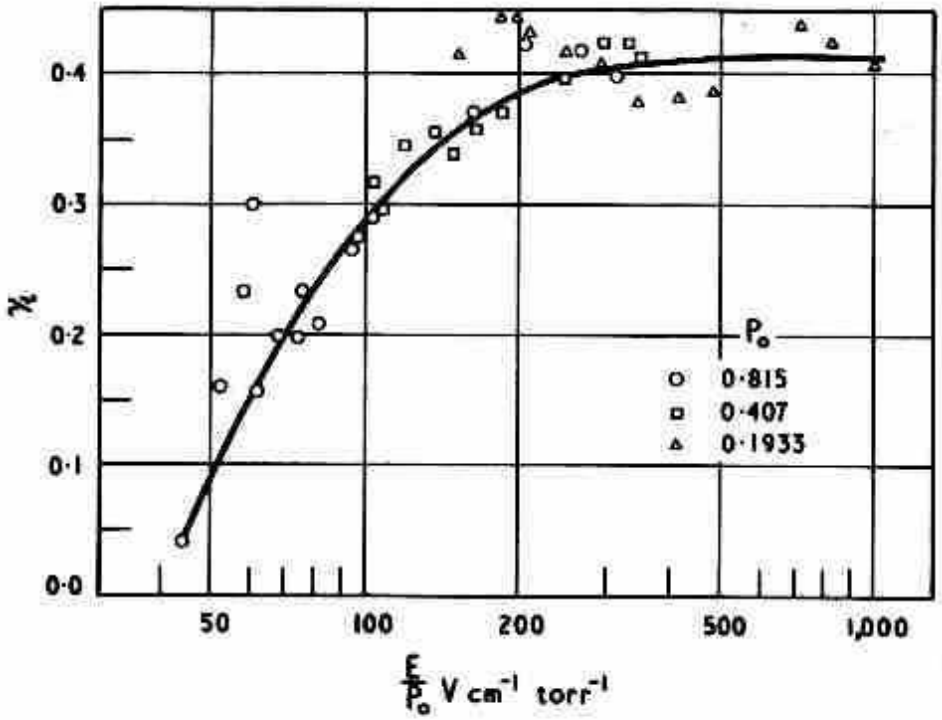


Fig. 1.11. Variation of γ_i, the number of secondary electrons emitted per positive argon ion, with E/p_0 for an activated caesium cathode according to Engstrom and Huxford[27] (γ_i calculated from Equations 1.12 and 1.13)

photon and positive ion actions, whilst Molnar looked at longer time intervals involving metastables. The basic analysis developed by Engstrom and Huxford for the latter case was extended by Molnar to include metastables destroyed in the gas. The destruction of the metastables may lead to photons which act at the cathode and contribute to the long term delay. Similar to Engstrom and Huxford, Molnar measured the total current and the fast instantaneous current to determine γ and γ_f (fast)—from Townsend's equation. γ_s for the slow component was derived from the relation $\gamma = \gamma_f + \gamma_s$ as before. If γ'_m and γ'_p are defined as the number of electrons emitted at the cathode by one metastable and one photon respectively, rather than as a function of the number of positive ions, then

$$\gamma_s = f_{ex}\left(\frac{\alpha_m}{\alpha}\right)(f_{mk}\gamma'_m + f_{mp}f_{pk}\gamma'_p) \tag{1.15}$$

where f_{ex} = fraction of electrons emitted escaping back-diffusion to the cathode.

f_{mk} and f_{pk} = fraction of metastables and photons respectively reaching the cathode.

f_{mp} = fraction of metastables giving rise to radiation in the gas.

and α_m = number of metastables formed in the gas per cm per electron.

By radiating the space between the electrodes with light of appropriate wavelength, it was possible to convert metastables into radiating atoms, thus varying γ'_p/γ'_m. In this way the contribution of γ'_p could be assessed. In Table 1.9 Molnar's values of γ_f, γ_s, γ_i, γ'_m and γ'_p are given for cathodes of tantalum, molybdenum and barium oxide coated tantalum, with argon gas. Molnar does not claim that the metal surfaces were particularly clean, but it shows that although γ'_m is in most cases of the same order as γ_i, due to the factors α_m/α, f_{mk}, the contribution by metastable action to the total γ is small compared with γ_i (usually less than 10%). For pure metal cathodes the contribution due to photons γ_p from the metastables was even smaller.

Hornbeck's experiments were based on the use of a very short light pulse of $0{\cdot}1\mu$ sec duration, and the examination of the current transient that followed. The transient consisted of a sharp spike attributed to the current carried by the initial photoelectrons and primary electrons formed in the gas followed by a smaller current

Table 1.9

Molnar's[43] values for γ_f γ_s γ_i γ'_m γ'_p for Ta Mo and BaO on Ta cathodes in Argon gas

E/p (V cm^{-1} torr^{-1})	*Cathode*	γ_f	γ_s	γ_i	γ'_m	γ'_p
195·4	Ta	0·026	0·0025	0·026	0·023	0·009
195·4	Mo	0·076	0·0074	0·071	0·060	0·005
195·4	BaO	0·23	0·050	0·23	0·31	0·031
117·2	Ta	0·022	0·0026	0·022	0·023	0·009
117·2	Mo	0·065	0·011	0·071	0·065	0·005
117·2	BaO	0·21	0·065	0·22	0·27	0·027
72·6	Ta	0·0060	0·0009	0·0053	0·0035	0·0021
72·6	Mo	0·030	0·00025	0·034	0·020	<0·001
72·6	BaO	0·073	0·013	0·085	0·078	—

lasting some tens of micro-seconds originating from the motion of the positive ions and the release of secondary electrons at the cathode by these ions. An idealised curve of this smaller current is shown in Fig. 1.12 on a time scale of ion transit time t_i. Secondary

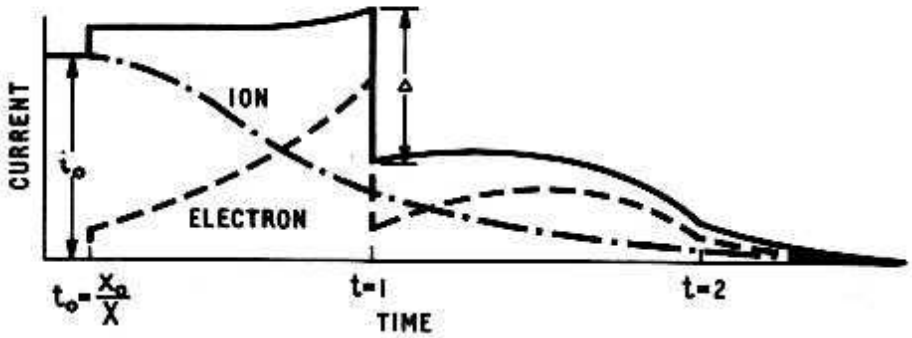


Fig. 1.12. Drawing of idealised transient current vs time following the formation at $t = 0$ of an exponential distribution of positive ions between the electrodes (time expressed in units of ion transit time)

electrons produced by photons are assumed to be indistinguishable from the initial photoelectric current, and metastable atoms would not reach the cathode on the time scale examined. Referring to Fig. 1.12, i_o represents the ion current due to the initial photoelectric current and Δ is the observed discontinuity at $t = 1$ ion transit time when all these initial ions have reached the cathode. γ_i can be derived in terms of i_o, t_i, Δ, and A the integrated area under the initial photocurrent by an approximate expression:

$$\gamma_i = \frac{\Delta}{i_o} \exp\left[-\frac{i_o t_i}{A}\left(1 - \frac{d_o}{d}\right)\right] \tag{1.16}$$

Since the measurement of i_o and A could not be obtained with the necessary accuracy, and in any case the analysis was not capable of evaluating γ_i to better than an order of magnitude, Hornbeck made very few γ_i measurements. However the method proved useful in evaluating ion drift velocities and in particular distinguishing ion species occurring at various E/p values.

The method, however, was taken up by Varney[45] using a rather better analysis. From his theory the current as a function of time normalised to the ion transit time (i.e. v^+t_i/d) gave curves similar to those in Fig. 1.13. When the first part of the curve was horizontal Varney showed that γ_i could be derived from the equation

$$\gamma_i = \left[\frac{\exp \alpha d}{\exp \alpha d - 1}\right]^{\frac{1}{2}} - 1 \tag{1.17}$$

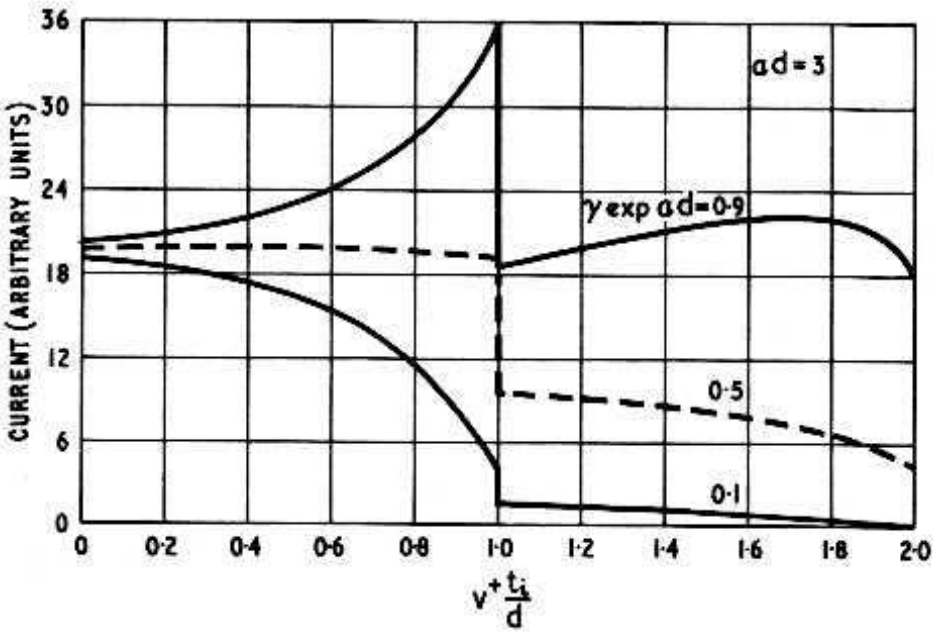


Fig. 1.13. Varney's[45] solution for the current as a function of v^+t_i/d for various values of γ_i exp αd where $\alpha d = 3$

Varney's method was to set the pressure and electrode separation on the parallel-plate discharge tube and to vary the voltage until a curve of the right contour was obtained for the current transient after the 10^{-7} sec light flash. Values of α were taken from Kruithof's work,[25] and γ_i found for molybdenum cathodes with neon, argon and krypton, and for a coating of barium oxide on molybdenum, for argon. The results are shown graphically in Fig. 1.14.

Rose[46] has also applied this method for obtaining γ_i of a molybdenum cathode with hydrogen and deuterium gas. Rose applied a correction to account for photoemission which was observed at the commencement of the pulse. His corrected values for the two gases are shown in Fig. 1.15. Electrons emitted by photon action

at the cathode accounted for approximately a third of the total γ, and are therefore significant in hydrogen discharges. In a further investigation, reported in the same paper, of the role of photons in the glow discharge, Rose concluded that although γ_p was

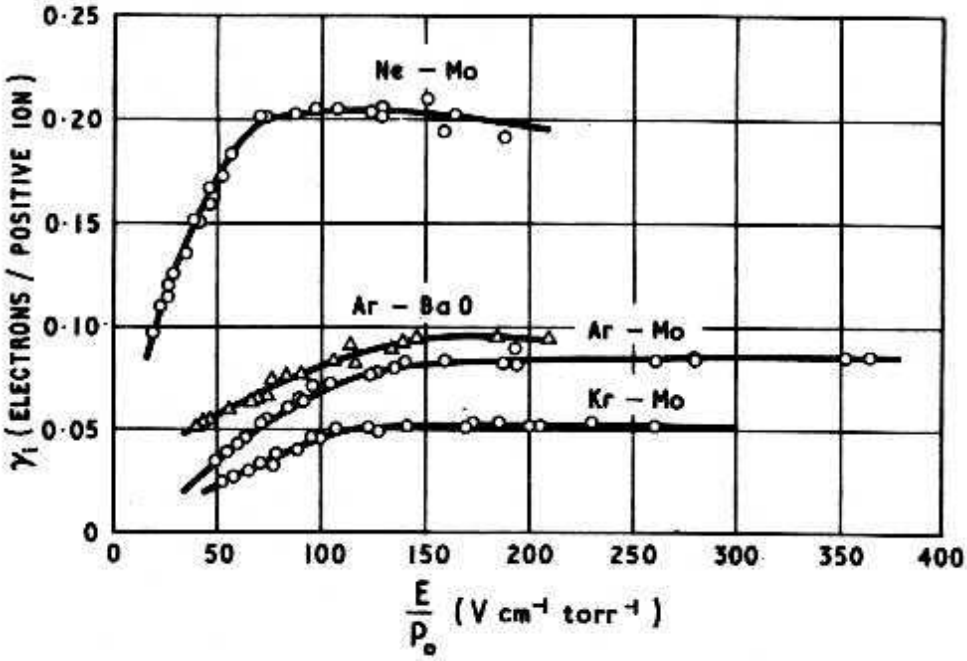


Fig. 1.14. The variation of γ_i with E/p_o for neon, argon and krypton ions at a molybdenum cathode, and also for argon ions at a barium oxide coated cathode (Varney[43])

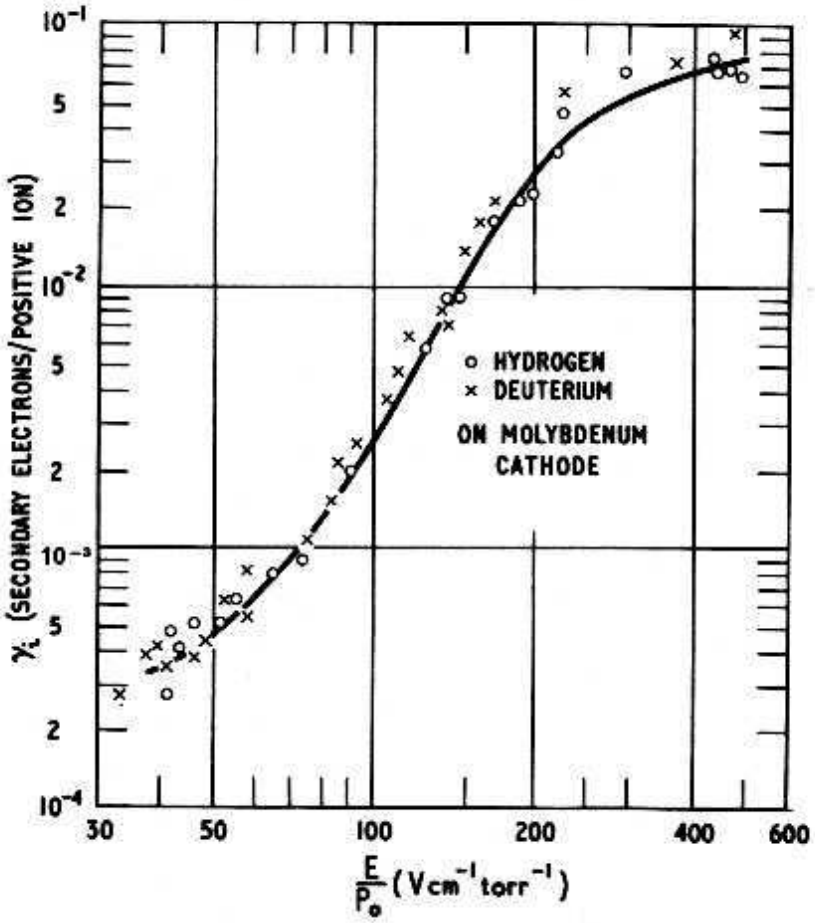


Fig. 1.15. The variation of γ_i with E/p_o for hydrogen and deuterium ions at a molybdenum cathode (Rose[46])

approximately $\frac{1}{2}$ of γ_i in hydrogen, in neon γ_p was less than $\frac{1}{10}$ of γ_i assuming previously published values of γ_i for neon. The relative unimportance of γ_p for pure metals with inert gases has been observed by other workers, in particular by Lauer[47] and Colli and Facchini,[48] using pulse methods with cylindrical geometry. This method was capable of rather higher resolution than the experiments of Hornbeck using plane parallel-plate geometry. Lauer observed no γ_p at all in argon with nickel cathodes, but Colli and Facchini showed in fact that there was a small effect which increased with pressure. Below 50 torr, γ_i was the major contributory factor in the establishment of a discharge, but above 200 torr the action was almost entirely due to γ_p.

Another method of separating the secondary mechanisms is to maintain a constant illumination to provide the initiating current i_o, and to pulse the potential across the electrodes. This method has been explored by Llewellyn Jones and his co-workers at Swansea University. The potential is pulsed above the static breakdown potential and the delay in breakdown determined as a function of the overvoltage. An analysis of the resultant curve has been given by Davidson,[49] to cover most secondary mechanisms, from which the prevailing action can be deduced. Llewellyn Jones and his colleagues have been interested in active gases at atmospheric pressure, and their results are not directly applicable to cold cathode tubes. However they were able to verify that the short time delays associated with discharges at high pressure are in accordance with Townsend's equation,[50] and that it is not necessary to invoke an alternative mechanism at a critical pressure. Such an alternative mechanism has been introduced by Raether[51] and expounded in detail by Loeb and Meek,[52] and is known as the 'streamer' mechanism. The arguments for and against the principles involved is outside the scope of this book and the interested reader is referred to Loeb and Meek's book and that of Llewellyn Jones.[53]

An interesting development of the work at Swansea has arisen out of an extension of Davidson's theory to include field distortion as space charge develops. In a paper by Davies, Evans and Llewellyn Jones[54] it was shown that short delay times at high pressures, hitherto associated with photon action either in the gas or at the cathode, could be explained entirely by a positive ion action at the cathode, γ_i, if the discharge is taken to be of a small diameter. Since the discharge path after breakdown is usually filamentary at such pressures the premise seems very likely. The results of Llewellyn Jones group are given in more detail in Chapter 5.

The main conclusion from the dynamic study of the secondary

process then, is that the ejection of electrons from the cathode by incident ions is the most important secondary process, especially in inert gases with pure metal cathodes under conditions pertinent to cold cathode tubes. Direct measurement of γ_i by ion beams, which allows very accurate and reliable data to be obtained, is therefore useful in assessing the characteristics of cathodes in glow discharge tubes.

1.3.2. DIRECT MEASUREMENT OF γ_i

It had been known for many years that electrons could be ejected from a metal surface in vacuum under bombardment by positive ions. Most of the early experiments, however, concerned beams of highly energetic ions, higher than encountered in a discharge, where the kinetic energy played an important role. In the early thirties, however, as a result of the study of gas discharges and the experiments of Oliphant[55] and others with low energy ion beams, it was evident that electrons could be ejected by ions having little or no kinetic energy. In these circumstances the potential energy of the ion, as a consequence of the ionisation potential, must be the controlling factor. Oliphant and Moon[56] suggested a two-step process based on this assumption; firstly the ion approaching the surface causes field emission of an electron which neutralises it to an excited atom, and then the excited atom interacts with a surface electron transferring its energy to the electron enabling it to escape into the vacuum. This physical picture explained the experimental facts, but data collected around that period is suspect due to the relatively poor vacuum systems and likelihood of the surfaces being contaminated. With the advent of better vacuum systems capable of obtaining pressures of less than 10^{-10} torr, reproducible and reliable values of γ_i have been obtained, mainly by Hagstrum at the Bell Laboratories who carried out a series of investigations covering a wide range of gas ions and metal surfaces, with particular reference to the behaviour with clean surfaces and the effect of monolayers of gas on them. From his measurements Hagstrum[57] concluded that a single-step process, first suggested by Shekhter,[58] was the most likely for the inert gases bombarding metal surfaces. The single step 'Auger' process can best be explained in terms of the schematic energy level diagram of Fig. 1.16. As the ion approaches the surface it captures an electron from the conduction band of the metal, neutralising the ion to the ground state. The energy released is absorbed by a second electron in the metal because the wave functions of the two electrons overlap. If the energy transferred is

large enough and the momentum correctly orientated the electron can escape. The maximum energy released is $E_i - \Phi$, where E_i is the ionisation energy of the ion and Φ the work function of the metal, this occurring when the captured electron is from the top of the conduction band. The minimum energy required for an electron to escape occurs when the second electron is similarly placed, and is equal to Φ. Thus the condition for secondary emission to occur at all is that $E_i - \Phi > \Phi$ or that $E_i > 2\Phi$. For the inert gases this will be fulfilled for all metal surfaces.

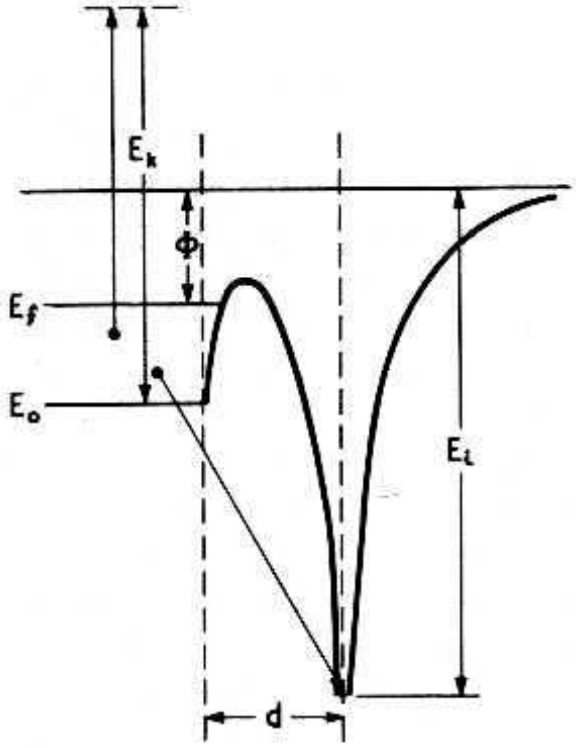


Fig. 1.16. Energy-level diagram for the ion-metal system illustrating the Auger neutralisation process for the emission of a secondary electron

E_i = ionisation energy of the ion
E_o = energy at the bottom of the conduction band
Φ = work function of the metal
E_f = Fermi level of the metal
E_k = energy of the excited electron

Hagstrum has measured γ_i for tantalum[59] molybdenum[60] and tungsten[61] surfaces with various inert gas ions. The values of γ_i for singly- and doubly-ionised atoms of 10 eV energy is given in Table 1.10; the energies of ions in a glow discharge are likely to be rather less than 10 eV. The variation with ion energy for singly-

Table 1.10

Electron yield for 10 eV ions according to Hagstrum[60, 61]

Gas	*γ_i for single ionised atom*			*γ_i for doubly ionised atom*		
	Mo	W	Ta*	Mo	W	Ta*
He	0·300	0·290	0·14	0·81	0·74	0·56
Ne	0·254	0·213	—	0·68	0·72	—
A	0·122	0·095	—	0·42	0·42	—
Kr	0·069	0·050	—	0·39	0·32	—
Xe	0·022	0·013	—	0·30	0·24	—

*The tantalum was probably covered by a monolayer of gas.

ionised atoms on molybdenum and tungsten is relatively small and is shown in Fig. 1.17.

From the early measurements of Hagstrum, and those of Parker[62] who carried out similar measurements at about the same time using platinum and tantalum targets, it was clear that adsorbed

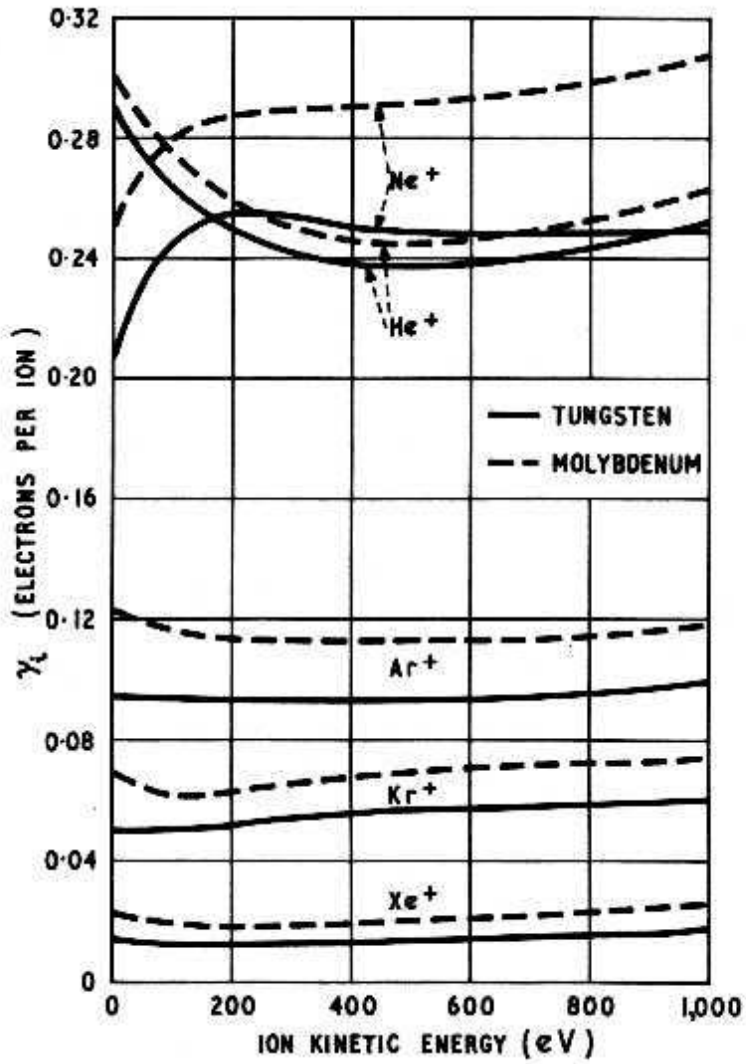


Fig. 1.17. Variation of γ_i with ion energy for the inert gas ions at molybdenum and tungsten targets (Hagstrum[61])

gas layers on the metal surface had a pronounced effect on γ_i, especially at low ion energies. Hagstrum[63] studied this effect in some detail, and showed that as the cathode was exposed to an active gas such as nitrogen N_2, hydrogen H_2 or carbon monoxide CO the value of γ_i decreased steadily with time as a monolayer formed, to level off at a value characteristic of the covered surface (See Fig. 1.18). The value for a covered surface depended on the ion energy, increasing with ion energy towards the value characteristic of the clean metal (See Fig. 1.19). The behaviour resulting in Fig. 1.18 is partly explained by the change in Φ of the metal surface due to the adsorbed layer; nitrogen and carbon monoxide

raise the work function of a metal and therefore fewer electrons escape. At higher ion energies the ions penetrate the gas layer to give the γ_i more characteristic of the pure metal. However, Hagstrum calculated that the expected change in Φ was not sufficient in itself to explain the change in γ_i and suggested that the adsorbed

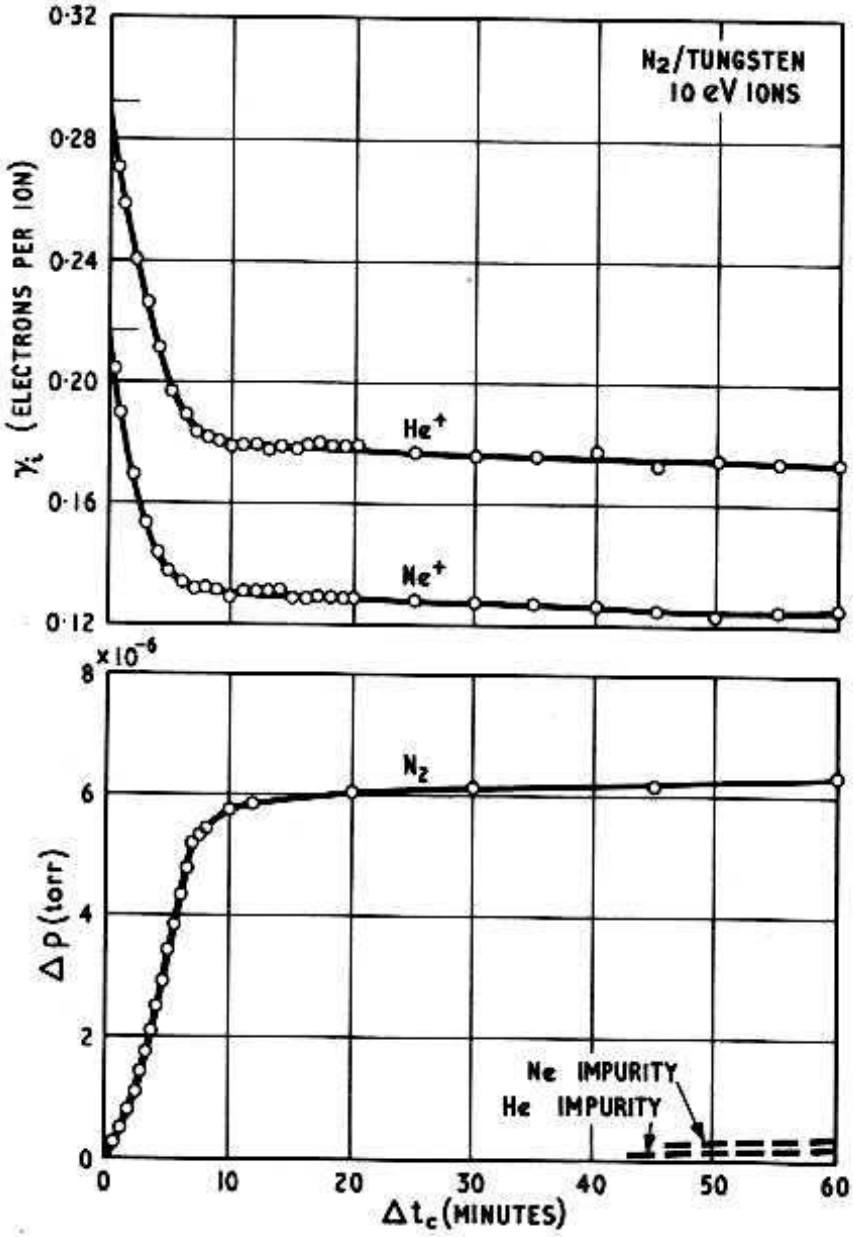


Fig. 1.18. Plots of γ_i for He^+ and Ne^+ ions at a tungsten target and the pressure rise on target flash as a function of the time as a monolayer of N_2 forms on the tungsten. The initial values give the γ_i for clean tungsten measured where N_2 was not admitted into the measuring apparatus (Hagstrum[63])

layer also changes the variation of atomic energy levels near the metal surface, and enhances the probability of electron ejection by non-Auger processes.

A similar decrease in γ_i for contamination of molybdenum cathodes by oxygen has been obtained by the Author[64] using a simple ion gauge comparison method. The method, described by Hall and Weston,[65] consists of comparing the collector current in

one ion gauge having a reference surface with that of a second ion gauge having a collector of the surface under investigation. The collector current is given by the ion current, i, plus the secondary electron current and thus the ratio R is given by

$$R = \frac{i_x(1+\gamma_{ix})}{i_R(1+\gamma_{iR})} \tag{1.18}$$

The suffixes x and R referring to the unknown and reference surface respectively. The method can be applied either when the unknown surface can be formed in the ion gauge, in which case i_x/i_R can be determined with a reference substrate before depositing

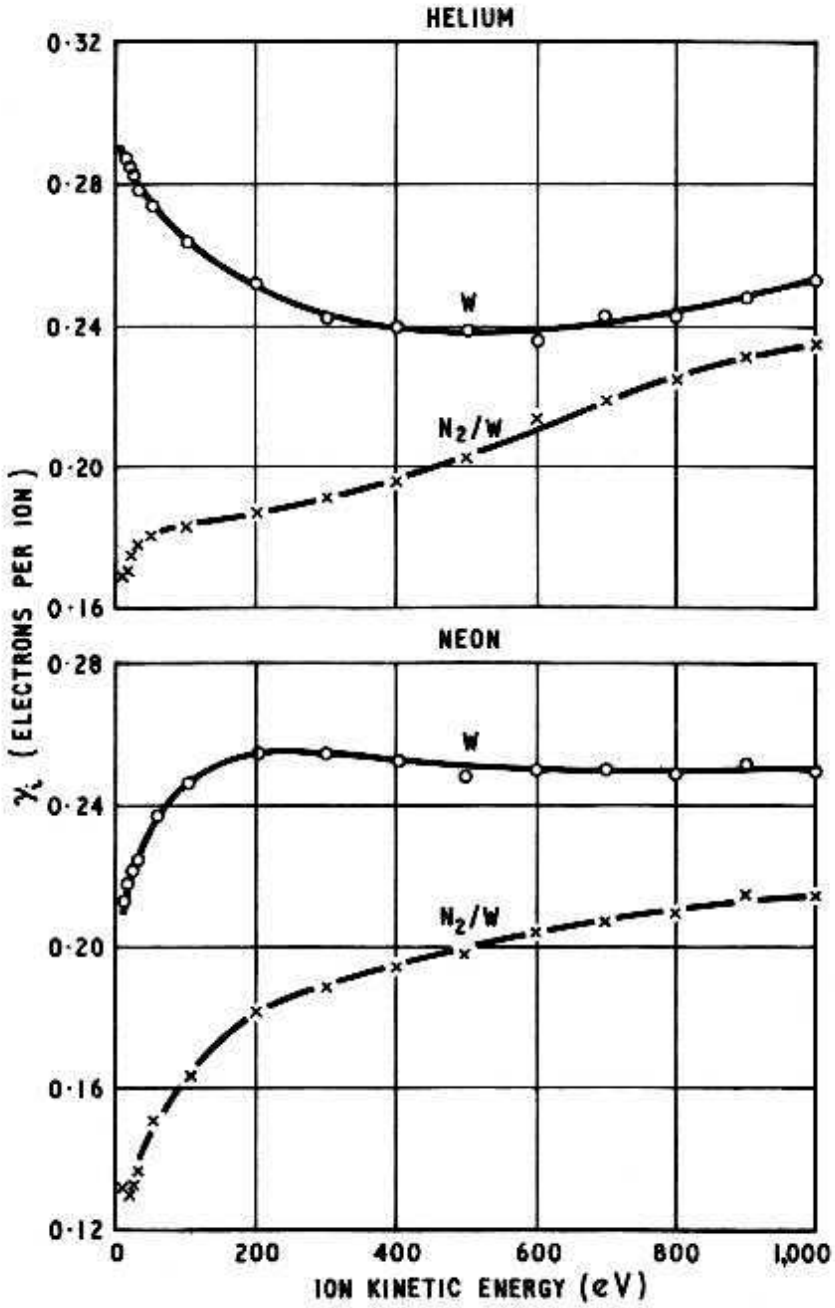


Fig. 1.19. Varation of γ_i with ion energy for He^+ *and* Ne^+ *ions at tungsten targets either clean* (W) *or covered with a monolayer of* N_2 (N_2/W) *(Hagstrum*[63]*)*

the cathode, or for the case of a cathode surface which changes with time. In the latter case the change in the secondary coefficient $\delta\gamma_i$ is given by

$$\delta\gamma_i = \frac{(R_t - R_o)(1+\gamma_i)}{R_o} \tag{1.19}$$

Where γ_i is the value of the secondary coefficient corresponding to the initial ratio R_o, and R_t is the ratio at time t after the surface has changed. Thus if γ_i of a pure metal is known, the change on admitting oxygen can be monitored. Of interest arising from this work was the rather surprising behaviour of nickel to oxygen contamination, which showed that as oxygen diffused in, γ_i first decreased, then increased above the value for clean nickel, and finally decreased to the value for fully oxidised nickel, Fig. 1.20.

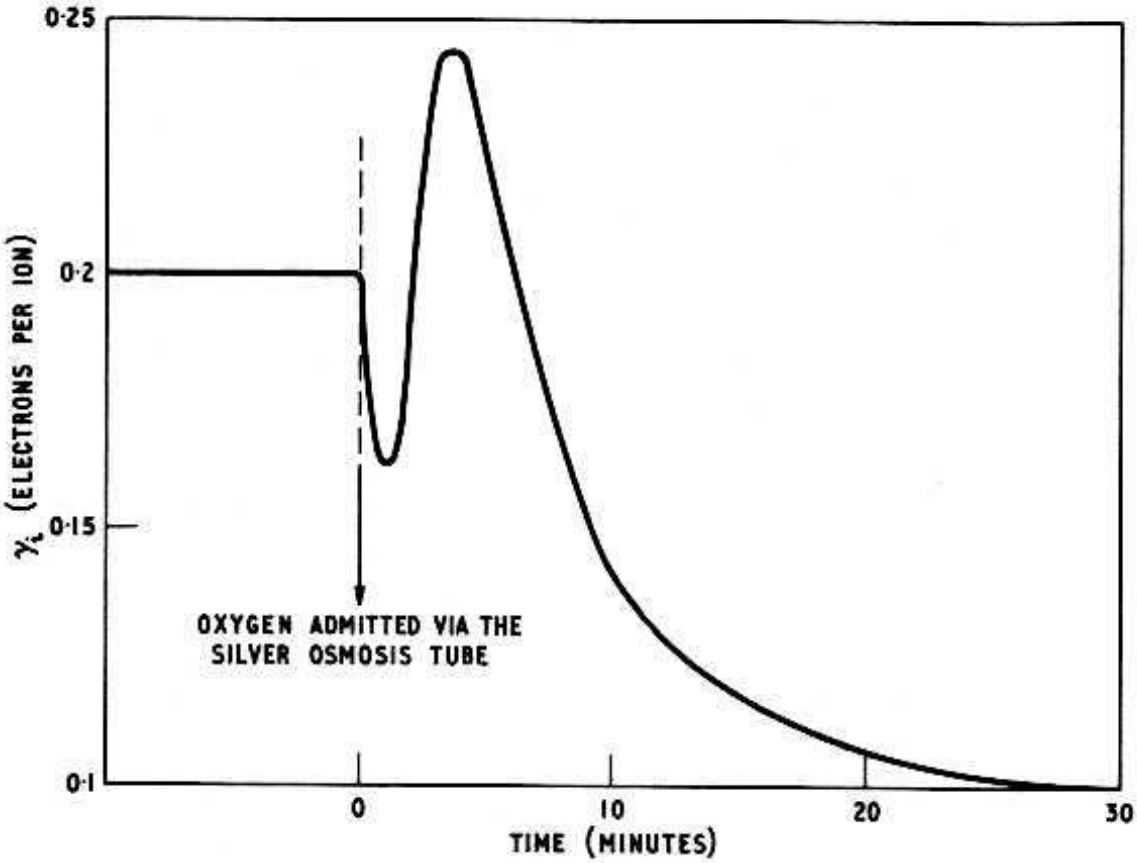


Fig. 1.20. Change in γ_i for Ne^+ ions at a nickel surface as oxygen is admitted. Measured by an ion gauge method

A study of oxygen sorption on nickel single crystals by Farnsworth and Madden[66] revealed a similar pattern of behaviour for the photoelectric work function, which was associated with a place exchange between some oxygen and nickel atoms preceding the formation of the oxide structure. It is likely that this explanation also applied to the γ_i variations.

From Hagstrum's theory γ_i is dependent on the product of the

energy distribution of the excited electrons within the metal, and of the probability of escape of these electrons. Calculated curves of energy distribution N_i and escape probability P_e as a function of energy E_k for tungsten are shown in Fig. 1.21. The distribution moves to higher energies for the ions of higher potential energy,

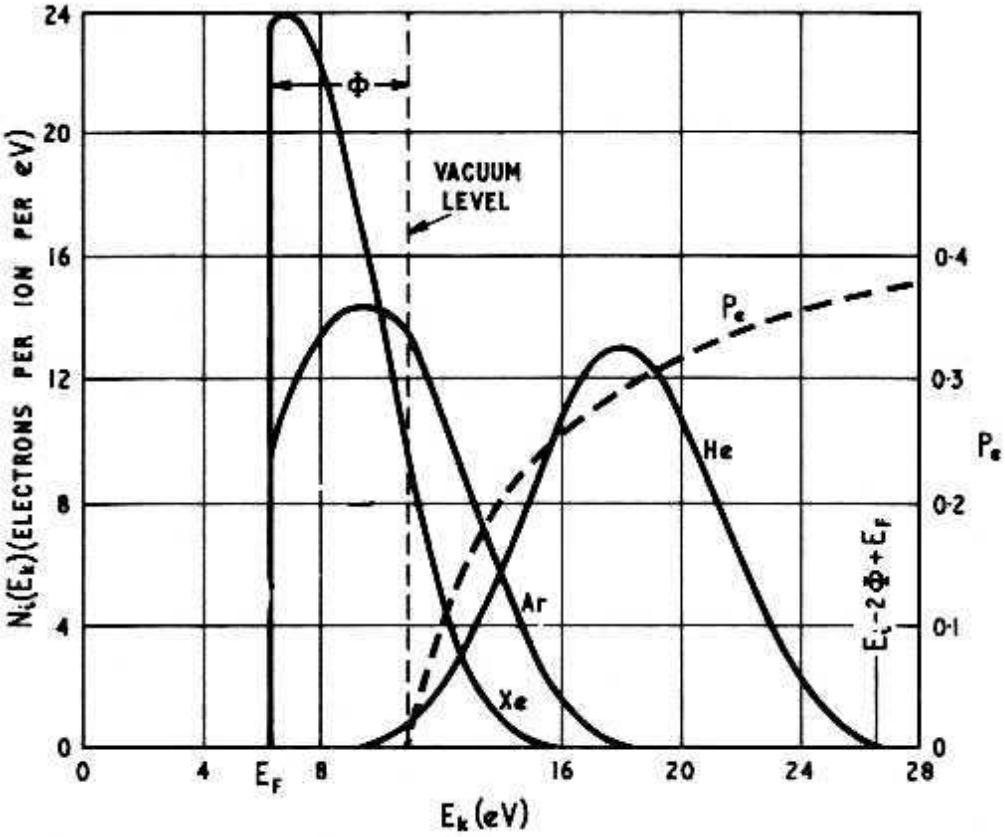


Fig. 1.21. Distribution of energy within the metal of electrons excited in the Auger neutralisation of He^+, Ne^+ *and* Xe^+ *ions at a tungsten surface. The energy levels are as shown in Fig. 1.16. The probability of escape of these electrons is given by the dashed curve* P_e

and the probability of escape depends on the work function. By combining the two appropriate curves of N_i and P_e the number of electrons emitted at each energy can be deduced, and a curve of energy distribution of emitted electrons plotted. γ_i will be given by the area under the resulting plot. Thus there is no simple theoretical relation between γ_i and either the work function of the metal Φ, or the potential energy of the incident ion E_i. Nevertheless, if the values of γ_i for pure metals bombarded by singly-ionised inert gas atoms are plotted in terms of these parameters reasonably smooth curves are obtained. This is demonstrated in Figs. 1.22 and 1.23 respectively where Hagstrum's and Parker's values of γ_i are plotted as a function of E_i for a tungsten surface, and as a function of Φ for argon ions. The value for tantalum does not lie on the curve, which might be expected in view of the likelihood of contamination. It is not suggested that the linear relations are a

universal law, but it is useful for obtaining approximate values for metals and ions not as yet investigated.

The ejection of electrons from a metal surface by metastable atoms is a similar process, i.e. depending on the potential energy of the metastable. In this case the de-exciting energy is transferred

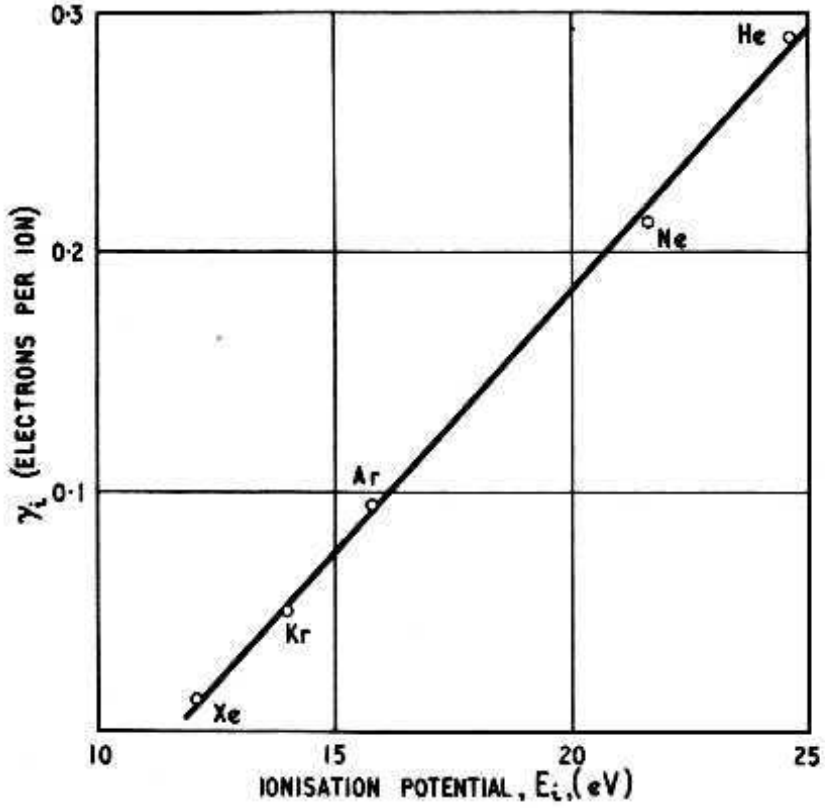


Fig. 1.22. Plot of γ_i values at a tungsten surface as a function of the potential energy of the ion (taken from Table 1.10)

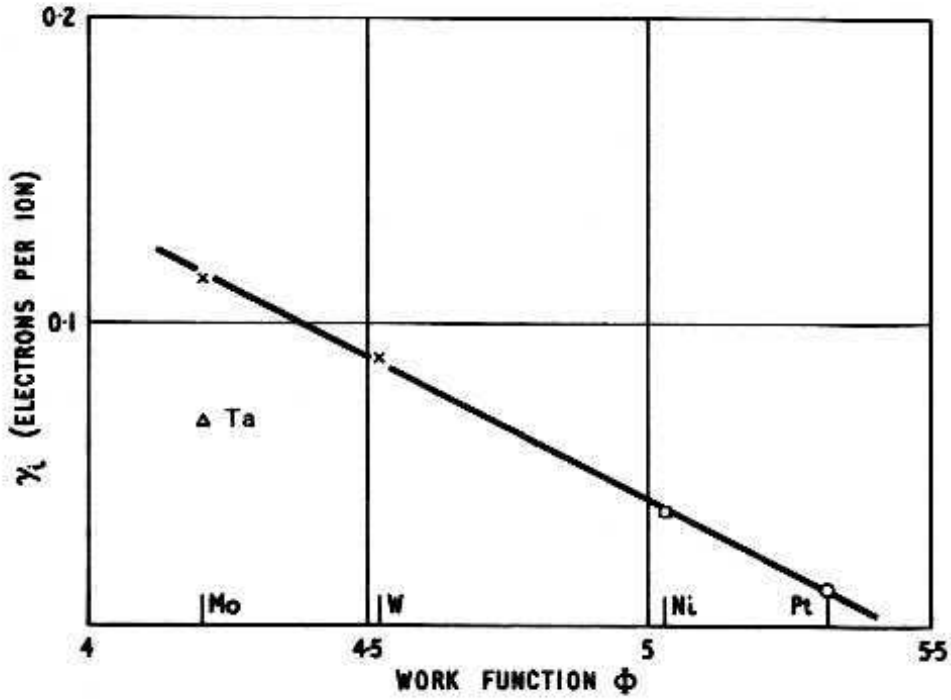


Fig. 1.23. Plot of γ_i values for A^+ ions at various metal surfaces as a function of the work function of the metal. Values taken from the literature of Hagstrum[60, 61] and Parker[62]

to an electron in the conduction band raising it above the work function barrier. Since the metastable atoms are not charged, it is difficult to produce a beam of metastables, and although the emission of electrons by metastables has been observed no values of γ_m by direct measurement have been obtained.

REFERENCES

1. TOWNSEND, J. S., *Phil. Mag.*, **1**, 198, (1901) and *Electricity in Gases*, Oxford University Press (1914)
2. KRUITHOF, A. A. and DRUYVESTEYN, M. J., *Physica*, **4**, 450, (1937)
3. LOEB, L. B., *Basic Processes of Gaseous Electronics*, University of California Press, U.S.A., (1955)
4. VON ENGEL, A. and STEENBECK, M., *Electrische Gasentladungen*, Springer, Berlin, (1932)
5. DRUYVESTEYN, M. J. and PENNING, F. M., *Rev. Mod. Phys.*, **12**, 102, (1940)
6. TAKEISHI, Y., *J. Phys. Soc., Japan*, **11**, 676, (1956)
7. EMELEUS, K. G., LUNT, R. W. and MEEK, C. A., *Proc. Roy. Soc.*, **A156**, 394, (1936)
8. TATE, J. L. and SMITH, P. T., *Phys. Rev.*, **39**, 270, (1932)
9. DEAS, H. D. and EMELEUS, K. G., *Phil. Mag.*, **40**, 460, (1949)
10. DRUYVESTEYN, M. J., *Physica*, **3**, 65, (1936)
11. WARD, A. L., *Phys. Rev.*, **112**, 1852, (1958) and *J. Appl. Phys.*, **33**, 2789, (1962)
12. PENNING, F. M., *Phil. Mag.*, **11**, 961, (1931)
13. AYRES, T. L. R., *Phil. Mag.*, **45**, 353, (1923)
14. TOWNSEND, J. S. and MCCALLUM, S. P., *Phil. Mag.*, **6**, 857, (1928)
15. TOWNSEND, J. S. and MCCALLUM, S. P., *Phil. Mag.*, **17**, 678, (1934)
16. PAAVOLA, M., *Arch. Elektrotech*, **22**, 443, (1929)
17. MASCH, K., *Arch. Elektrotech*, **26**, 589, (1932)
18. SANDERS, F. H., *Phys. Rev.*, **41**, 667, (1932)
19. SANDERS, F. H., *Phys. Rev.*, **44**, 1020, (1933)
20. POSIN, D. O., *Phys. Rev.*, **50**, 650, (1936)
21. BOWLES, W. E., *Phys. Rev.*, **53**, 293, (1938)
22. PENNING, F. M. and ADDINK, C. C. J., *Physica*, **1**, 1007, (1934)
23. KRUITHOF, A. A. and PENNING, F. M., *Physica*, **3**, 515, (1936)
24. KRUITHOF, A. A. and PENNING, F. M., *Physica*, **4**, 430, (1937)
25. KRUITHOF, A. A., *Physica*, **7**, 519, (1940)
26. HUXFORD, W. S., *Phys. Rev.*, **55**, 754, (1939)
27. ENGSTROM, R. W. and HUXFORD, W. S., *Phys. Rev.*, **58**, 67, (1940)
28. DAVIES, D. E. and MILNE, J. G. C., *Brit. J. Appl. Phys.*, **10**, 301, (1959)
29. GOLDEN, D. E. and FISHER, L. H., *Phys. Rev.*, **123**, 1079, (1961)
30. CHANIN, L. M. and RORK, G. D., *Phys. Rev.*, **132**, 2547, (1963)
31. RIESZ, R. and DIEKE, G. H., *J. Appl. Phys.*, **25**, 196, (1954)
32. SMITH, J., *Proc. Intern. Conf. Ionization Phenomena Gases*, 5th, *Munich*, (1961), North Holland Publish. Co. 449
33. FLETCHER, J. and DAVIES, D. E., *Proc. Intern. Conf. Ionization Phenomena Gases, 6th, Paris*, (1963), **2**, 217
34. DAVIES, D. K., LLEWELLYN JONES, F. and MORGAN, C. G., *Proc. Phys. Soc.*, **80**, 898, (1962)
35. CHANIN, L. M. and RORK, G. D., *Phys. Rev.*, **133**, 1005, (1964)
36. HALE, D. H., *Phys. Rev.*, **54**, 241, (1938); and **55**, 815, (1939)

37. ROSE, D. J., *Phys. Rev.*, **104,** 273, (1956)
38. JONES, E. and LLEWELLYN JONES, F., *Proc. Phys. Soc.*, **72,** 363, (1958)
39. HAYDON, S. C. and ROBERTSON, A. G., *Proc. Phys. Soc.*, **78,** 92, (1961)
40. GOLDEN, D. E., NAKANO, H. and FISHER, L. H., *Phys. Rev.*, **138,** A1613, (1965)
41. CHANIN, L. M. and RORK, G. D., *Phys. Rev.*, **135,** A71, (1964)
42. VAN GUGELBERG, H. L., *Helv. Phys. Acta*, **20,** 307, (1947)
43. MOLNAR, J. P., *Phys. Rev.*, **83,** 933, 940, (1951)
44. HORNBECK, J. A., *Phys. Rev.*, **83,** 374, (1951)
45. VARNEY, R. N., *Phys. Rev.*, **93,** 1156, (1954)
46. ROSE, D. J., *Proc. Intern. Conf. Ionization Phenomena Gases, 3rd Venice*, (1957), Societa Italiana di Fisica, Milan, 888
47. LAUER, E. J., *J. Appl. Phys.*, **23,** 300, (1952)
48. COLLI, L. and FACCHINI, U., *Phys. Rev.*, **96,** 1, (1954)
49. DAVIDSON, P. M., *Brit. J. Appl. Phys.*, **4,** 170, (1953), *Phys. Rev.*, **99,** 1072, (1955), *Proc. Roy. Soc.*, **A.249,** 237, (1958), *Proc. Phys. Soc.*, **80,** 143, (1962)
50. DUTTON, J., HAYDEN, S. C. and LLEWELLYN JONES, F., *Brit. J. Appl. Phys.*, **4,** 170, (1953)
51. FLEGLER, E. and RAETHER, H., *Z. Physik.*, **99,** 635, (1936) and **103,** 315, (1936)
52. LOEB, L. B. and MEEK, J. M., *The Mechanism of the Electric Spark*, Standard University Press, (1941)
53. LLEWELLYN JONES, F., *Ionization and Breakdown in Gases*, Methuen, London, (1957)
54. DAVIES, A. J., EVANS, C. J. and LLEWELLYN JONES, F., *Proc. Roy. Soc.*, **A281,** 164, (1964)
55. OLIPHANT, M. L. E., *Proc. Roy. Soc.*, **A127,** 373, (1930)
56. OLIPHANT, M. L. E. and MOON, P. B., *Proc. Roy. Soc.*, **A127,** 388, (1930)
57. HAGSTRUM, H. D., *Phys. Rev.*, **96,** 336, (1954)
58. SHEKHTER, S. S., *J. Exp. Theor. Phys. U.S.S.R.*, **7,** 750, (1937)
59. HAGSTRUM, H. D., *Phys. Rev.*, **91,** 543, (1953)
60. HAGSTRUM, H. D., *Phys. Rev.*, **89,** 244, (1953), *Phys. Rev.*, **104,** 672, (1956)
61. HAGSTRUM, H. D., *Phys. Rev.*, **96,** 325, (1954), *Phys. Rev.*, **104,** 317, (1956)
62. PARKER, J. H., *Phys. Rev.*, **93,** 1148, (1954)
63. HAGSTRUM, H. D., *Phys. Rev.*, **104,** 1516, (1956)
64. WESTON, G. F., *Proc. Intern. Conf. Ionization Phenomena Gases, 6th, Paris*, (1963) **2,** 83.
65. HALL, R. F. and WESTON, G. F., *J. Sci. Instr.*, **40,** 573, (1963)
66. FARNSWORTH, H. E. and MADDEN, H. H., *Solid Surfaces and the gas solid Interface*, Am. Chem. Soc., Washington, (1961), p. 114

CHAPTER 2

Breakdown at Low Pressures

2.1. Paschen's Law

In Chapter 1 it was explained that the discharge between two parallel-plate electrodes, distance d apart, became self-sustaining at a critical potential V_s known as the breakdown potential. From Equation 1.9, this critical potential occurs when i approaches ∞, i.e. when the denominator becomes zero.

$$\gamma[\exp \alpha(d-d_o)-1] = 1 \tag{2.1}$$

Both γ and α/p are functions of E/p (field divided by the pressure) and therefore Equation 2.1 can be written as

$$f_1\ (V_s/pd)\{\exp\,[pd\, f_2(V_s/pd)]-1\} = 1 \tag{2.2}$$

Where f_1 and f_2 are the respective functions of E/p for γ and α/p, and V_s/d is the field at breakdown. d_o is small and can be omitted for most values of V_s. Thus the breakdown potential can be expressed as a function of the product pd. This agrees with the experimental observation known as Paschen's law, which states that if the interelectrode spacing and the gas pressure in a discharge tube are altered to keep their product constant then there is no change in the breakdown potential. The curve relating V_s to pd is known as the Paschen curve, and can be deduced theoretically if f_1 and f_2 are known. From Chapter 1 it was seen that at low pressures, high E/p, γ does not vary appreciably with E/p (Fig. 1.14), and that an empirical relationship between α and E/p for inert gases is given by the equation

$$\alpha/p = C\exp(-Dp^{\frac{1}{2}}/E^{\frac{1}{2}}) \tag{2.3}$$

Inserting this function in Equation 2.1 gives

$$\gamma\{\exp[C\,pd\exp(-Dp^{\frac{1}{2}}/E^{\frac{1}{2}})]-1\} = 1 \tag{2.4}$$

Which reduces to

$$(V_s/pd)^{\frac{1}{2}} = \frac{D}{K+\log_e pd} \tag{2.5}$$

Where $$K = \log_e\left[\frac{C}{\log_e(1+1/\gamma)}\right]$$

For large values of pd V_s rises nearly linearly with pd since the logarithmic term varies only slowly. At low values of pd, the logarithmic term becomes the controlling factor and V_s rises as pd decreases. Thus a minimum occurs at an intermediate pd value, when $dV_s/d(pd) = 0$

i.e. when $$K+\log_e(pd)_{min} = 2$$

giving $$(pd)_{min} = (7{\cdot}39/C)\log_e(1+1/\gamma) \tag{2.6}$$

and minimum breakdown potential

$$V_{s\,min} = D^2/4(pd)_{min} \tag{2.7}$$

A curve for V_s as a function of pd, which is characteristic of experimental curves for pure gases, has been deduced from Equation 2.5 in Fig. 2.1 for argon with molybdenum electrodes. The values of C and D used are those given by Ward,[1] whilst γ has been taken as 0·025–0·085 from Varney's experiments.[2] The resulting curve compares favourably with experimental values by Frouws[3] also given in Fig. 2.1.

The effect of changes in γ and α due to varying the gas filling or electrode material can best be appreciated by expressing Equation 2.1 in terms of η

$$\gamma\,[\exp\eta(V_s-V_o)-1] = 1$$

or $$V_s = V_o+\frac{1}{\eta}\log_e(1+1/\gamma) \tag{2.8}$$

Thus the breakdown potential is an inverse function of both η and γ, having the lowest values for gases and cathodes with the highest values of η and/or γ. As an example, η and γ are both higher for molybdenum electrodes in neon gas than for the same electrodes in argon at pd values greater than 8 cm torr. As a result the breakdown potentials for the former combination are at least

100V below those for the latter at the same pressure. Because of the logarithmic term, changes in γ with pd have a smaller effect than changes in η and the Paschen curve is almost the inverse of the η versus pd curve, with the minimum occurring very close to the pd value for maximum η. Varying the cathode material, and

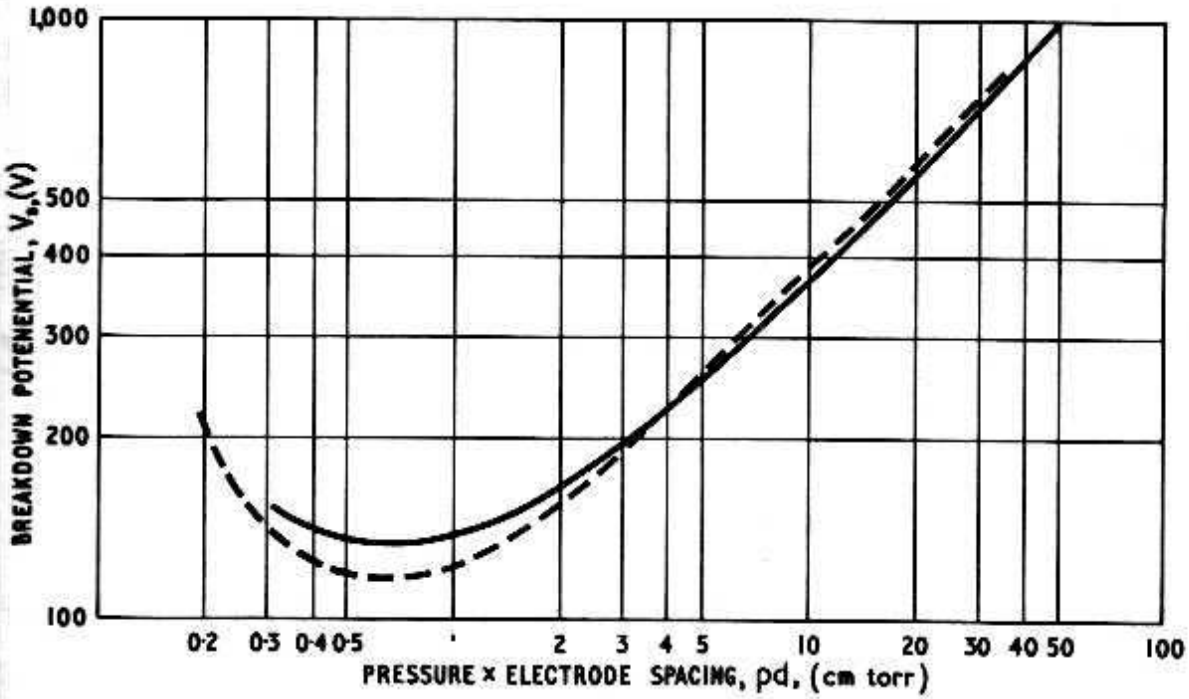


Fig. 2.1. Paschen curves on a logarithmic scale for argon with molybdenum electrodes
– – – Calculated from Equation 2.5
—— Experimental values by Frouws[3]

thus γ, displaces the Paschen curve along the voltage axis with little change in the slope. On the other hand changing the gas, which affects η and γ, shifts the curve along both axes, and can materially alter the general shape. This is particularly true for 'Penning mixtures'. (See Section 2.3).

Paschen curves are a convenient method of presenting breakdown characteristics of discharge tubes, and are particularly useful in the design of cold cathode devices. However it is pointed out that values pertaining to uniform field conditions, for which most of the data exists, cannot be directly applied to electrode geometries resulting in non-uniform field distribution. This is discussed in Section 2.4.

2.2. Uniform-field Breakdown Characteristics of Inert Gases

Paschen curves have been published for most of the inert gases with a variety of cathode materials, often, however, as a subsidiary experiment in the investigation of other discharge parameters such

as the first ionisation coefficient or delay times. Most of the early results refer to massive electrodes of metals such as copper and silver which are unsuitable for commercial tubes, and in some cases of doubtful surface cleanliness. In compiling this section, therefore, results have been chosen which are believed to be reasonably reliable, and some of the earlier results have been omitted.

The breakdown between parallel plates in neon has been investigated by Penning and Addink[4] using iron electrodes, and by Kruithof[5] using copper electrodes, and these are given in Fig. 2.2.

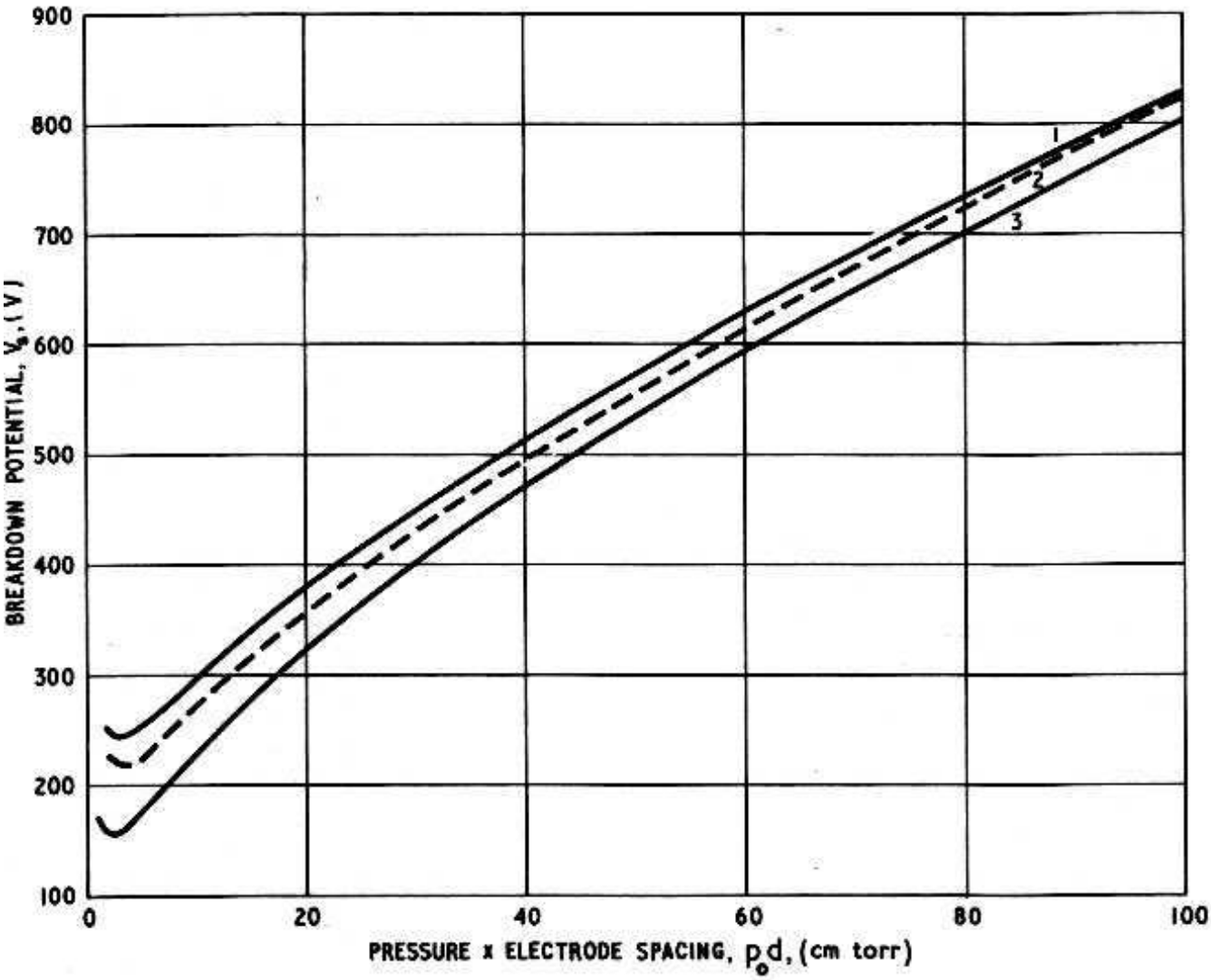


Fig. 2.2. Paschen curves for neon with various cathode materials
(1) Iron cathode[4]
(2) Copper cathode[5]
(3) Molybdenum cathode[3]

In both of these experiments massive cathodes were used, and in view of the variation in γ recorded by Kruithof on processing his cathodes, a shift in the curves along the voltage axis might be expected with different surface conditions. Townsend and McCallum[6] also measured breakdown in neon but the slope of the

Paschen curve they obtained is rather different from other results in the literature, and, in view of the discrepancies between their α/p values and later measurements, it is possible that their neon was contaminated. More recently the author[7] and Frouws[3] have determined breakdown potentials for molybdenum cathodes which had been thoroughly sputtered to obtain a clean surface. The results are in good agreement, and Frouws' curve is reproduced in Fig. 2.2.

Paschen curves for argon have been given by several authors. Of early values of V_s before 1940, the results of Ehrenkranz[8] for platinum cathodes and of Kruithof[5] for copper are given in Fig. 2.3. Ehrenkranz's values closely follow those of Penning and Addink[4]

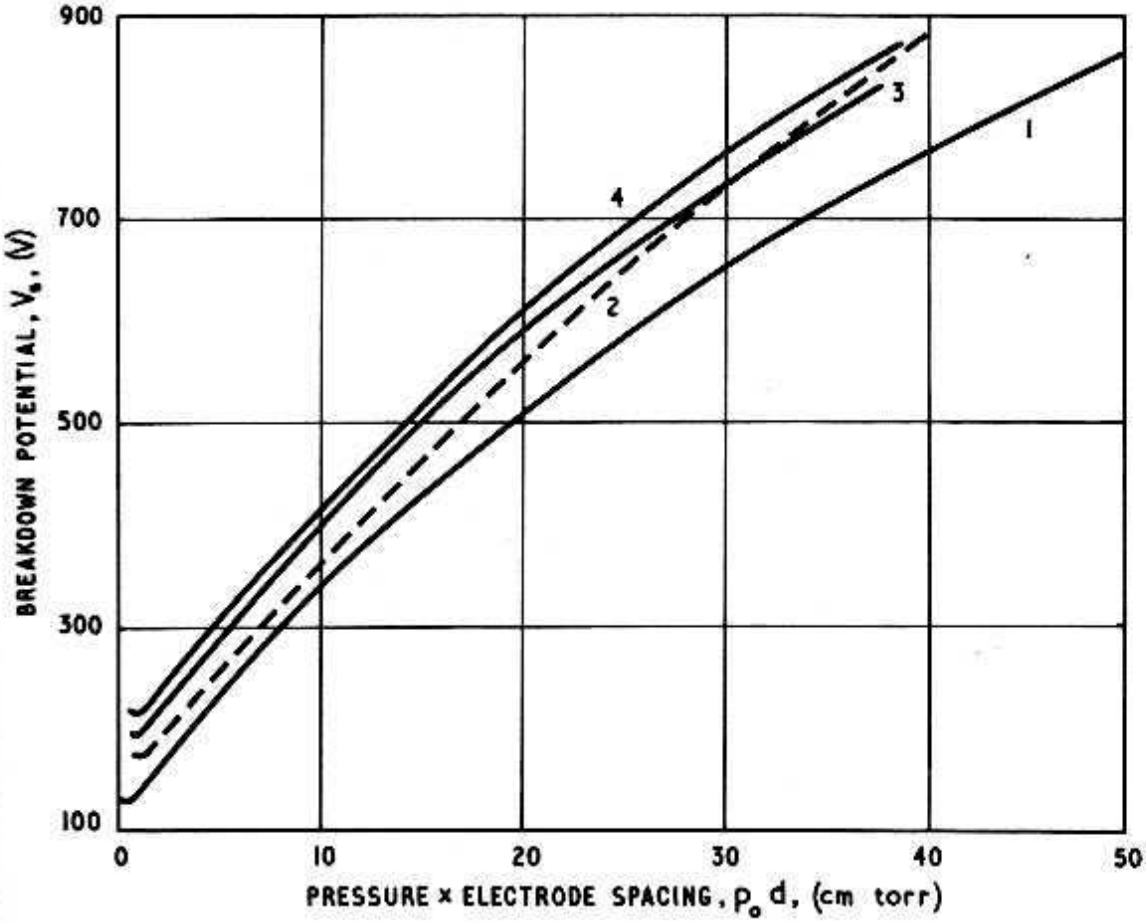


Fig. 2.3. Paschen curves for argon with various cathode materials
(1) Molybdenum cathode[10]
(2) Nickel cathode[9]
(3) Platinum cathode[8]
(4) Copper cathode[5]

using iron cathodes. More recently Heymann[9] has published Paschen curves for a nickel cathode. He pointed out that for small diameter electrodes with low pressures, and therefore relatively

large electrode spacing, higher values of V_s are obtained than with larger diameter cathodes due to diffusion losses. He therefore used cathodes of 6 cm diameter and fairly high pressures. Unfortunately it is more difficult to obtain clean surfaces under these conditions, and above the Paschen minimum the breakdown potential with each electrode as cathode differed appreciably. The V_s values given in Fig. 2.3 for molybdenum were obtained by the author[10] using sputtering techniques for cleaning the cathode. They are

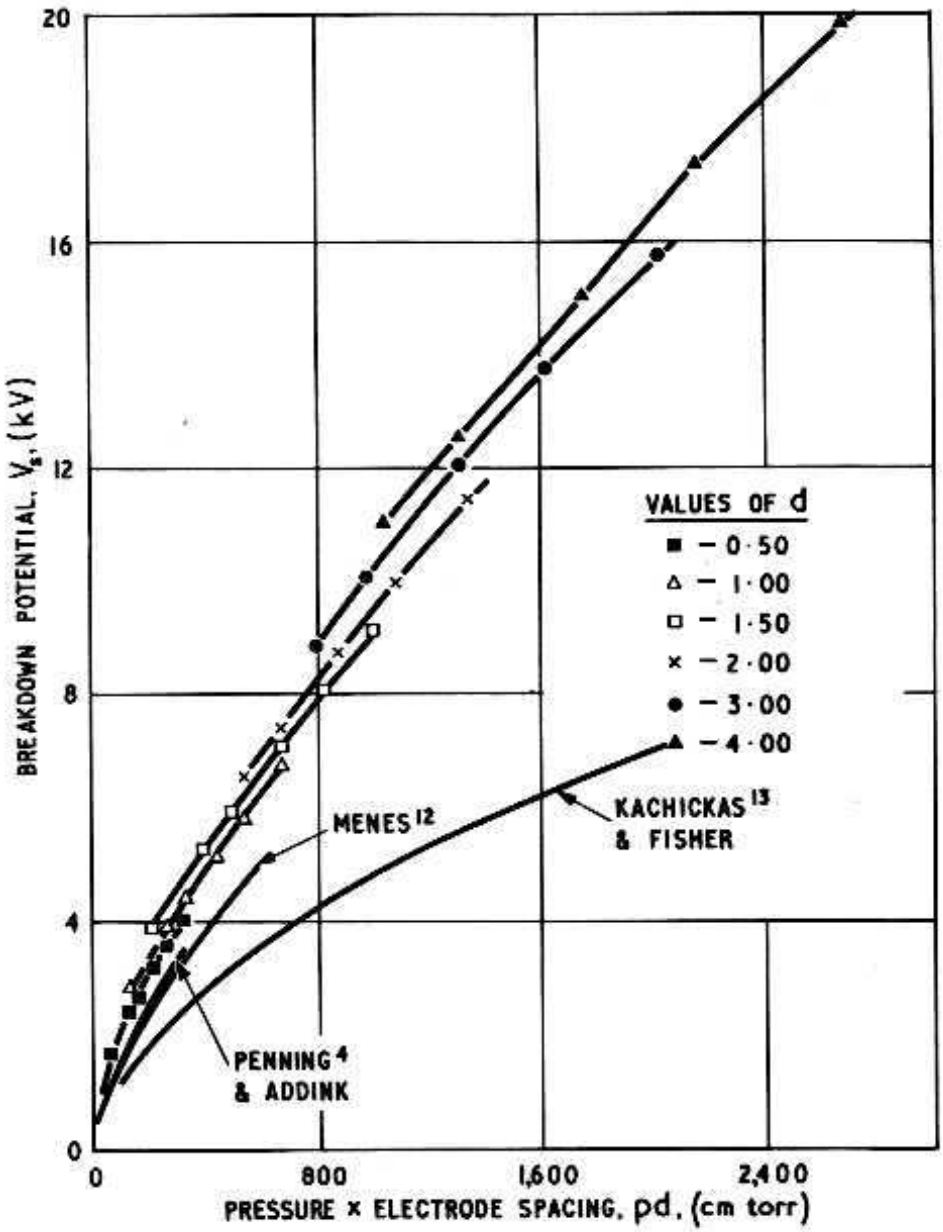


Fig. 2.4. Paschen curves for argon showing the discrepancy at higher pd values between the published data, and the deviation from Paschen's law observed by Golden and Fisher[11]

slightly lower than values given by Frouws in similar experiments. At higher pressures where γ effects should be less evident there are wide discrepancies between published V_s values. This is shown in Fig. 2.4, taken from a paper by Golden and Fisher.[11] The discrepancies must be due to impurities in the gas, which in most

cases was of commercial grade. Golden and Fisher stated that they needed to run a discharge of 1 μA for a few days in order to stabilise the potentials, and that during the stabilising period the breakdown potential rose with time in a spectacular manner. Kachickas and Fisher[13] did not run their tubes before taking measurements. Golden and Fisher also reported a deviation from Paschen's law in that the breakdown potential at a given *pd* increased with increasing electrode spacing. They ascribed the effect to the dependence of α/p on d which has been discussed in Chapter 1.

For helium, the author[7] has measured breakdown values over a limited range with molybdenum cathodes. Since it is difficult to sputter the cathode in helium, the results are less reliable than for neon or argon. They are given in Fig. 2.5, together with the earlier

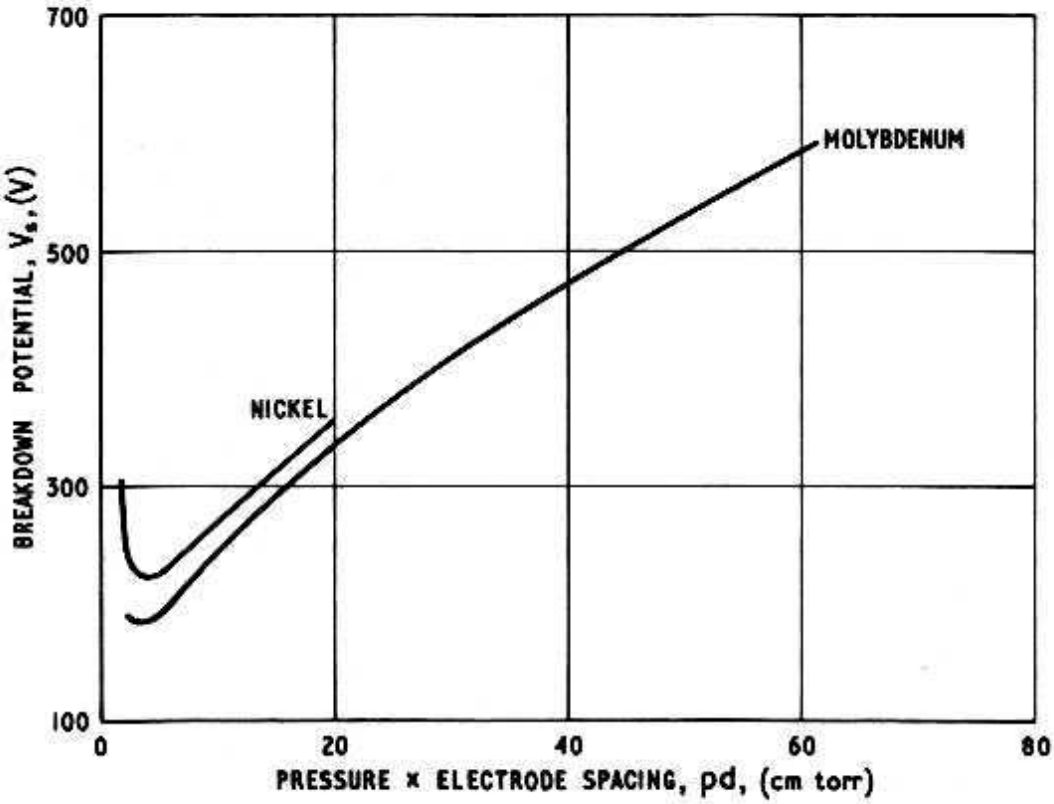


Fig. 2.5. Paschen curves for helium with nickel[14] and molybdenum[7] cathodes

values of Townsend and McCallum[14] for nickel cathodes. The measurements were made at room temperatures and are not corrected to 0°C.

Apart from the above data there has also been some published results over a much more limited range of *pd*. Jacobs and LaRocque[15] have measured breakdown values in the inert gases near the Paschen minimum for cathodes of barium, magnesium and aluminium formed by evaporation on to a nickel substrate. Measurements below the *pd* minimum have been made by Quinn[16] with nickel electrodes in helium and in the active gases hydrogen,

carbon dioxide and air. Measurements near the Paschen minimum for neon with a composite gold-caesium-oxide cathode have recently been described by Hall and Stocker.[17] The results showed a marked deviation from Paschen's law, which they suggest is due to a dependence of V_s on the field in the gap at breakdown. The minimum values recorded by these workers are given in the next section.

2.2.1. THE MINIMUM BREAKDOWN POTENTIAL

The minimum breakdown potential is particularly interesting, not only because it defines the lower limit for a given gas-cathode combination, but also because the semi-theoretical equations are reduced to simple relations at this point of the curve, i.e. Equations 2.6 and 2.7 apply. Thus it should be possible to estimate $(pd)_{min}$ and $V_{s_{min}}$ readily from knowledge of γ and the constants C and D. However, although the equations give the general variation found in practice of $(pd)_{min}$ and $V_{s_{min}}$ with γ and the ionisation potential of the gas, if the values of C and D for the inert gases, given in Table 1.2, are used in conjunction with measured γ values, quoted in Chapter 1, then the calculated values of $V_{s_{min}}$ are at least 20% lower than are found experimentally (cf. Fig. 2.1).

An alternative method of estimating $V_{s_{min}}$ has been suggested by Acton and Swift[18] on the assumption that at $V_{s_{min}}$ η is a maximum, and that for the pure gases the constant V_o in Equation 2.8 can be taken as the ionisation potential V_i.

$$V_{s_{min}} = V_i + 1/\eta_{max} \log_e(1 + 1/\gamma) \tag{2.9}$$

Substituting γ, η_{max}, and V_i values, however, tends to give breakdown minima which are again appreciably lower than experiment. This appears to be mainly due to the assumption that V_o is equal to V_i at E/p for η_{max}, whereas Kruithof[5] points out that in fact V_o rises markedly with E/p and only at much lower E/p is $V_o = V_i$. If Kruithof's values of η_{max} and V_o are used with Varney's[2] values of γ for molybdenum cathodes then agreement is obtained with experimental results. This is shown in Table 2.1.

Unfortunately, there are insufficient data on either V_o or γ pertaining to a discharge to calculate the minimum breakdown for more than a few gas-cathode combinations, and one has to rely on interpolating from experimentally determined values. Table 2.2 gives most of the experimental minimum breakdown voltages available in the literature.

Table 2.1

Calculated values of $V_{s\min}$ for molybdenum cathode based on Equation 2.8. Assuming $V_{s\min}$ occurs at E/p for η_{max}

Gas	*E/p for* η_{max} (V cm^{-1} $torr^{-1}$)	η_{max}	V_o	γ	$V_{s\min}$ (V) *Calculated*	$V_{s\min}$ (V) *Measured*
Ne	85	0·015	31·5	0·2	151	155
Ar	200	0·022	18	0·085	130	128
Kr	220	0·024	16	0·05	141	—

Table 2.2

Minimum breakdown potential $V_{s\min}$ for inert gases with various cathode surfaces (V)

Cathode	*Helium*	*Neon*	*Argon*	*Krypton*	*Xenon*
Al[15]	185	160	154	131	151
†*Au-Cs-0[17]	62	58			
*Ba[15]	145	130	92	102	82
Cu[5]		218	216	203	255
Fe[4]		244	270		
*Mg[15]	160	150	122	116	120
Mo[7][10]	184	155	128		
*Na[8]			94		
Ni[14][18]	223		174		
Pt[8]			195		
*Cs[19]		64			

*V_s(min) for these surfaces may well depend on the metal substrate upon which they are deposited if the film is very thin.
†These values V_s(min) are for 2 mm electrode spacing—the value was found to vary with d.

In general the minimum breakdown voltage decreases with decreasing ionisation potential of the gas filling, and increases with decreasing γ of the cathode material. Because of this pattern of behaviour attempts have been made to express $V_{s\min}$ as a linear function of the ionisation potential of the gas and of the work function Φ of the cathode.[15, 20, 21] Although the general trend is observed if the values of $V_{s\min}$ given in Table 2.2 are plotted as a function of V_i and Φ, Figs 2.6 and 2.7 respectively, the spread in plotted points makes it impossible to draw any conclusions. In the case of Φ, the values chosen from the literature are somewhat arbitrary since they are dependent on surface conditions, and the work function applying to the metal in the gas is likely to differ from that determined under different conditions in vacuum. Realising this, Davies and his co-workers measured the breakdown

potential and the work function of the cathode in the same envelope. The latter was determined by a Kelvin method, which gives the contact potential difference between the cathode and a reference surface. A gold film on a glass substrate was used as the reference

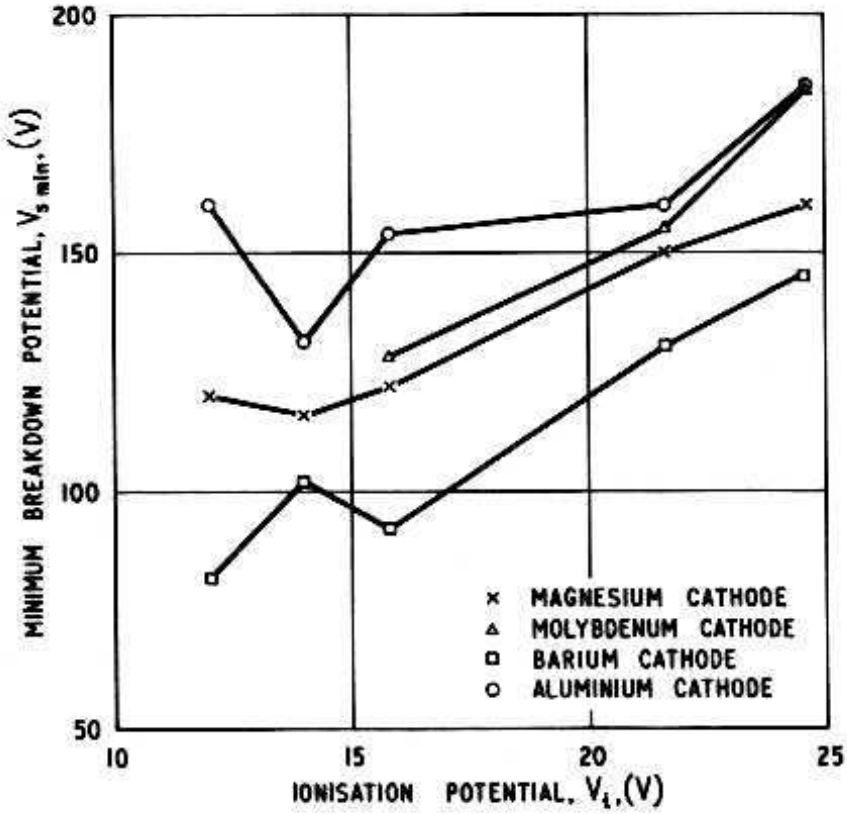


Fig. 2.6. Minimum breakdown potential as a function of the ionisation potential of the gas. $V_{s\min}$ taken from Table 2.2

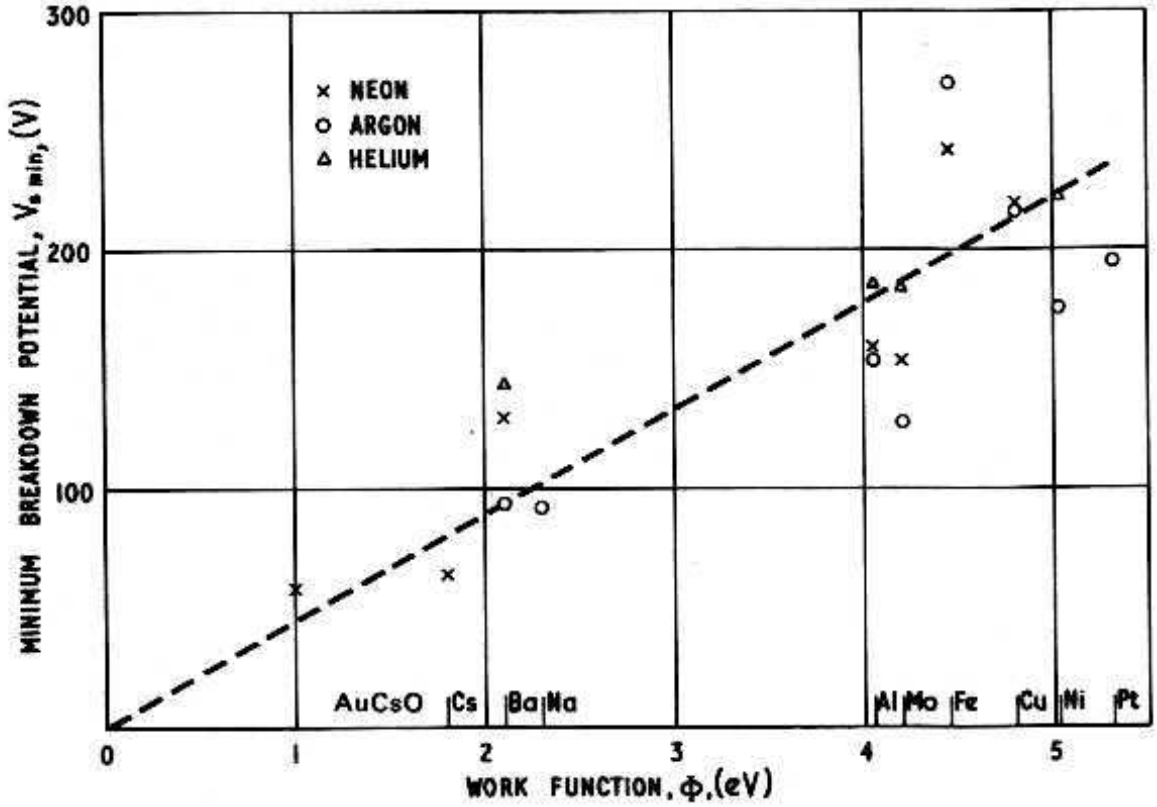


Fig. 2.7. Minimum breakdown potential as a function of the cathode work function. $V_{s\min}$ taken from Table 2.2

surface and the gas filling was hydrogen. In one paper Davies and Fitch[22] monitored Φ and V_s as functions of time for a copper film cathode which they believed was contaminated by oxygen. The change in Φ and V_s on the passage of a small discharge current of 10^{-7} A for 10 sec at times 0 and 50 min is shown in Fig. 2.8. The explanation for the lowering of the work function is probably that the positive ions formed in the discharge deposit on to the oxide layer to produce an electrical dipole layer with the positive charge outermost.

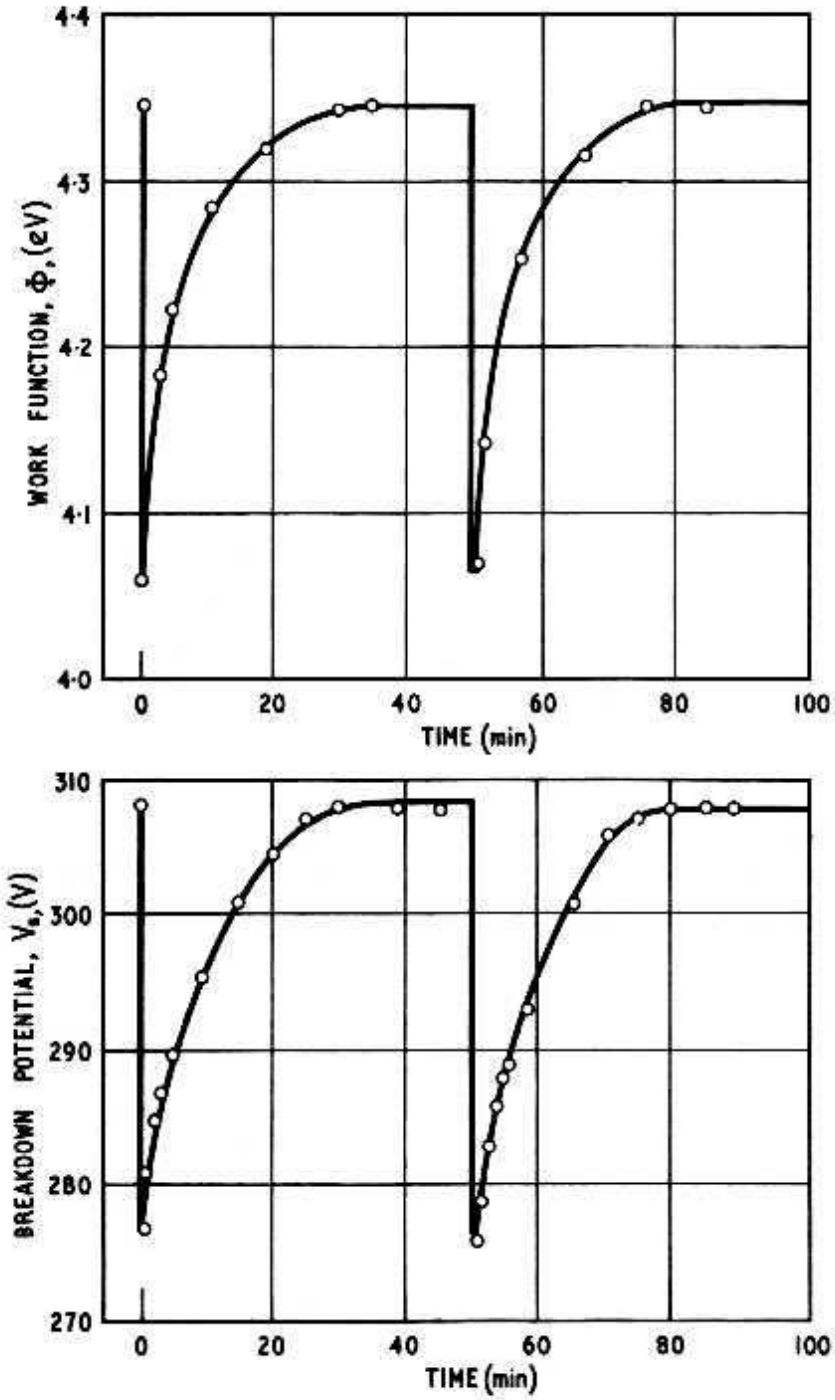


Fig. 2.8. Variation of work function and the breakdown potential in hydrogen for a copper cathode with time after passing 10^{-7} A current for 10 sec at time 0 and 50 min. Davies and Fitch[22]

Davies and Hopkins[23] extended this work using oxide-coated cathodes, the work function of which could be varied from 1–4 eV by treatment of the cathode or the choice of oxide. They also determined the work function of the cathodes by thermionic emission measurements under vacuum conditions. Their results are summarised in Table 2.3 and shown graphically in Fig. 2.9.
The spread about the straight line is considerably less than in Fig. 2.7, but this may be fortuitous particularly as a different slope

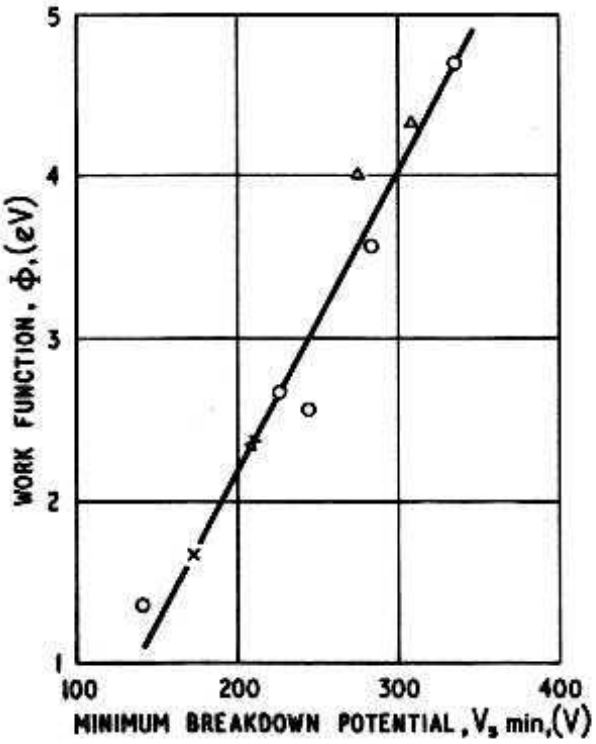


Fig. 2.9. Work function of various cathodes as a function of their minimum breakdown potential in hydrogen. Values taken from Table 2.3. Davies and Hopkins[23]

Table 2.3

$V_{S_{min}}$ and Φ for various cathodes in hydrogen according to Davies and Hopkins[23]

Tube No.	*Cathode*	*Condition*	Φ		$V_{s_{min}}$(V)
			Kelvin Method eV	*Emission Measurements* eV	
2	BaO	Activated	1·4	—	146
5	BaO	Activated	—	1·65	174
5	BaO	Unactivated	—	2·37	212
3	CaO	Activated	2·6	—	246
4	CaO	Activated	2·7	—	230
1	BaO	Unactivated	3·6	—	285
4	Au	—	4·7 assumed	—	335
Davies and	Cu	Oxide layer	4·34	—	308
Fitch	Cu	Oxide layer + H^+ ion layer	4·00	—	275

is obtained from Davies and Fitch's values. From theoretical considerations a simple linear relation is not expected (cf. Section 1.3.2) and one could only hope for an approximate rule which could be applied by the valve engineer.

2.3. Breakdown Potentials for Inert-gas Mixtures

As discussed in Chapter 1, if one gas is added to another, such that the ionisation potential of the added gas is lower than the metastable level of the parent gas (a Penning mixture), then interaction takes place manifest in a large increase in the Townsend first coefficient η. Because of this the breakdown potential is decreased, showing the same sensitivity to small admixtures of the lower ionisation

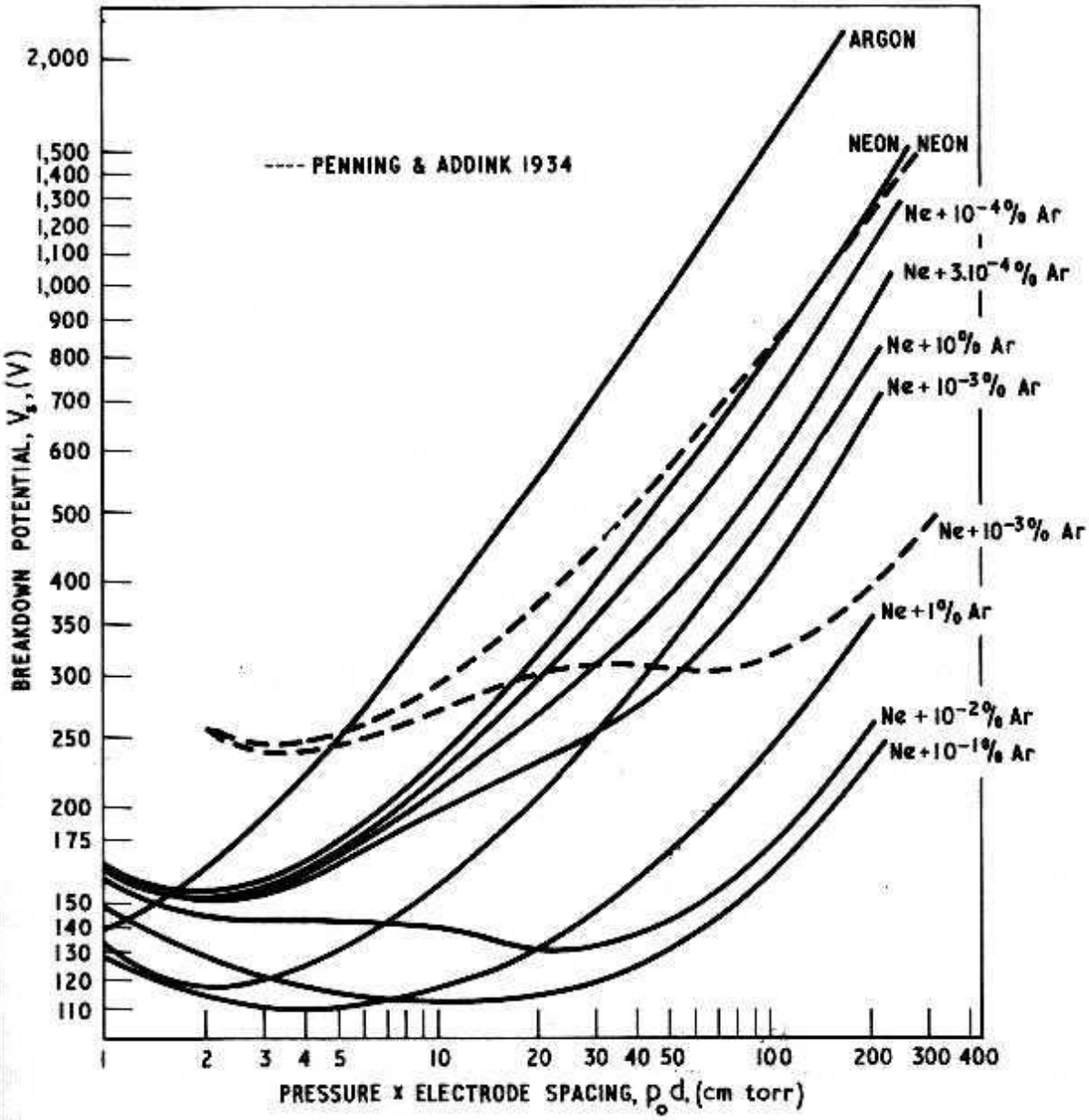


Fig. 2.10. Paschen curves on a logarithmic scale for neon-argon mixtures with molybdenum electrodes. Frouws[3]

potential gas as exhibited by the η values. Also for certain concentrations a double minimum is observed corresponding to the two maxima of η.

The effect on the breakdown potential of the Penning reaction was first demonstrated by Penning and Addink[4] for neon-argon mixtures in a tube having iron electrodes. Their curves for a range of mixtures are reproduced in most text books dealing with gas discharges. Since then Frouws[3] has repeated the measurements using molybdenum electrodes, which is of more direct interest for cold cathode devices. His results are given in Fig. 2.10. He found slight discrepancies between his and Penning and Addink's results which he considered were due to a dependence of the 'Penning effect' on pressure. This is illustrated by a curve of breakdown potential as a function of argon concentration for a given *pd* but with different electrode spacings, Fig. 2.11. A similar dependence

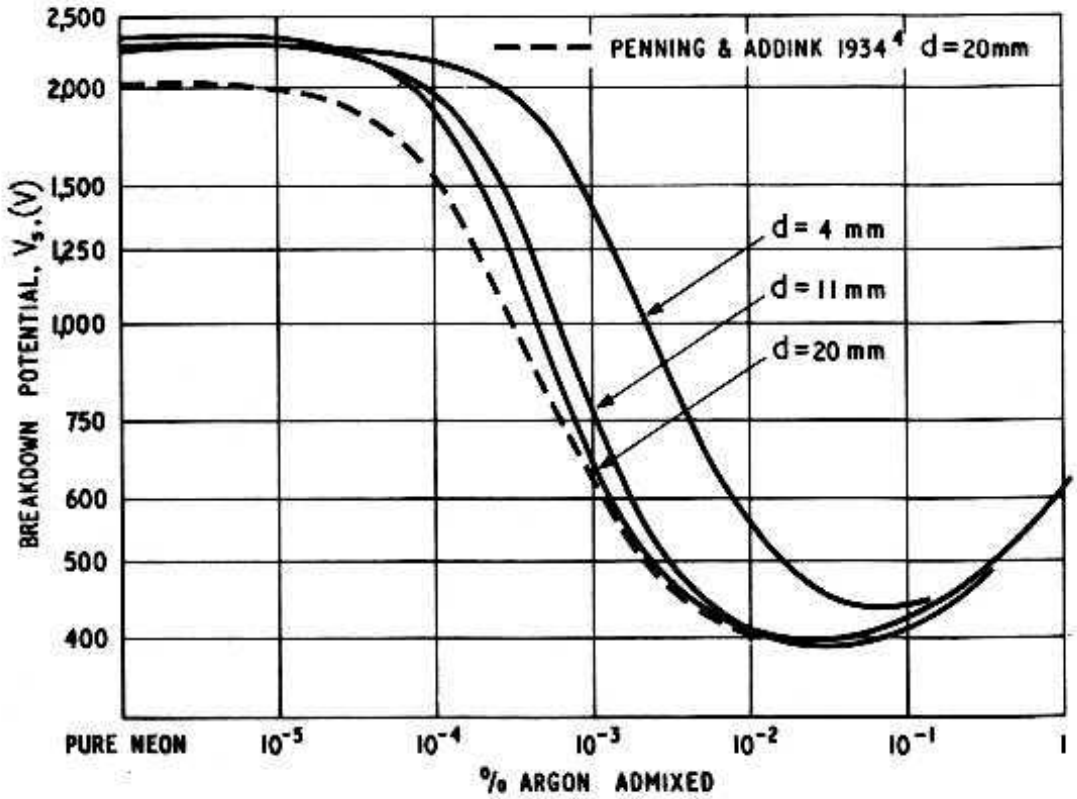


Fig. 2.11. Breakdown potential as a function of the percentage of argon admixed in neon at $p_0d = 450$ cm torr with iron electrodes, showing the departure from the similarity law. Frouws[3]

on pressure was observed by Chanin and Rork[24] who made a more detailed study with helium-hydrogen, neon-hydrogen and neon-argon mixtures. They showed that the effect was caused by a decrease of γ with pressure. The reasons for this were not clear,

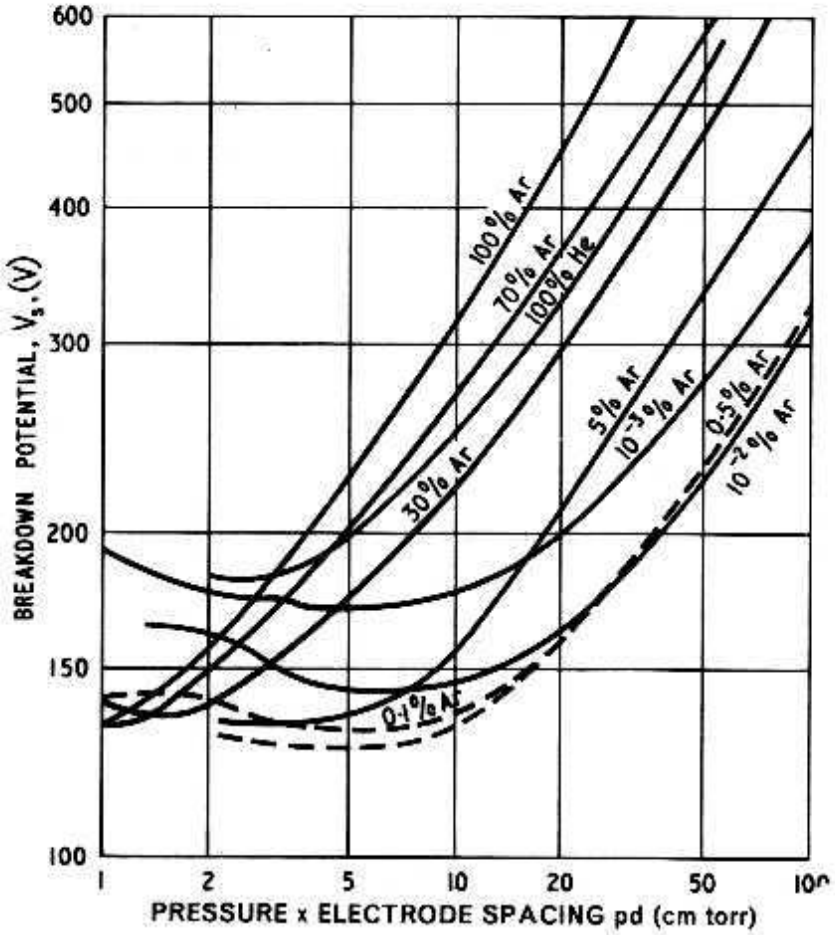


Fig. 2.12. Paschen curves on a logarithmic scale for helium-argon mixtures with molybdenum electrodes

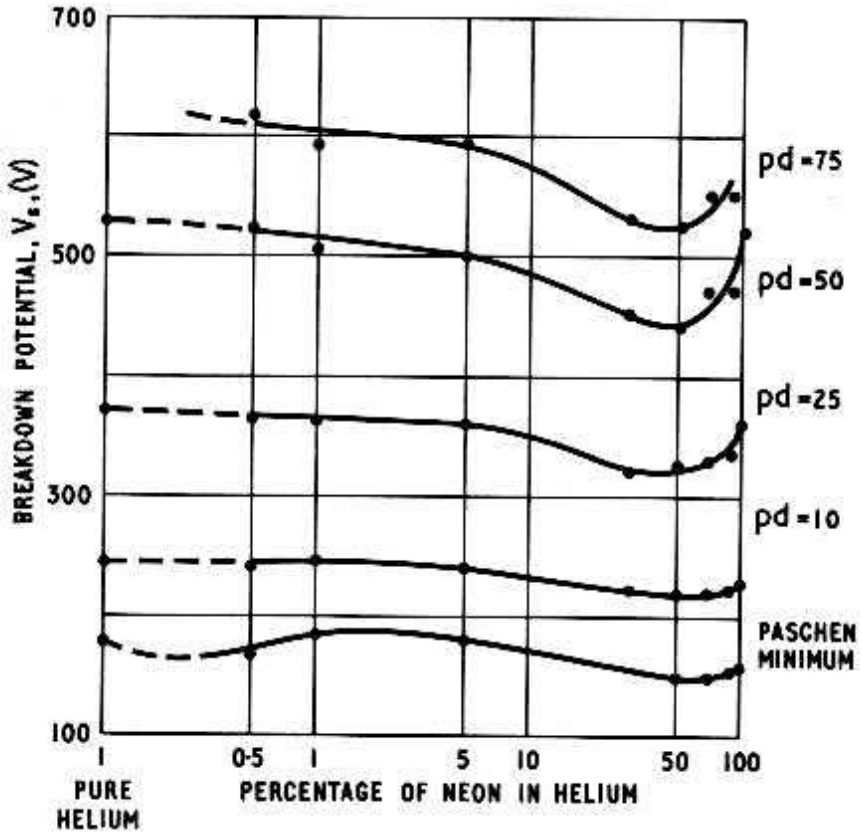


Fig. 2.13. Breakdown potential as a function of the percentage of neon admixed in helium at various pd values

although photon absorption processes seem to be a likely possibility.

The departure from Paschen's law for small concentrations of argon was not observed by the author[10] when measuring breakdown potentials of helium-argon mixtures in tubes with molybdenum electrodes. The Paschen curves obtained for helium-argon are shown in Fig. 2.12.

As far as is known no other Paschen curves for Penning mixtures have been published, although the effect on the breakdown voltage, at a single *pd* value, has been reported by Benson and Burdett[25] for admixtures of krypton and xenon in neon, and also some triple mixtures using electrode geometries which were unfortunately non-uniform. The results, given in Table 2.4, show the expected decrease in V_s for small admixtures in pure neon. The effect is less efficient than for argon in neon, and a smaller reduction in V_s is therefore observed. Adding krypton or xenon to neon + 0·3 % argon decreases the effectiveness of the argon and raises the breakdown potential.

An increase in the breakdown potential above those of the constituent gases has been observed by the author[7] in helium-neon mixtures containing 1–5 % neon, near the Paschen minimum, with a lowering of the potential at higher *pd* values for larger percentages of neon. The effects are small but nevertheless significant, and are shown in Fig. 2.13. Although ionisation of neon by the helium metastable is energetically impossible, two other reac-

Table 2.4

Breakdown potentials for various gas mixtures in stabiliser tubes filled to 40 torr with 1 mm electrode gap, according to Benson and Burdett[25]

Gas	*Breakdown Potential* (V)	*Gas*	*Breakdown Potential* (V)
Neon	220–234	Neon +0·3% Argon	108–115
Neon Plus		Neon + 0·3% Argon plus	
0·001% Kr	191–200	0·001% Kr	118–122
0·01% Kr	142–151	0·01% Kr	112–122
0·1% Kr	121–127	0·1% Kr	114–124
1·0% Kr	138–162	1·0% Kr	136–158
Neon plus		Neon +0·3% Argon plus	
0·001% Xe	194–200	0·001% Xe	120–121
0·01% Xe	128–136	0·01% Xe	114–120
0·1% Xe	136–138	0·1% Xe	134–141
1·0% Xe	150–155	1·0% Xe	140–145

tions in the gas are known.[26] The first is a charge exchange between the helium molecular ion, which is known to be the main ion in pure helium discharges at such pressures, and the neon atom

$$He_2^+ + Ne \rightarrow Ne^+ + 2He$$

The second reaction is the formation of a helium-neon ion by the three-body reaction:

$$Ne^+ + Ne + He \rightarrow (NeHe)^+ + Ne$$

Both these processes will alter the mobility and energy of the positive charge carrier thus at least affecting γ. It is suggested therefore that the variations in V_s are a result of one or other of these reactions.

2.4. Breakdown in Non-uniform Fields

In the previous sections it has been assumed that the applied field controlling the ionisation processes resulting in breakdown has been uniform throughout the interelectrode gap. If, however, due to electrode geometry the field is non-uniform, then η will vary across the gap and electron multiplication will be governed by the integral of η over the path travelled. Further, if the secondary process is that of electron ejection by ions then the value of γ will be controlled by the field at the cathode surface. Thus the breakdown criterion will be given by

$$\int_{V_o}^{V_s} \eta \mathrm{d}V = \log_e \left(1 + \frac{1}{\gamma}\right) \tag{2.10}$$

The simplest form of non-uniform field geometry is that of coaxial cylinders, and this has been examined by several workers (see for example Meek and Craggs).[27] If the inner cylinder is of radius a and the outer cylinder of radius b then Equation 2.10 can be written as

$$\int_a^b \eta E_r \mathrm{d}r = \log_e \left(1 + \frac{1}{\gamma}\right) \tag{2.11}$$

where E_r is the field at radius r, which at breakdown is given by

$$E_r = \frac{V_s}{r \log_e b/a} \tag{2.12}$$

If the value of η at any point in the gap is the same as it would be for a uniform field of the same magnitude, the integral and thus the breakdown potential would be independent of polarity. In practice, however, the breakdown potential with the inner cylinder negative is normally lower than with it positive. This is illustrated in Fig. 2.14 where curves of breakdown potential against

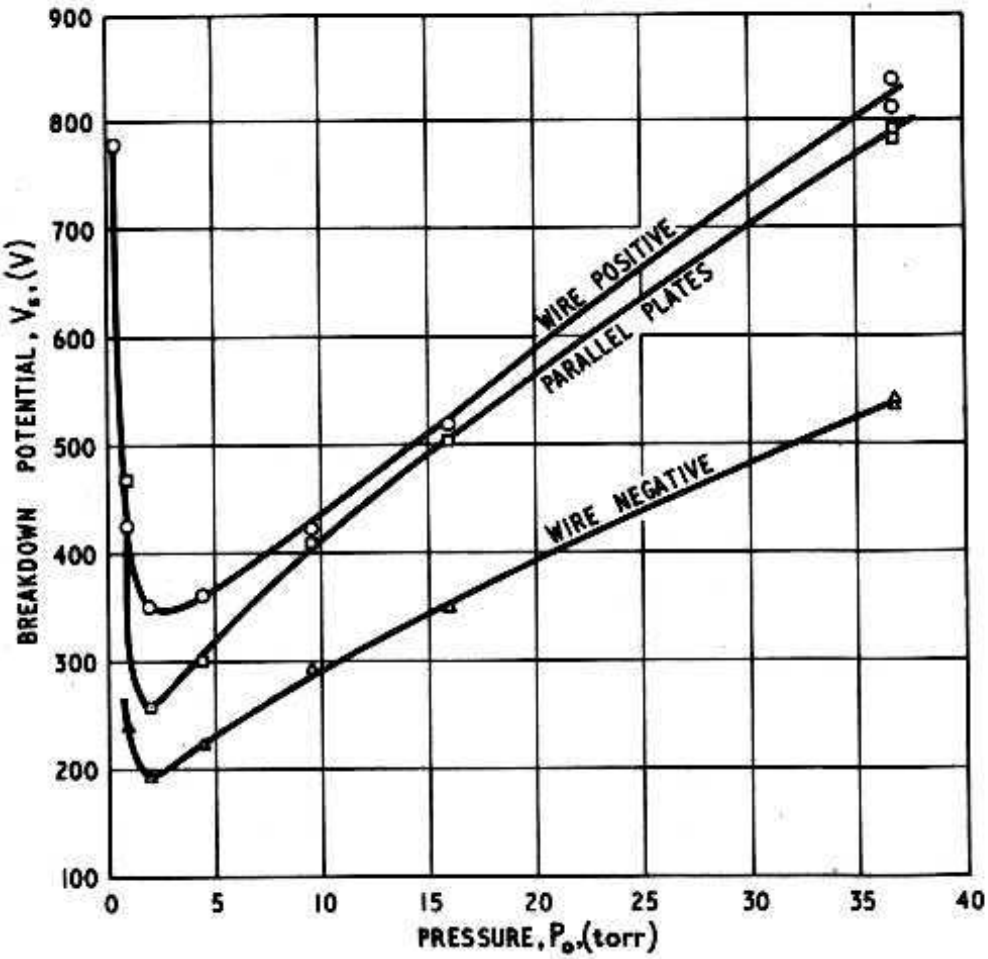


Fig. 2.14. Breakdown potential as a function of helium pressure for coaxial cylindrical geometry (0·175 cm diameter nickel wire mounted in a 4·6 cm diameter iron cylinder), and for parallel plate geometry. Penning[28]

pressure are plotted for the two cases and compared with parallel-plate geometry for helium.[28] The lower value for the inner cylinder negative is attributed to γ being higher because of the greater field at the surface. In fact it has been shown by Morton[29] that η also depends on polarity, particularly when the mean free path is large compared with one electrode dimension, i.e. for fine wire cathodes, or when gas mixtures exhibiting the Penning effect are involved (see later).

For the cylindrical geometry a similarity relation can be obtained if b/a is kept constant:[18]

$$V_s = aE_s \log_e b/a = f(ap_o) \tag{2.13}$$

where E_s is the field at $r = a$ at breakdown. Further, for the case of the inner cylinder negative and $b \gg a$, the field at the outer cylinder can be taken as zero and the similarity relation Equation 2.13 can be applied for any value of b. Thus by plotting V_s or aE_s against (ap_o) universal curves can be drawn relating the breakdown in cylindrical geometries of differing sizes.

Referring to Fig. 2.14 it is seen that for the inner cylinder or wire positive, i.e. the field concentrated round the anode, the values are comparable or slightly higher than for the plane-parallel case over most of the pressure range, whilst for the wire negative they are lower. A similar result is obtained in other non-symmetric geometries, for example with a wire anode and plate cathode the breakdown potential is higher than for the uniform field case with the same electrode separation, whereas it is possible to obtain breakdown potentials lower than for a uniform field with a wire or spike cathode and plate anode.

It is emphasised that for the case of non-uniform field Paschen's law relating breakdown to gap width and pressure no longer holds, and the minimum breakdown potential no longer has a unique value for two given electrodes. This is demonstrated in Fig. 2.15

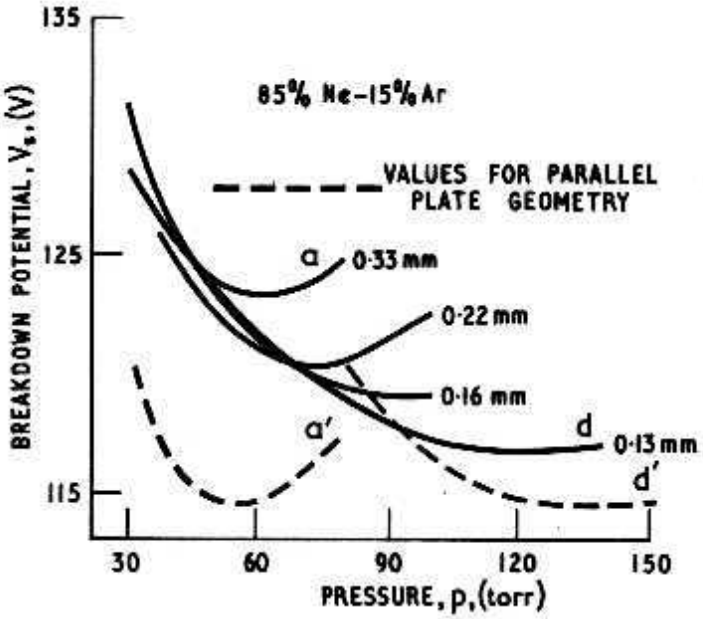


Fig 2.15. Breakdown potential as a function of pressure for a plate cathode and wire anode in the shape of a hook, for various electrode spacings

where the breakdown potential between a plane cathode and a wire anode bent in the form of a hook is plotted as a function of pressure for various interelectrode spacings. The broken lines a' and d' are the curves for parallel-plate geometry corresponding to gaps a and d. As the distance is decreased it is seen that the curve approaches that for a uniform field geometry.

Of interest is the Penning effect in non-uniform geometry. Penning[28] showed that the effect on breakdown of introducing argon into neon in a tube with coaxial cylinder electrode geometry depended on polarity. Breakdown potentials with the central cylinder or wire negative are plotted as a function of pressure in Fig. 2.16, and with the wire positive in Fig. 2.17. The lowering of

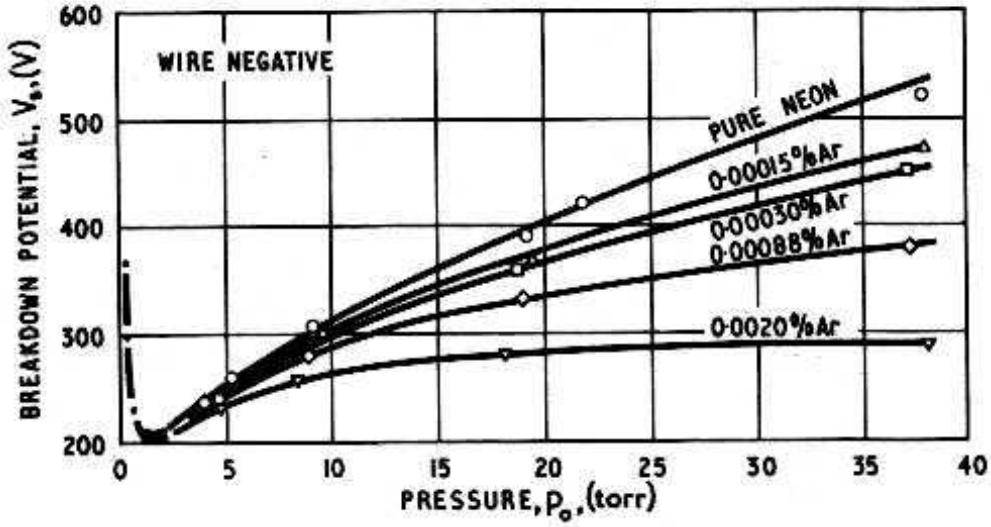


Fig. 2.16. Breakdown potential between coaxial cylinders with neon-argon mixtures. Centre wire negative. Penning[28]

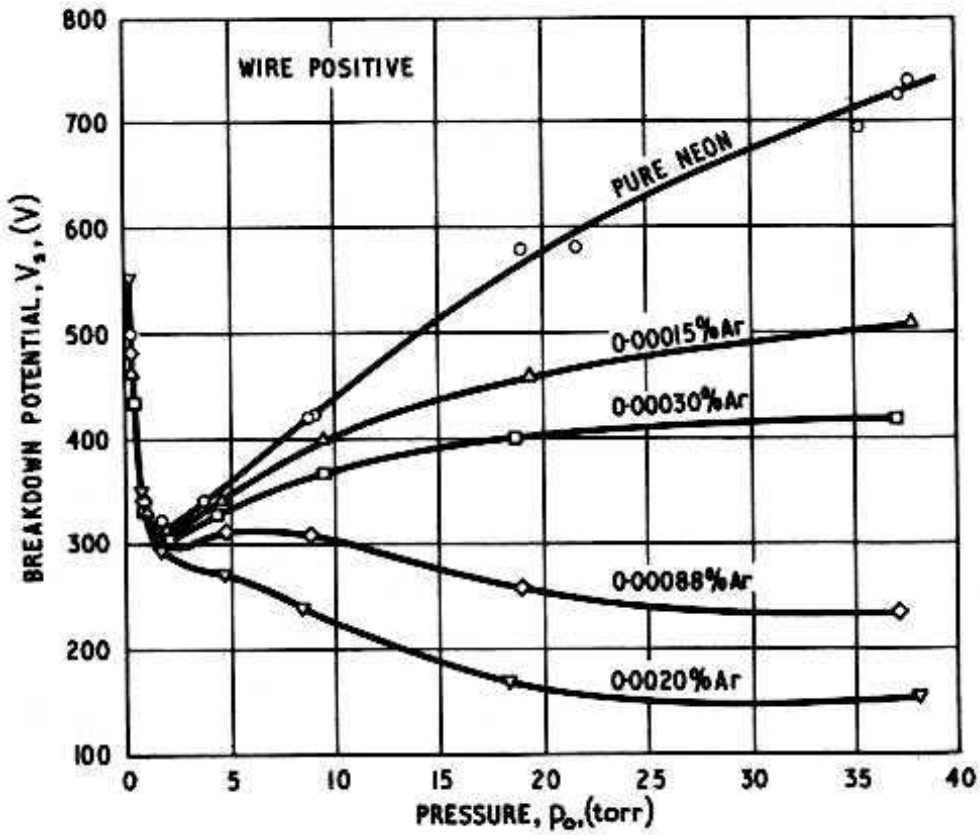


Fig. 2.17. Breakdown potential between coaxial cylinders with neon-argon mixtures. Centre wire positive. Penning[28]

the breakdown potential by small admixtures of argon is greater for the wire positive than for the wire negative and is explained as follows: for the wire positive the ionisation and excitation coefficients increase towards the anode due to increasing E/p, and therefore the majority of the neon metastables will be produced near the anode. These, however, will drift away from the anode area before experiencing ionisation collisions with the argon atoms, and thus the electrons formed by these collisions will be produced in the cathode area and result in considerable multiplication on

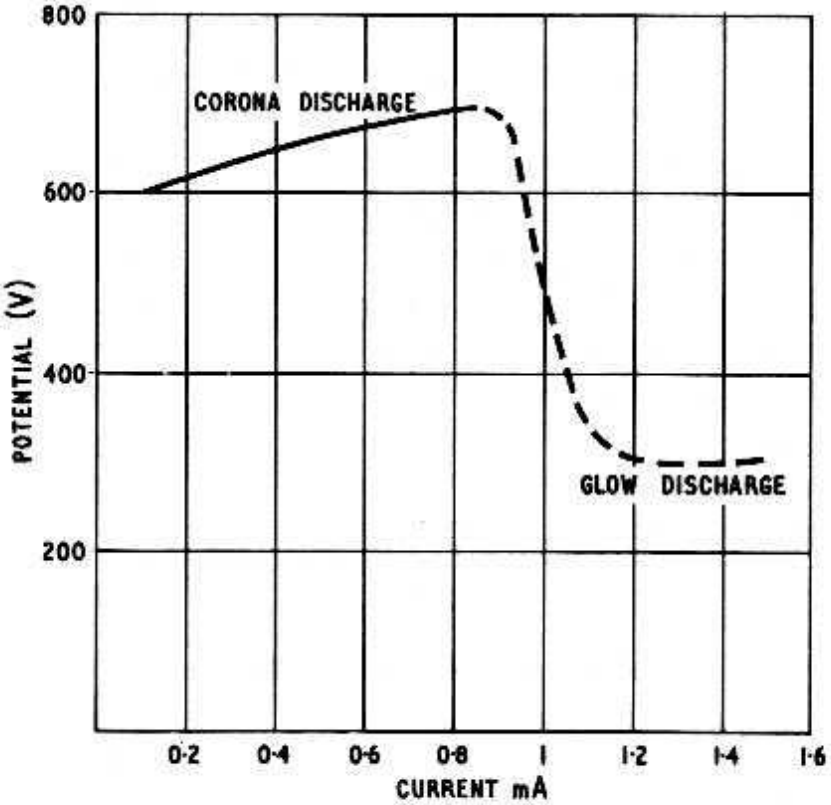


Fig. 2.18. Voltage-current characteristic for coaxial cylindrical geometry with the centre wire positive, filled with hydrogen at a pressure of 20 torr

travelling to the anode. With the wire negative the converse is true; the neon metastables will be produced near the cathode and the resulting multiplication of the electrons produced in the low field region near the anode will be small. Because of this the breakdown potentials for mixtures having argon concentrations of around 10^{-3}% are lower for the wire positive than for the wire negative.

Before leaving non-uniform field breakdown, mention must be made of the corona discharge. This phenomenon occurs when the field at one electrode is particularly high and/or the gas is at a high pressure. It is distinguished by a concentration of the glow at low currents around the electrode where the high field exists, for example around the central wire in a coaxial cylinder geometry. The latter geometry with the central wire positive gives a stable

corona discharge around the anode over a reasonable current range, and with suitable filling can be used as a stabiliser. As the potential is raised the gap breaks down into a corona discharge. Increasing the current increases the voltage across the tube, and there is a maximum in the V-i characterstic before it becomes negative with the development of a glow discharge, i.e. with characteristics similar to those for parallel-plate geometry at breakdown. Thus the breakdown voltage for the establishment of a glow discharge exceeds that required to produce the corona, see Fig. 2.18. In the corona discharge it is assumed that most of the ionisation takes place in the immediate vicinity of the electrode where the high field occurs, the anode in this case. The increase in voltage with current is then explained in terms of the ion space charge reducing the field at the anode, and thus the number of ionising collisions. At larger currents the space charge is sufficient to increase the field in the cathode region and the assumption that ionisation is confined to the vicinity of the anode is no longer valid, and the increase in ionisation at the cathode will more than compensate the decrease in ionisation rate at the anode due to the reduction in field strength. In the case of the negative wire, ion formation will tend to increase the field at the wire surface and thus give a negative V-i characteristic.

2.5. Effect of a Priming Current on the Breakdown Potential

Townsend's equation and the breakdown criterion assumes that the build-up of current before the breakdown potential V_s is reached does not introduce space-charge distortion of the field. However if the initiating current i_o is increased a critical value will be reached where this is no longer true. This can be seen from Fig. 2.19, where the V-i characteristics of a parallel-plate discharge tube are plotted as a function of the initiating current i_o. The critical i_o value is shown as the dotted curve at i_c. Above this value space-charge distortion occurs before the discharge becomes self-sustaining, and Equation 2.1 is no longer valid since i/i_o is a function of i_o as well as of V. As a result the maximum potential reached falls below the normal breakdown potential.

Several approximate methods have been proposed for determining the effect of space charge, and the consequent electric field distortion, upon the V-i characteristic of a uniform field discharge. Von Engel and Steenbeck[30] assumed that the space-charge field was a linear function of distance from the cathode and deduced that, for the negative V-i characteristic, the change in voltage due

to space charge was proportional to current density squared. Measurements of the negative characteristics by Büttner[31] and the author[32] suggest however that the change in voltage is linear with current. Acton and Swift[18] suggest that these results are accounted for by some ionisation process which is proportional to the square of the current such as collision between two metastable atoms. It is interesting that the resulting equation, when used to express V_s as a function of i_o, predicts the reduction of voltage produced by large

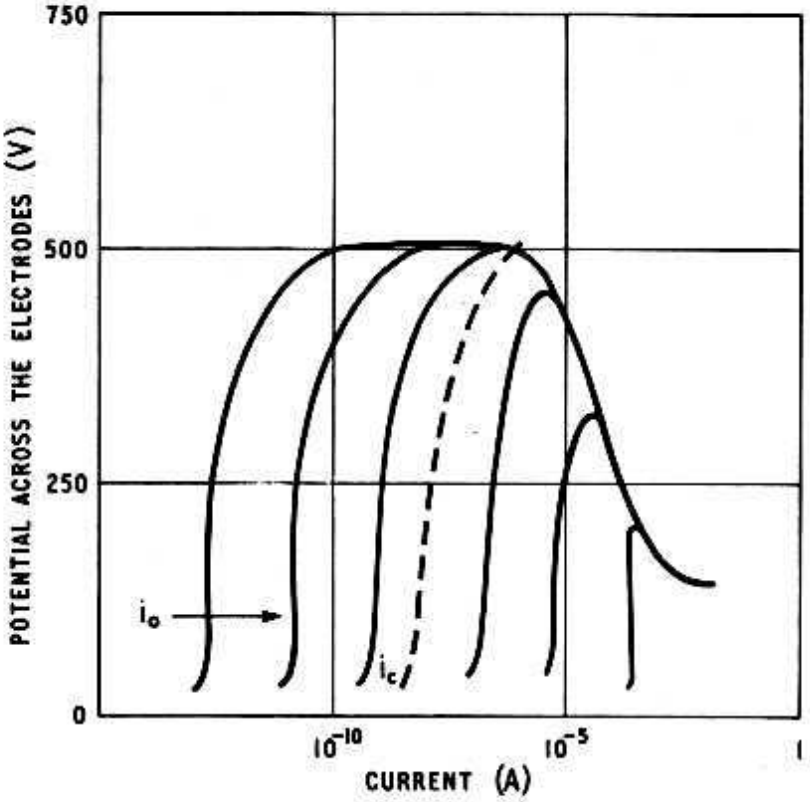


Fig. 2.19. The dependence of tube voltage on current for different i_o values. To the left of the dotted line $i = i_o f(V)$, to the right $i = i_o f(V, i_o)$

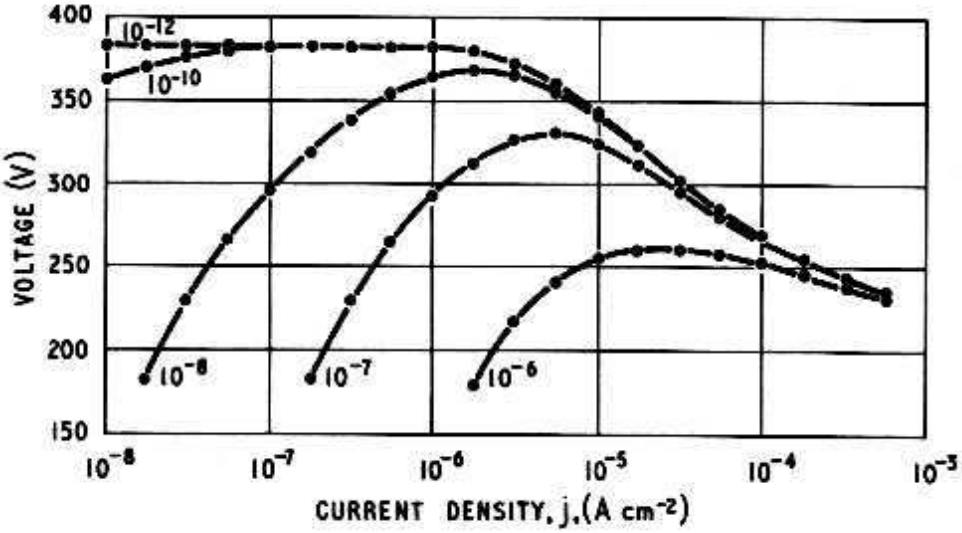


Fig. 2.20. Calculated voltage-current density characteristics for argon at pd = 10 cm torr by Ward[1] showing the effect of the variation of the initial current density j_o

primary current i_o as proportional to $i_o^{\frac{1}{2}}$, which agrees with the experimental findings of Fucks.[33]

A more rigorous study of the effect of space charge, in which Townsend's equation is combined with a one-dimensional Poisson's

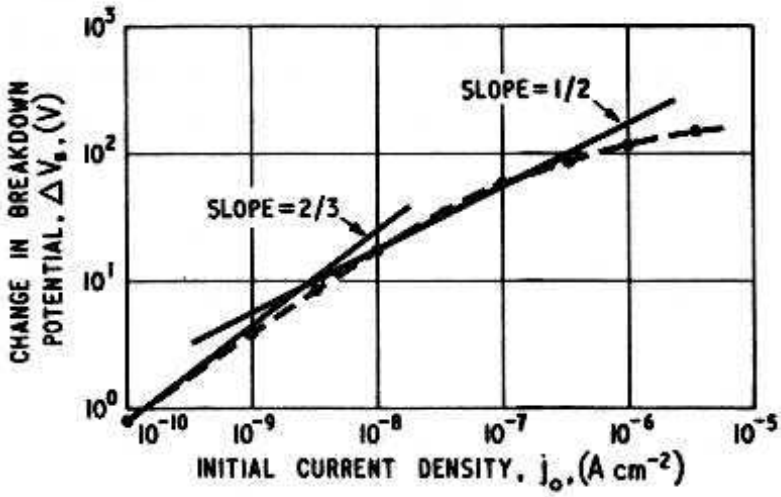


Fig. 2.21. The change of breakdown potential ΔV_s as a function of the initial current density j_o, for argon at pd = 10 cm torr. Ward[1]

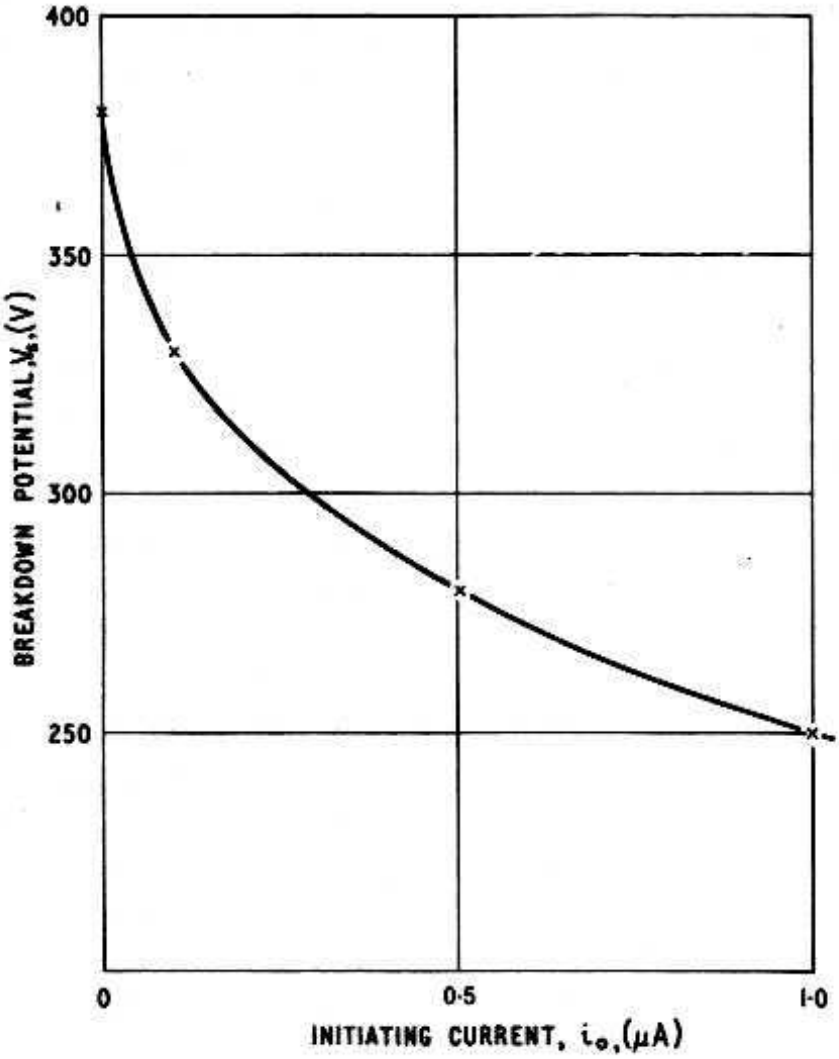


Fig. 2.22. Variation of breakdown voltage with initial current for argon deduced from Ward's[1] *data*

equation, has been made by Ward,[1] which explains quantitatively the negative V-i characteristic for hydrogen solely on the basis of space-charge distortion.[34] The analysis also gives the variation of the V-i characteristics for i_o values greater than the critical value i_c, and these are given as a voltage versus current density (j) plot in Fig. 2.20 for argon at a pd of 10 cm torr. The reduction of the breakdown potential ΔV_s plotted against initial current density j_o on a log scale, Fig. 2.21, shows that ΔV_s is proportional to $j_o^{\frac{3}{4}}$ at current densities from 10^{-10}—10^{-9} Acm^{-2}, but gives a decreasing power law with increasing current density, being proportional to $j_o^{\frac{1}{2}}$ from 10^{-8}–10^{-7} Acm^{-2}, and less than $\frac{1}{2}$ power law at higher currents. An approximate curve based on Ward's curve for V_s against i_o is given in Fig. 2.22.

In practical devices the lowering of the breakdown potential is made use of in many applications by providing the high initiating current by means of a subsidiary or priming discharge. Although the situation is complicated because the initiating current is no longer formed at the cathode and also it will only be a fraction of the total priming current, if V_s is plotted against the priming current a curve similar to that of Fig. 2.22 is obtained. The curve approaches asymptotically the maintaining potential V_m of the discharge at very high priming currents. This is discussed in more detail in Chapter 7.

REFERENCES

1. WARD, A. L., *Phys. Rev.*, **112**, 1852, (1958)
2. VARNEY, R. N., *Phys. Rev.*, **93**, 1156, (1954)
3. FROUWS, S. M., *Proc. Intern. Conf. Ionisation Phenomena Gases, 3rd, Venice*, (1957) Societa Italiana Di Fisica, 341.
4. PENNING, F. M. and ADDINK, C. C. J., *Physica*, **1**, 1007, (1934)
5. KRUITHOF, A. A., *Physica*, **7**, 519, (1940)
6. TOWNSEND, J. S. and MCCALLUM, S. P., *Phil. Mag.*, **6**, 857, (1928)
7. WESTON, G. F., *Brit. J. Appl. Phys.*, **10**, 523, (1959)
8. EHRENKRANZ, F., *Phys. Rev.*, **55**, 219, (1939)
9. HEYMANN, F. G., *Proc. Phys. Soc.*, **B63**, 25, (1950)
10. WESTON, G. F., *Brit. J. Appl. Phys.*, **11**, 305, (1960)
11. GOLDEN, D. E. and FISHER, L. H., *Phys. Rev.*, **123**, 1079, (1961)
12. MENES, M., *Phys. Rev.*, **116**, 481, (1959)
13. KACHICKAS, G. A. and FISHER, L. H., *Phys. Rev.*, **91**, 775, (1953)
14. TOWNSEND, J. S. and MCCALLUM, S. P., *Phil. Mag.*, **17**, 678, (1934)
15. JACOBS, H. and LAROCQUE, A. P., *Phys. Rev.*, **74**, 163, (1948)
16. QUINN, R. B., *Phys. Rev.*, **55**, 482, (1939)
17. HALL, R. F. and STOCKER, B. J., *Rad. Electron. Engr.*, **28**, 313, (1964)
18. ACTON, J. R. and SWIFT, J. D., *Cold Cathode Discharge Tubes*, Heywood, London, (1963)
19. CUEILLERON, J., *C. R. Acad. Sci., (Paris)*, **226**, 400, (1948)

20. LEMAIGRE-VOREAUX, M. P., *Bull. Soc. Franc. Electrns.*, **2**, 331, (1961)
21. RITOW, H., *J. Electron. Control*, **12**, 273, (1962)
22. DAVIES, D. E. and FITCH, R. K., *Brit. J. Appl. Phys.*, **10**, 502, (1959)
23. DAVIES, D. E. and HOPKINS, B. J., *Brit. J. Appl. Phys.*, **10**, 498, (1959)
24. CHANIN, L. M. and RORK, G. D., *J. Appl. Phys.*, **36**, 1515, (1965)
25. BENSON, F. A. and BURDETT, G. P., *Proc. Inst. Elec. Engrs.*, **108B**, 501, (1961)
26. OSKAM, H. J., *Philips, Res. Rept.*, **13**, 335, (1958)
27. MEEK, J. M. and CRAGGS, J. D., *Electrical Breakdown of Gases*, Clarendon Press, Oxford, 100, (1953)
28. PENNING, F. M., *Phil. Mag.*, **11**, 961, (1931)
29. MORTON, P. L., *Phys. Rev.*, **70**, 358, (1946)
30. VON ENGEL, A. and STEENBECK, M., *Electrische Gasentladungen*, Springer, Berlin (1932)
31. BÜTTNER, W., *Z. Physik*, **111**, 750, (1939)
32. WESTON, G. F., *Proc. Inst. Elec. Engrs.*, **111**, 1546, (1964)
33. FUCKS, W., *Appl. Sci. Res.*, **B5**, 109, (1955)
34. WARD, A. L. and JONES, E., *Phys. Rev.*, **122**, 376, (1961)

CHAPTER 3

The Glow Discharge

3.1. General Features

The V-i characteristic for a glow discharge is shown in Fig. 3.1, which represents region C of Fig. 1.1 expanded to higher currents. Increasing the current from the low-current self-sustained discharge results in a fall of potential due to the introduction of space charge, until a minimum is reached. At the minimum, the discharge is said to be a 'normal' glow discharge, and is characterised by a glow around the cathode. In this region the current density is constant, the glowing area spreading over the cathode surface as the current is increased with practically no variation in voltage. Raising the current above that required for the glow completely to cover the cathode increases the current density and the glow discharge becomes 'abnormal'. The potential in the abnormal glow region rises markedly with current until eventually the cathode heats up and an arc discharge develops. A glow discharge can be defined, therefore, as a self-sustained discharge in which the current density is sufficient to cause space charge distortion yet not large enough to heat the cathode to the temperature at which it will emit thermal electrons. The potential between the electrodes in a glow discharge is known as the maintaining potential V_m, or sometimes the running or burning voltage.

Although there is a wide variety of glow discharges, and their appearance varies with current, gas pressure, electrode geometry, etc., there are certain general features of light emission and potential distribution which are similar. These are shown in Fig. 3.2 for a long cylindrical tube with plane electrodes at each end at a pressure of the order of 1–10 torr; the potential distribution and the field along the tube are shown with the corresponding glow zones. Close to the cathode is a dark area in which most of the potential

drop occurs. At the edge of this dark space is a bright glowing zone known as the negative glow. At pressures of the order of 20 torr this glow extends 1 mm or so from the cathode, following the contours of the cathode as a glowing sheath. Beyond the cathode glow is a second and usually rather larger dark area known as the Faraday dark space, where the field is low and sometimes in the

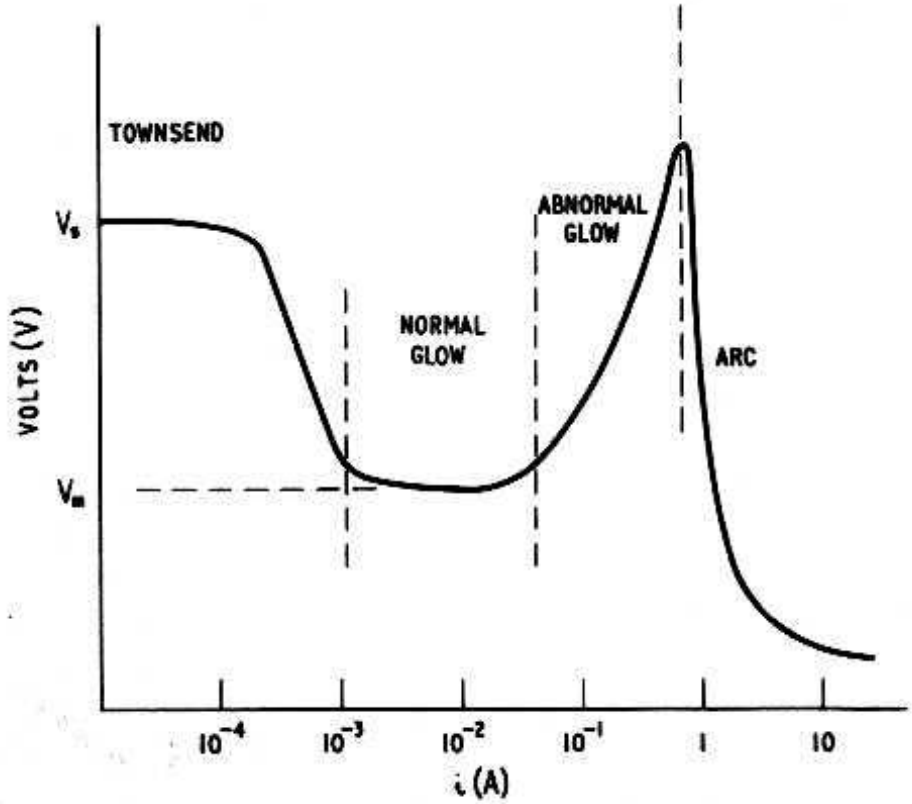


Fig. 3.1. Voltage-current characteristic of a gas-filled diode in the glow and arc discharge regions

reverse direction to that at the cathode. This is followed by the positive column, a glowing zone of low field which extends from the negative glow to the anode however far apart they may be. Although the field is low in the negative glow, Faraday dark space and positive column the ion and electron densities are appreciable and almost equal. Such a neutral zone of high electron and ion density is called a plasma. At the positive end of the positive column there is often a dark space followed by an anode glow close to the anode surface, where a rise in potential occurs. Other glow and dark spaces have been observed in the negative zone, and for a fuller description the reader is referred to the survey on the glow discharge by Francis.[1]

When the interelectrode distance is varied only the positive column changes its length; the zones near the cathode remain unaltered. Further, if the cathode is rotated the negative glow moves with the cathode as though attached to it, whilst the positive column simply joins the Faraday dark space to the anode. Only

the cathode dark space and negative glow are essential for the maintenance of the discharge, and in fact most cold cathode devices have their electrodes close enough for the positive column to be absent. The potential drop across the cathode dark space is known as the cathode fall V_c. The rise in potential at the anode is termed the anode fall.

A physical picture which helps to explain these general features of a glow discharge can be described by considering the transport of current through the glow discharge. Electrons emitted from the cathode, e.g. as a result of ion bombardment, are accelerated in the strong field of the cathode dark space. However, initially they will execute few ionising or excitation collisions because of their

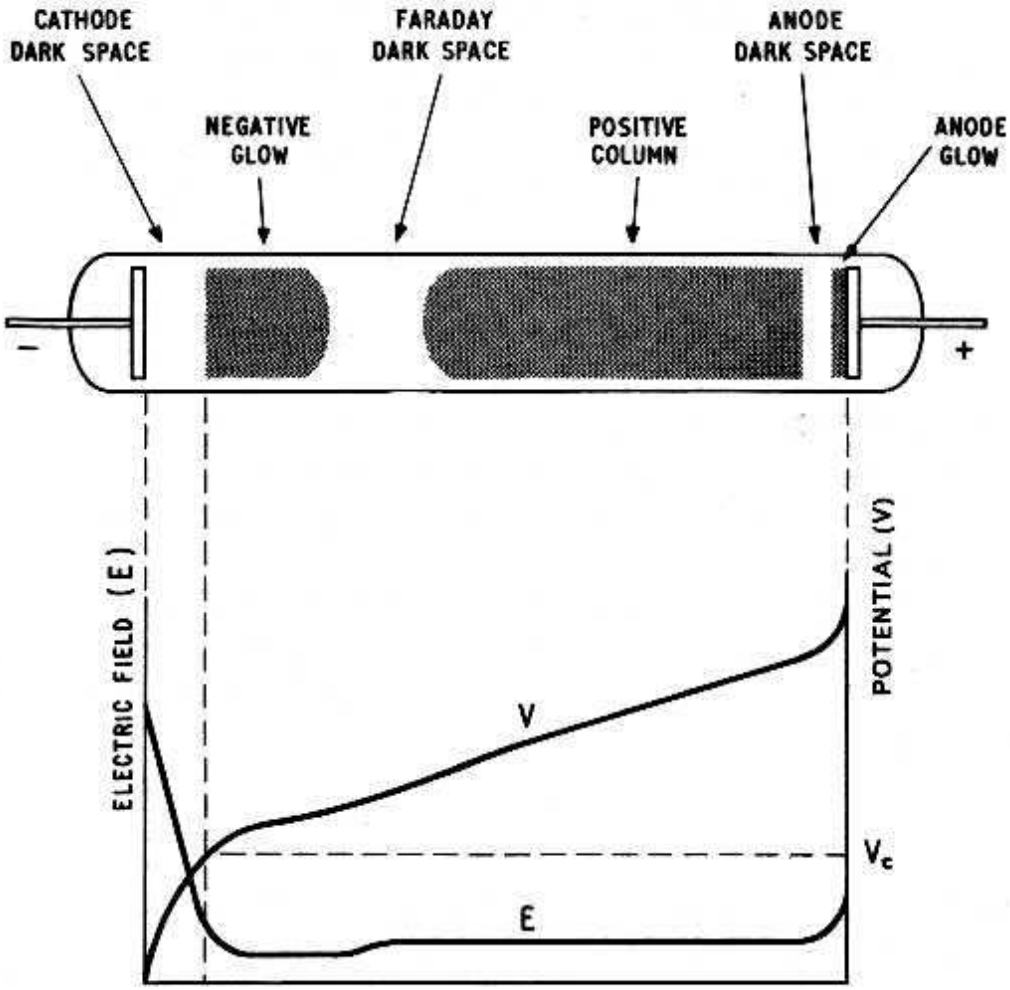


Fig. 3.2. Spatial distribution of luminous and dark zones in a glow discharge tube and the associated electric field and potential distribution

low energy, and only further from the cathode will the electrons ionise efficiently causing strong multiplication. The positive ions so formed move towards the cathode and are responsible for the space charge field, whereas the electrons, which have lost their energy in collision, will pass into the negative glow zone. Although

only a few fast electrons which have not lost their energy will be able to ionise in this glow zone, the total number of electrons will be large because of the efficient ionisation at the boundary, and many will have sufficient energy to cause considerable excitation. With increasing distance from the cathode the electron energy decreases since the field is weak, the number of excitation collisions falls off, and the Faraday dark space develops. Towards the positive edge of the Faraday dark space the field rises slowly and the positive column is formed. Here, the concentration of ions and electrons is equal, so that the net space charge is zero, but since the mobility of the electrons is some hundred times that of the ions the electric current is carried mainly by the electrons. The field is several orders smaller than that found in the cathode dark space, and therefore it is likely that ionisation in the positive column is a result of the random velocity of the electrons rather than the drift velocity in the field. At the anode, electrons are attracted to the electrode whilst the positive ions are repelled, and a negative space charge is set up in front of the anode giving the anode fall.

3.2. The Cathode Zone of the Glow Discharge

From the above description it is evident that the cathode zone is the most important section of the discharge, particularly the cathode fall zone where the electron multiplication and secondary processes required to sustain the discharge take place. For the discharge to be self-sustaining each electron leaving the cathode must produce a second electron at the cathode as a result of multiplication in the gas and the secondary processes at the cathode i.e.

$$\gamma(M-1) = 1 \tag{3.1}$$

Where M is the multiplication in the gas which can be expressed in terms of the Townsend first coefficient as

$$\int_{o}^{d_k} \alpha \mathrm{d}x \text{ or } \int_{o}^{V_c} \eta \mathrm{d}V$$

d_k is the thickness of the cathode fall zone.

Although Equation 3.1 appears to be identical to the Townsend equation pertaining to breakdown, the field across the gap which controls the value of α and γ is non-uniform. In the case of α, there may be errors introduced in applying Equation 3.1 if values

obtained under uniform conditions are applied to the glow discharge field situation (cf. Section 2.4).

However it is clear that, since both the gas pressure and composition will affect α and γ (the latter also being affected by the cathode material), then the cathode fall and indeed the maintaining potential of any discharge will, like the breakdown potential, depend on the gas-cathode combination. The cathode fall will also be affected by the current, and it is reasonable to search for similarity relations such as expressed in Paschen's law in order to reduce the number of interdependent parameters. Francis[1] deals with this in some detail and shows that from the similarity principle*

$$V_c = f_3(j/p^2) \text{ and } V_c = f_4(pd_k)$$

where j is the current density. Knowledge of the form of these two functions f_3 and f_4 would enable universal curves to be drawn up which could be applied to practical cases.

For an electrode system where the gap is small enough to eliminate the positive column and anode fall (for example in most stabiliser tubes) the maintaining potential will be a measure of the cathode fall.

3.2.1. THEORY OF THE CATHODE FALL

Numerous attempts have been made to derive the functions f_3 and f_4 from elementary processes and known constants. Essentially the cathode fall can be expressed in terms of three basic equations (given here for the x direction only).

1. Poisson's equation

$$\frac{dE}{dx} = -4\pi e\,(N_+ - N_-) \tag{3.3}$$

Where N_+ and N_- are the positive ion and electron concentrations.

2. The continuity equation

$$j = (N_+ v_+ + N_- v_-)\,e \tag{3.4}$$

where j is the total current density and v_+ and v_- are the velocities of the positive ions and electrons.

*The similarity principle depends on the gas temperature remaining constant. Since the temperature will depend on i/p^2 the relations are not strictly accurate.

3. The electron balance equation which is virtually Townsend's equation

$$\frac{dj_-}{dx} = j_-\alpha \tag{3.5}$$

Where $j_- = N_- ev_-$ = the electron current.
(When considering the negative glow as well as the dark space a diffusion term must be included in Equation 3.5.)

Clearly for these equations to be solved the variations of α, N and v with electric field have to be known, and also the boundary conditions. Unfortunately these relationships are not established, and in any case are likely to be different at different points in the gap due to the wide range of E across the gap. As a result assumptions have to be made with regard to these relationships which to some extent are arbitrary, and the various theories which have been propounded differ in their solutions according to the choice made.

Early theories considered the cathode dark space as an independent unit, and often omitted one or other of the equations, using instead empirical formulae. For example von Engel and Steenbeck[2] assumed the field to be linear in the dark space and used the empirical expression

$$\alpha/p = \mathrm{A}\exp(-\mathrm{B}p/E)$$

In this way they could substitute directly into the maintenance Equation 3.1 to obtain V_c as a function of pd_k. By assuming that the ion velocity was linearly dependent on E, they used Poisson's equation to relate j to d_k, and obtained V_c in terms of j/p^2

$$(\mathrm{C}_2 j/p^2)^{\frac{3}{2}} = (\mathrm{C}_1 V_c)^{\frac{1}{2}} \mathrm{f}(\mathrm{C}_1 V_c . \mathrm{C}_2 j/p^2)^{\frac{1}{2}} \tag{3.6}$$

where C_1 and C_2 are constants involving the properties of the gas and electrodes and the function $\mathrm{f}(\mathrm{C}_1 V_c . \mathrm{C}_2 j/p^2)^{\frac{1}{2}} = \mathrm{f}(z)$ is of the form

$$\mathrm{f}(z) = \int_o^z \exp\left(-\frac{1}{y}\right) dy \tag{3.7}$$

the solution of which can be obtained from tables. Equation 3.6 is thus a universal expression applying to all gases and cathode materials. The curve deduced by von Engel and Steenbeck from the equation is shown in Fig. 3.3. As j/p^2 is varied V_c passes through a minimum which corresponds to the normal cathode fall. The value of V_n according to von Engel and Steenbeck is

$$V_n = \frac{3\mathrm{B}}{\mathrm{A}} \log_e\left(1 + \frac{1}{\gamma}\right) \tag{3.8}$$

The curve to the left of the minimum is of no practical interest since, at currents in the normal cathode fall and below, the cathode area involved in the discharge varies with current and the current density cannot be measured. Druyvesteyn[3] has compared the theory with experiment using inert gases at various pressures with

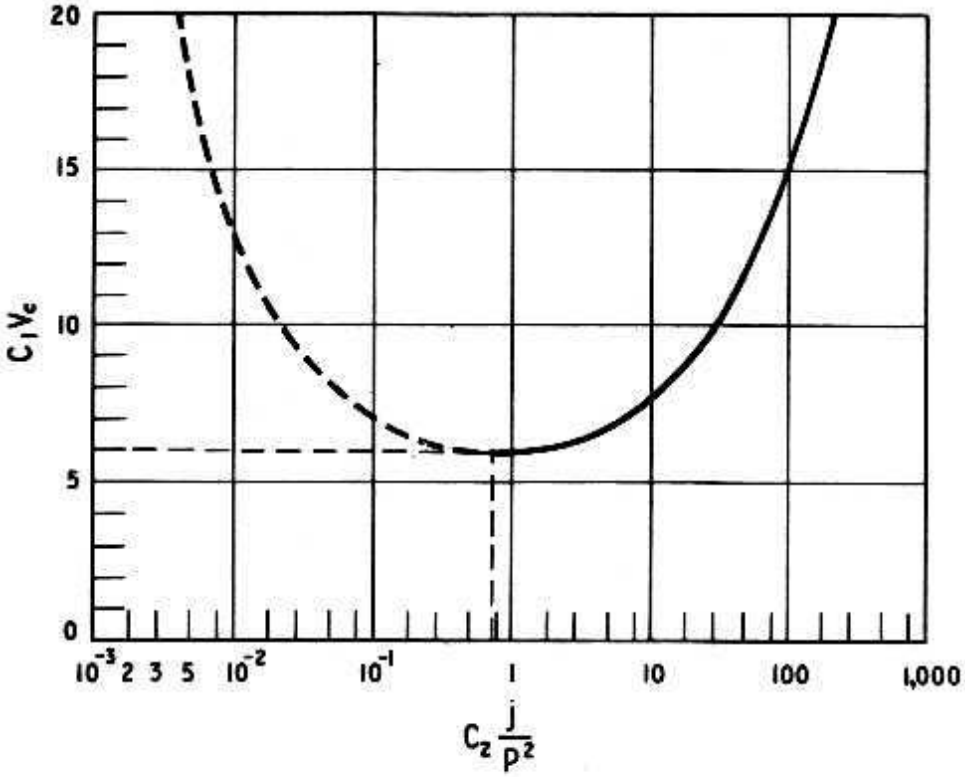


Fig. 3.3. Calculated curve relating cathode fall to the current density and pressure according to von Engel and Steenbeck[2]

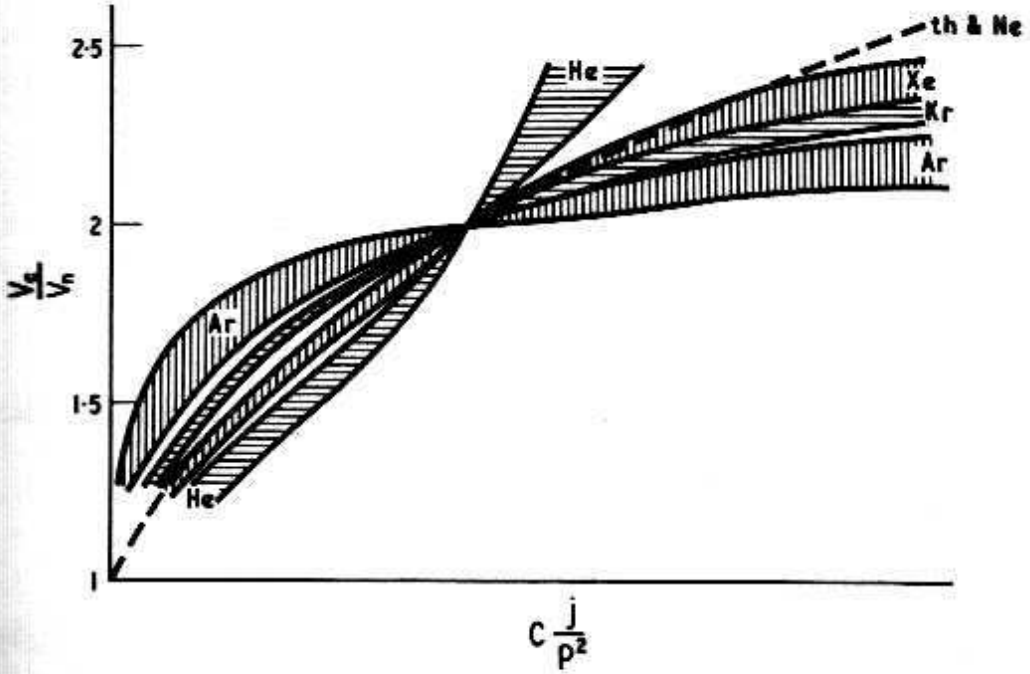


Fig. 3.4. Plot of V_c/V_n against Cj/p^2 showing the deviation from von Engel and Steenbeck's theory. The dashed curve gives the theoretical characteristic, and also that for He which is near to it. The domains for He *and* Kr *are shaded horizontally and those for* Ar *and* Xe *vertically. Druyvesteyn*[3]

a graphite cathode. Druyvesteyn plotted his results as V_c/V_n against $j/p^2 \times$ (a constant), the constant being chosen such that all the curves pass through a common point at $V_c/V_n = 2$. The curves are reproduced in Fig. 3.4 and show that although neon, xenon and krypton exhibit only small deviations from the theory, large deviations occur for helium and argon. This is not surprising in view of the assumptions that had been made.

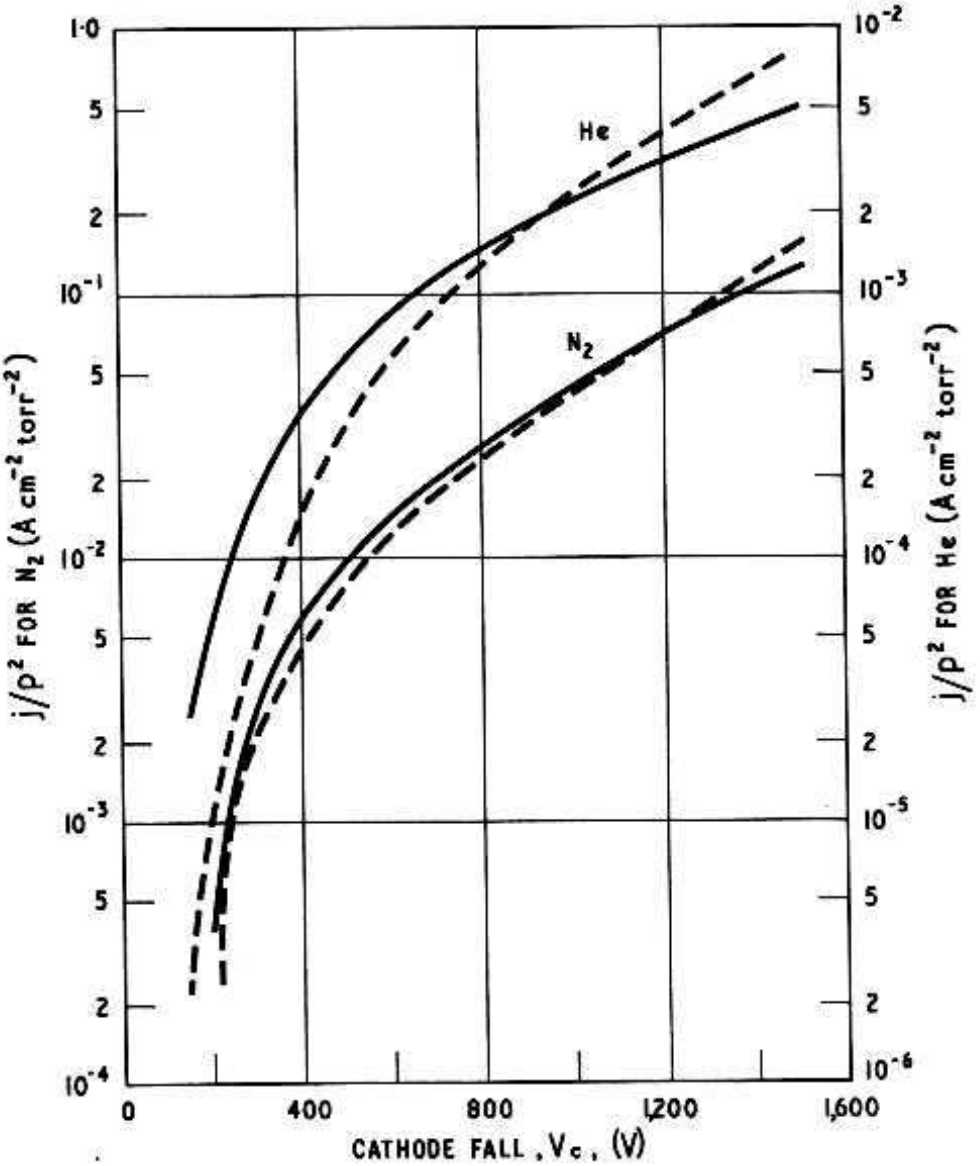


Fig. 3.5. j/p^2 as a function of the cathode fall for helium and nitrogen in discharge diodes with iron electrodes.
------experimental (Güntherschulze 1930)
——— Theory (Little and von Engel[4])

A summary of the other theories put forward at about the same time is given by Francis.[1] More recent theories have often considered the cathode dark space together with the negative glow, the latter acting as a source of ions, photons and metastable atoms. For example Little and von Engel[4] have expanded the theory

outlined above to include ions and also photons from the negative glow.

They consider the secondary process at the cathode as due to ions and photons originating in both the dark space and negative glow, and in deducing the relation between v_+ and E they include the

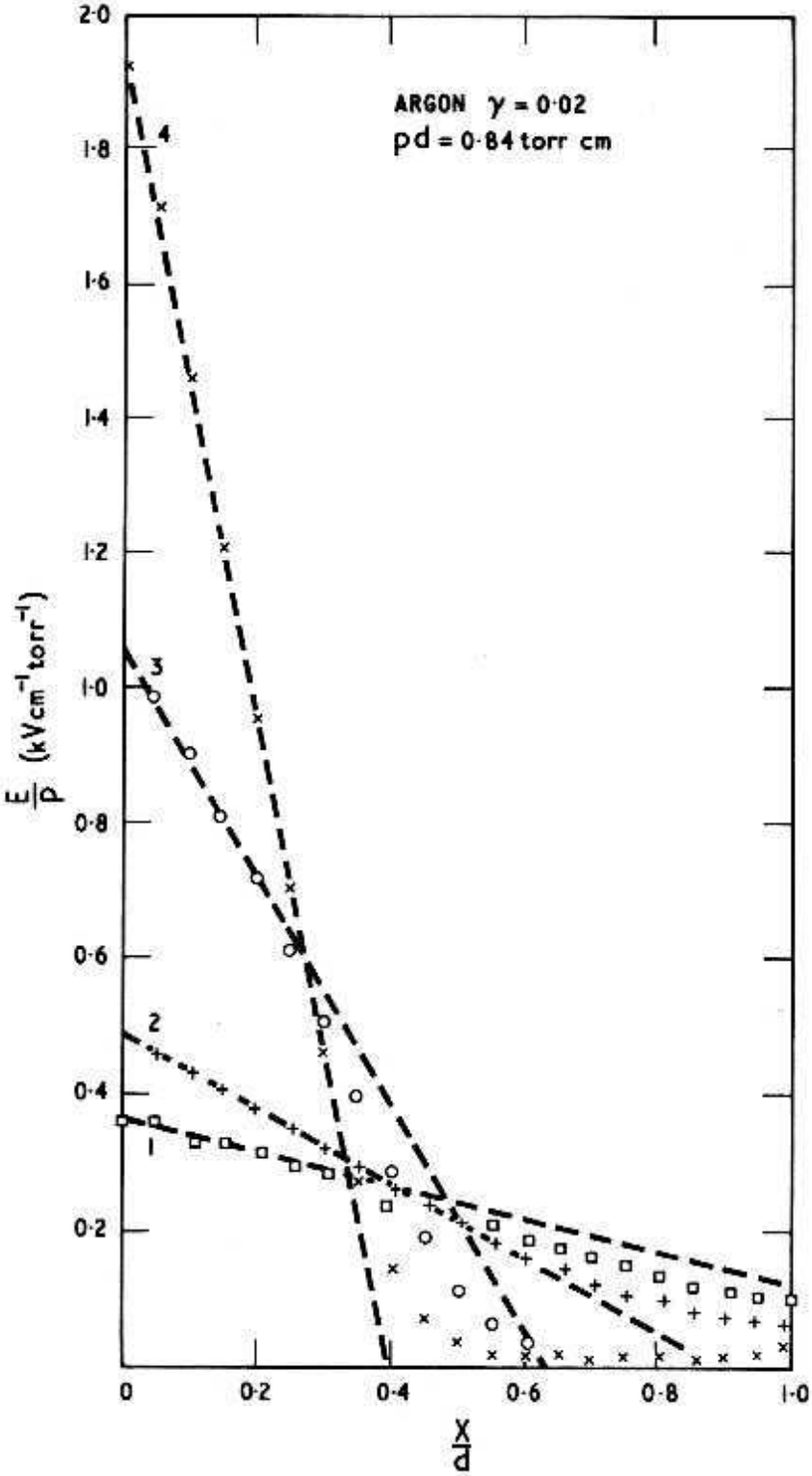


Fig. 3.6. Calculated ratio of electric field to pressure as a function of the ratio of distance from the cathode to total gap-distance for argon at pd = 0·84 cm torr. Curves 1 and 2 are for $j/p^2 = 1{\cdot}0 \times 10^{-5}$ *A cm*$^{-2}$ *torr*$^{-2}$ *and 3 and 4 are for* $j/p^2 = 1{\cdot}78 \times 10^{-4}$ *A cm*$^{-2}$ *torr*$^{-2}$. *Curves 1 and 3 are calculated using a constant positive ion mobility, whilst curves 2 and 4 use the expressions for* v_+ *as a function of E/p given in the text. Ward*[5]

effects of charge transfer and forward scattering. The final equation is given as

$$j/p^2 = \sqrt{\left[\frac{e\lambda_o}{\pi M}\right]} \cdot \frac{V_c^{3/2}}{(pd_k)^{5/2}} \cdot \sqrt{\left(1-\frac{\lambda_o}{2pd_k}\right)} \cdot S(\gamma) \quad (3.9)$$

where λ_o is the mean free path of the ions of mass M for charge transfer, and $S(\gamma)$ is the secondary process term involving a complex function of the various γ mechanisms. By making simplifying assumptions regarding $S(\gamma)$, the relation between V_c and j/p^2 has been calculated and compared with experiment in Fig. 3.5 for helium and nitrogen. Little and von Engel's theory also applies to a hollow cathode discharge (Section 3.1) where support for the general premise that

$$j/p^2 \propto \frac{V_c^{3/2}}{(pd_k)^{5/2}} \text{ was found empirically.}$$

In general the functions that have been derived relating V_c to pd_k and j/p^2 are complicated, and in many cases can only be expressed graphically.

In 1962 Ward[5] used a computer to calculate the parameters relating to the cathode-fall phenomena. He assumed the relation

$$\alpha/p = A \exp\ [-B(p/E)^r] \quad (3.10)$$

where $r = 1$ for molecular gases and $\frac{1}{2}$ for the atomic inert gases (cf. Chapter 1). By expressing v_+ as dependent on $(E/p)(1-C_3E/p$ at low values of E/p but as $(E/p)^{\frac{1}{2}}(1-C_4(p/E)^{3/2})$ at higher values (similar relations have been found empirically for uniform fields), Ward obtained an almost linear relationship between E and x, shown in Fig. 3.6. The assumption that the field distribution is linear in the cathode fall region has been justified by experimental investigations,[4] but Ward's calculations have been the only ones to deduce the premise theoretically. Although the calculated total gap voltages were not a unique function of j/p^2 for varying pd values, the calculated cathode fall was. The total gap voltages and cathode fall for helium at several pressures for a 1 cm gap are shown in Fig. 3.7. Comparison with experimental curves of Güntherschulze and Druyvesteyn for neon and helium are given in Fig. 3.8.

Several other attempts at calculating the cathode fall in terms of the current density have appeared in the 1960's.[6–8] However, often in attempting a more exact analysis refinements are introduced which must later be neglected, or estimated with little experimental basis in order to obtain the best fit with measured results. Further, the measured results with which they are compared are usually

those of Güntherschulze made in the early 1930's, which, although comprehensive, are for gases which would be considered contaminated by present day standards, and were certainly obtained with metal electrodes now known to give unstable discharge characteristics (i.e. copper, iron, aluminium).

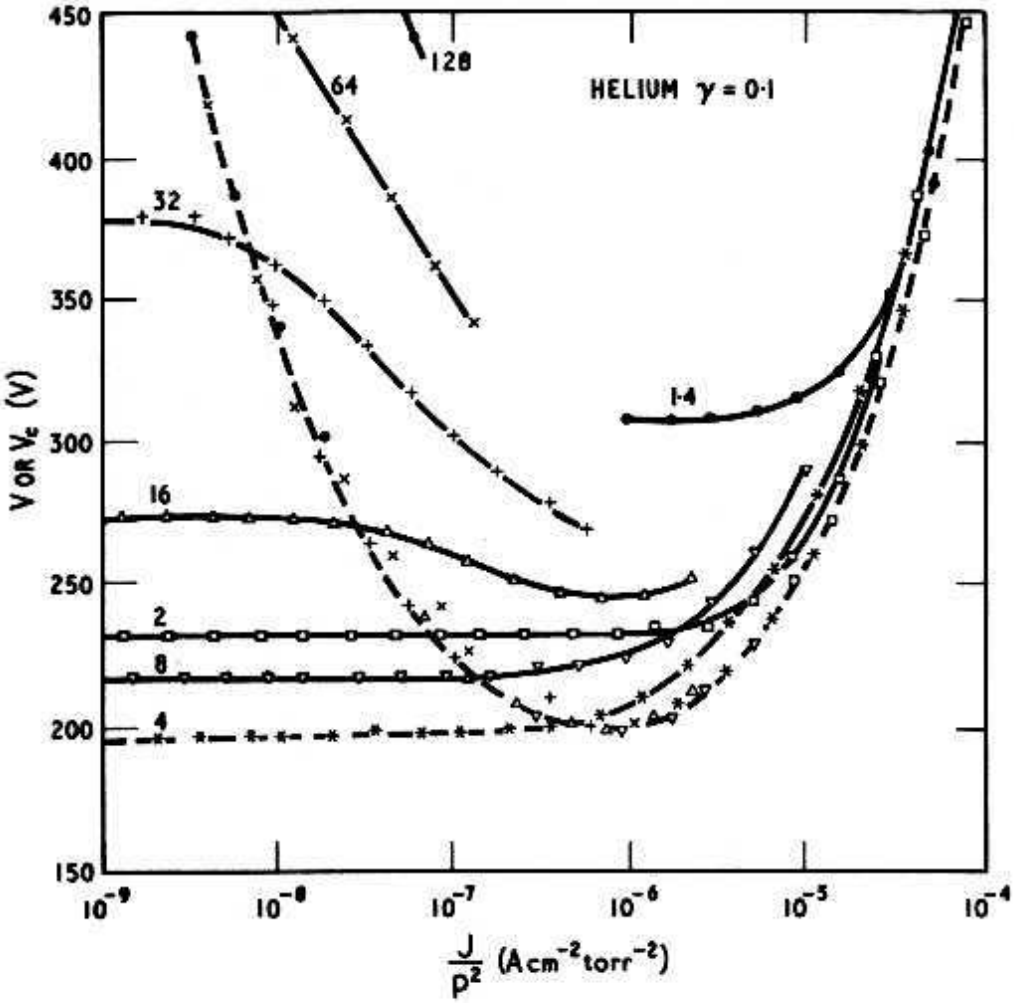


Fig. 3.7. Calculated tube voltages and cathode falls in helium as a function of j/p^2. *Points connected by solid lines are total voltage for the pd values labelled. Cathode fall is indicated by the same symbols but shows a unique function of* j/p^2 *i.e. dashed curve. Ward*[5]

It is concluded, therefore, that although a better understanding of the mechanism in the cathode fall region is now available, it is doubtful if the curves relating V_c to j/p^2 are any more reliable than those obtained from earlier theories. Clearly, experiments as comprehensive as those of Güntherschulze, but using modern vacuum techniques with cathodes cleaned by sputtering, are needed before theories of the cathode-fall region can be adequately tested. One factor which must be taken into account in obtaining the necessary data is the cathode temperature. Unless the cathode is adequately cooled the gas density across the inter-electrode space will be non-uniform, and as Smith[9] has pointed out this affects the cathode

fall. The theories so far proposed assume a uniform gas density, and usually use pressure corrected to 0°C as their parameter. This would be incorrect if a temperature gradient occurred across the tube. This is discussed in Section 3.4.

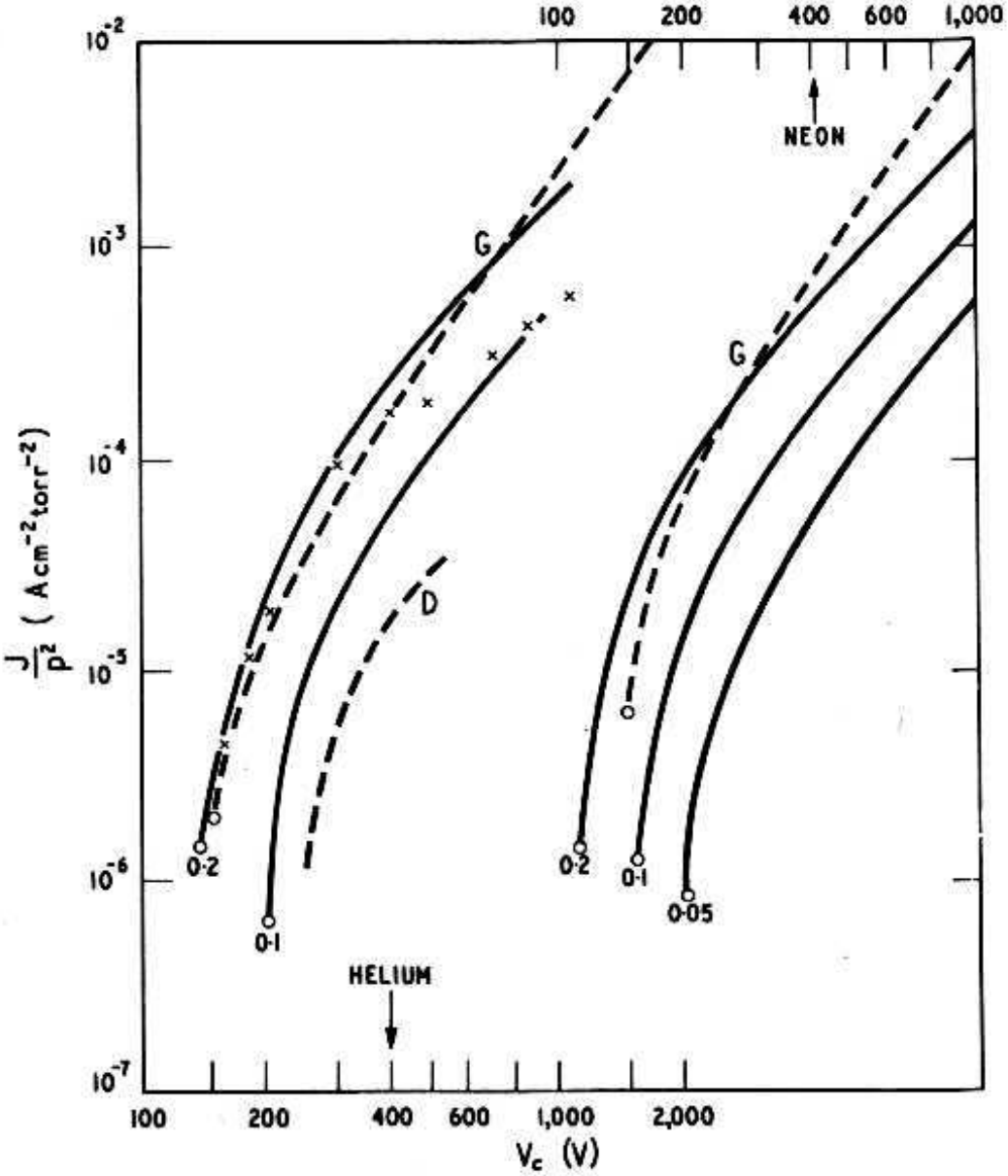


Fig. 3.8. j/p^2 plotted as a function of cathode fall. Solid curves calculated by Ward[5] with γ values shown. Dashed curves G are average values obtained by Güntherschulze (see Reference 1) whilst crosses are for one sequence of measurements $p = 0{\cdot}8$ to 14 torr. Dashed curve D from Druyvesteyn[3]

3.2.2. THE NORMAL CATHODE FALL FOR INERT GASES

Although there are insufficient measurements of the cathode fall as a function of the current density there are considerable data on the normal cathode fall, determined as the minimum maintaining potential in tubes where the electrode spacing precludes the development of a positive column or an anode fall. As V_n is a func-

tion of α and γ it may be expected that the values will be sensitive to gaseous impurities and surface contamination. Since this was not realised until the 1930's, the comprehensive table of V_n for various gases and cathode materials which was compiled in 1927[10] and which, unfortunately, is still being reproduced in physics textbooks today, is likely to be unreliable. Certainly comparison with recent data for inert gases shows the 1927 values of V_n to be from 10–40% high. Experiments by Penning and Moubis[11] have shown the importance of the cathode surface, and have led to a technique of cathodic sputtering which allows accurate and reproducible measurements of V_n. They observed that for a molybdenum cathode in neon gas V_n was reduced by 60V after sputtering the cathode, even though the tubes had been thoroughly baked and the electrodes outgassed by R.F. eddy current heating. They attributed the voltage drop to the removal by the bombarding ions of a thin layer of oxide on the surface, thus increasing γ.

Table 3.1

Normal cathode fall for molybdenum and zirconium cathodes in the inert gases

Gas	*Pressure (Torr.)*	V_n *(average)* (V)	*Reference*
Molybdenum Cathodes			
He	42	109·5	Jurriaanse, Penning and Moubis[12]
He	40	114	Weston[14]
Ne	57	108	Weston [13]
Ne	40	107·1	Jurriaanse, Penning and Moubis[12]
Ne	27	106	Milovanova and Chistyakov[15]
Ne	20	110·6	Jurriaanse, Penning and Moubis[12]
Ar	40	105	Weston[14]
Ar	20	103·4	Jurriaanse, Penning and Moubis[12]
Ar	12	100	Milovanova and Chistyakov[15]
Ar	10	103·6	Jurriaanse, Penning and Moubis[12]
Kr	13	114·8	Jurriaanse, Penning and Moubis[12]
Kr	5·8	120·2	Jurriaanse, Penning and Moubis[12]
Xe	13	133·6	Jurriaanse, Penning and Moubis[12]
Xe	5·4	134·4	Jurriaanse, Penning and Moubis[12]
Zirconium Cathode			
Ne	42·7	103·7	Jurriaanse, Penning and Moubis[12]
Ne	39·5	105·7	Jurriaanse, Penning and Moubis[12]
Ne	20·4	105·1	Jurriaanse, Penning and Moubis[12]
Ne	19·5	108·0	Jurriaanse, Penning and Moubis[12]
Ar	10·2	92·1	Jurriaanse, Penning and Moubis[12]
Ar	7·1	92·2	Jurriaanse, Penning and Moubis[12]

Penning and his co-workers followed up this work by determining the normal cathode fall of molybdenum and zirconium[12] in several inert gases. They found that very stable values reproducible to ± 1 or 2 V could be obtained with the molybdenum cathode in all the inert gases but, with zirconium, only neon and argon gave reliable results. The values of V_n are given in Table 3.1, together with the pressure, since it was found that V_n for some gases varied with pressure (see Section 3.4). Also included in Table 3.1 are results for molybdenum cathodes obtained by the author[13, 14] in a study of inert gas mixtures, and some results obtained by Milovanova and Chistyakov.[15] The agreement is within 5%.

A further paper by Penning and Moubis[16] gave V_n for a wide range of cathode materials in neon gas. In some cases a rod cathode was used instead of a plane cathode, and geometrical correction was made. The results given in Table 3.2 are reduced to a pressure of 40 torr assuming an empirical variation of V_n with pressure.

More recently Takeishi,[17] using similar techniques, has measured V_n for nickel, tungsten and barium oxide-coated nickel cathodes, with the inert gases, and compared results with V_n calculated from

Table 3.2

Values of the minimum maintaining voltage $V_{m\,min}$ and normal cathode fall V_n (reduced to plate cathode and 40 torr pressure) in neon—according to Penning and Moubis[16]

Cathode			*Pressure (torr)*	$V_{m\,min}$ (V)	ΔV (V)	*Correction*		V_n (40 *torr*) (*plate cathode*)
Material	*Shape*	*diam. (mm)*				*Shape*	*Pressure*	
C	—	—	20	200	±20	—	—	(195)
Si	rod	1·8	40	142	± 5	−5	—	137
Ti	plate	—	20	116	± 2	—	−3	113
V	plate	—	40	118	± 3	—	—	118
Cr	plate	—	40	123	± 5	—	—	123
Fe	rod	1·5	20	137	+10, −2	−4	−4	129
Ni	rod	2	20	145	± 2	−5	−4	136
Cu	rod	3·7	20	156	± 5	−3	−5	148
Zr	plate	—	40	104	± 3	—	—	104
Nb	plate	—	40	106	± 3	—	—	106
Mo	plate	—	40	107	± 2	—	—	107
Rh	plate	—	40	124	± 3	—	—	124
Ag	rod	2·3	20	152	± 5	−5	−4	143
Hf	rod	2	40	108	± 2	−3	—	105
Ta	plate	—	20	109	± 3	—	−3	106
W	plate	—	20	114	± 2	—	−3	111
Th	plate	—	20	115	± 3	—	−3	112

von Engel and Steenbeck's theory (Equation 3.8). The values of A and B employed are given in Table 1.1, and the γ values are either those of Hagstrum for 10 eV electrons or values calculated by Takeishi[18] in an earlier paper. The results are shown in Table 3.3

Table 3.3

Normal cathode fall in the inert gases according to Takeishi[17]

Cathode	Ni			BaO			W		
	p (*torr*)	V_n(*obs*) (V)	V_n(*calc*) (V)	*p* (*torr*)	V_n(*obs*) (V)	V_n(*calc*) (V)	*p* (*torr*)	V_n(*obs*) (V)	V_n(*calc*) (V)
He	20·6	145·0	149·2	15·1	80·5	84·4	20·0	121·0	116·4
Ne	17·7	143·0	146·5	15·0	73·5	76·5	36·0	114·5	118·1
Ar	16·5	142·5	142·0	14·9	56·0	53·0	10	117·0	114·4
Kr	—	—	—	—	—	47·7	—	—	128·3
Xe	15·1	196·5	205·0	6·6	46·5	45·6	6·6	250	173·3

and show good agreement with the calculated values. In view of the assumptions made in the theory, however, it cannot be assumed that similar agreement will be obtained with other gas-cathode combinations.

An improved stability for copper and silicon cathodes has been obtained by Ahsmann and van Gelder[19] using single-crystal cathodes. They discovered that the normal cathode fall depended on the orientation of the crystal, the most stable values being obtained with the (111) crystal face. Providing the tube was kept at a constant temperature, a voltage drift of less than 1 part in 10^5 was maintained over several hours for a copper single crystal. Germanium was also examined, intrinsic as well as p- and n-type samples being used. Significant differences in V_n between the different types of germanium however were not found. The results are given in Table 3.4. Unfortunately these single crystals showed

Table 3.4

The normal cathode fall for single crystals according to Ahsmann and van Gelder[19]

Metal	*Normal Cathode fall in Volts (40 torr Neon)*		
	Crystal plane (100)	Crystal plane (110)	Crystal plane (111)
Ge	132	129·8±0·5	131·5±0·5
Si	134·5	134·0	133·4
Cu	143·6	143·7±0·5	149·9±0·3

large variations of V_n with temperature which were not always reversible. Stability of V_n for single crystals of molybdenum has also been reported by Benson and Rigg[20] for a limited number of samples using a neon-argon mixture.

As with the minimum breakdown potential, attempts have been made to correlate the normal cathode fall with the work function Φ of the cathode material. Penning and Moubis showed that although a linear relation $V_n = C\Phi$ could be applied to many of the cathode materials examined in neon, large deviations occurred for others,

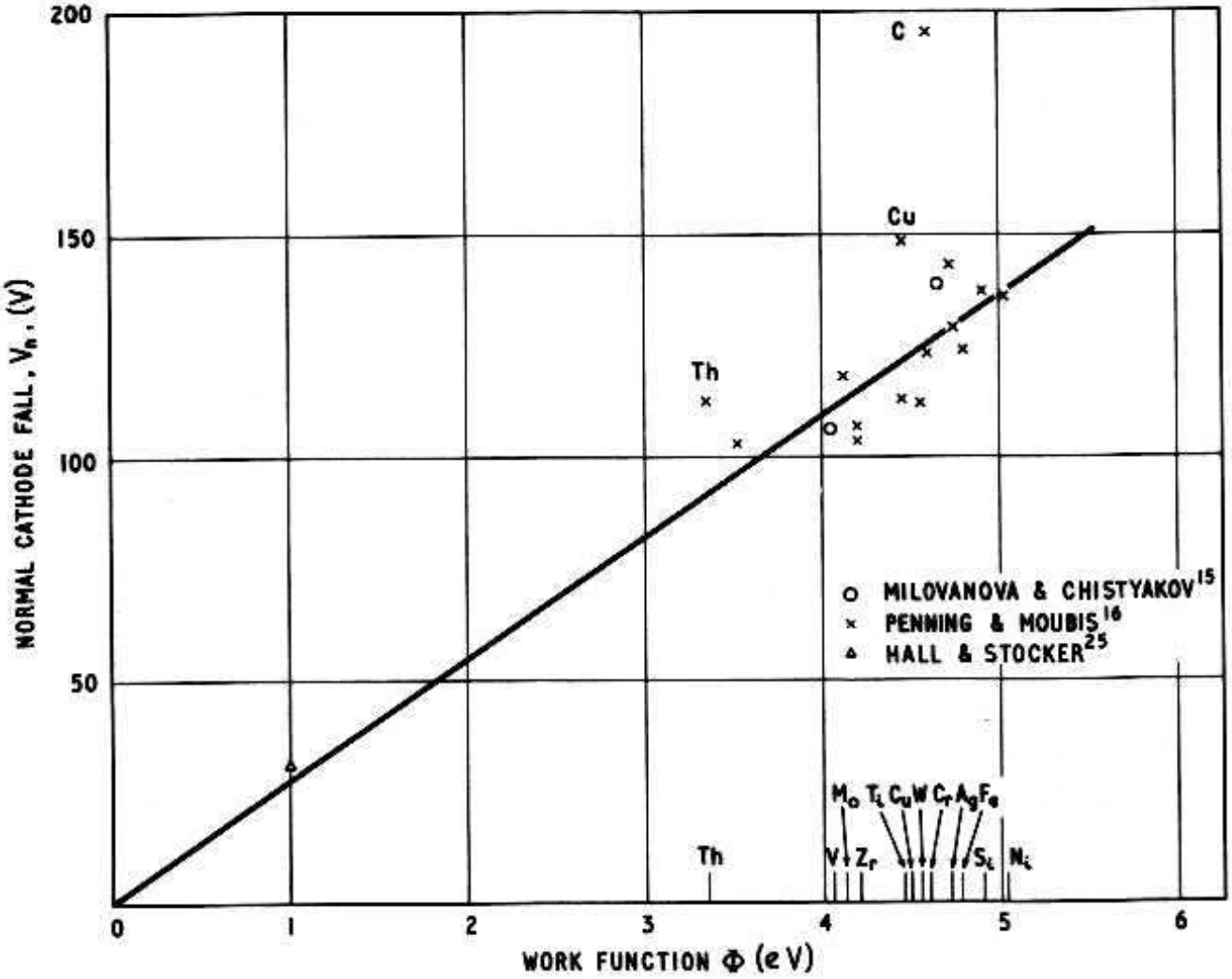


Fig. 3.9. Normal cathode fall in neon as a function of the cathode work function

in particular carbon, silicon, thorium and silver. However these were cathodes which were particularly difficult to clean, and the values of Φ quoted in the literature are rather uncertain, probably for the same reason. For example Penning and Moubis have taken the value of $\Phi = 3{\cdot}6$ eV for silicon measured in 1947,[21] whereas more recent values have tended to be about 4 eV, and for single crystals as much as 4·9 eV depending on the crystal face.[22] Lemaigre-Voreaux[23] has used the latter value for Φ in compiling a curve relating V_n to Φ for neon, and obtains only a small spread

about the straight line $V_n = 27\Phi$ (where V_n and Φ are in volts) even though very early data on V_n have been included. However he has omitted the anomalous cathode materials carbon, silver and thorium, and the work function for copper has been taken as 5·4 eV which is much higher than is generally accepted. Since for many of the metals the range of Φ reported in the literature is wide, the choice of Φ will influence the resulting curve. In Fig. 3.9 Penning and Moubis' results (Table 3.2) have been replotted against more recent work function data. The 'preferred values' of Φ, or if not available the latest results in the *Handbook of Chemistry and Physics*[24] have been used, with the exception of silicon where 4·9 eV has been taken. Although the general trend is evident the spread is quite large. Since it is possible that the vacuum work function cannot be applied to the case of a metal in an inert gas atmosphere the only way to investigate the relation between V_n and Φ is to measure both parameters in the same envelope and gas atmosphere. This has been accomplished by Milovanova and Chistyakov[15] for molybdenum and nickel in neon and argon, where Φ was measured by a photoelectric method. The values of Φ for the same cathode material but with different gases were very close, but rather lower than the published values measured in vacuum. For example in neon they obtained $\Phi = 4{\cdot}06$ eV for Mo and 4·65 eV for Ni as opposed to the generally accepted vacuum values of 4·2 eV and 5·03 eV respectively. The value of the constant C in the relation $V_n = C\Phi$ depended both on the gas and the cathode, and thus they concluded that the values of V_n and Φ showed considerable deviation from direct proportionality. Their results for molybdenum and nickel in neon are included in Fig. 3.9.

Comparing data in the various inert gases Lemaigre-Voreaux[23] suggests that a universal equation applicable to all cathodes and inert gases can be given as

$$V_n = (25{\cdot}5 + 0{\cdot}08\, V_i)\Phi \tag{3.11}$$

where V_n, V_i and Φ are expressed in volts.

He considers the fact that this formula does not agree with V_n measured in xenon and krypton is because the cathodes are not completely cleaned in these gases and still retain an oxide layer on the surface. However, Hall and Stocker[25] found that the minimum maintaining potential for a composite gold-caesium oxide cathode in helium was lower than in neon (22V and 30V respectively) which also contradicts Equation 3.11. At such low voltages the multiplication factor in the gas is less than 1 so that γ is all important, and since the potential energy of the helium ion is greater than that

of the neon ion the results are not unexpected. Therefore if an approximate equation such as 3.11 can be applied it is likely to be restricted to metal cathodes where V_n is several times the ionisation potential of the gas. It is interesting that the Hall and Stocker value for neon fits in with the general trend; this is seen in Fig. 3.9.

3.2.3. THE NORMAL CATHODE FALL FOR INERT GAS MIXTURES

The variation of cathode fall on mixing two inert gases shows similar effects to that of the breakdown potential discussed in

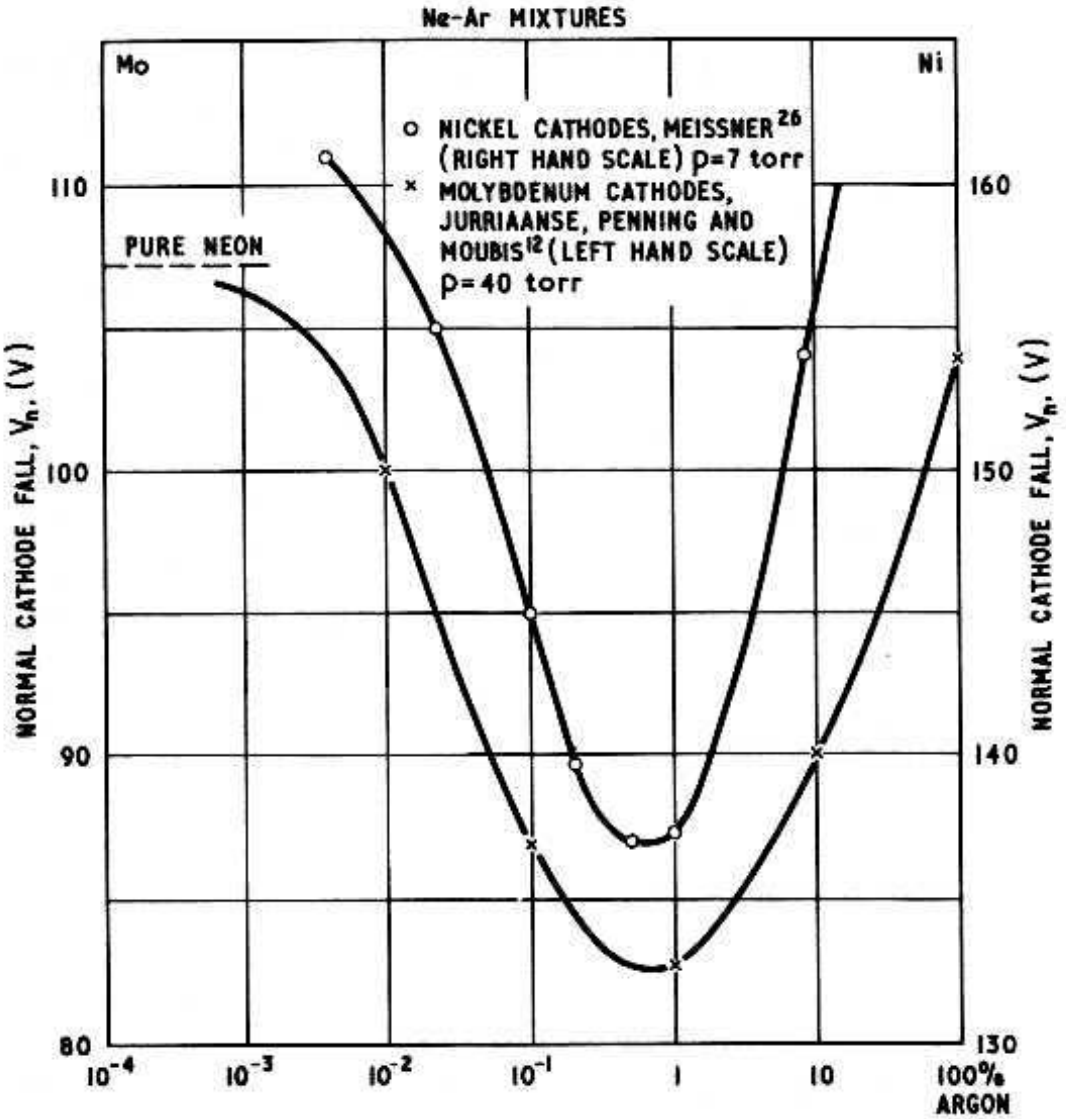


Fig. 3.10. Normal cathode fall as a function of the percentage of argon admixed in neon

Section 2.3. In particular, for a Penning mixture where the metastable level of the parent gas is higher or equal to the ionisation potential of the added gas (see Section 1.2.3), small concentrations of the admixed gas results in a lowering of the cathode fall below that found for either of the constituent gases. For example in neon-

argon where the effect is most pronounced due to the relevant potentials being so close, 1% of argon in neon lowers the cathode fall by approximately 20%. This is shown in Fig. 3.10 for a molybdenum cathode according to Jurriaannse, Penning and Moubis[12] and for a nickel cathode according to Meissner.[26] Similar curves for neon-krypton and neon-xenon have been obtained by Benson and his co-workers[27, 28] for a molybdenum cathode at 40 torr gas pressure and are replotted in Fig. 3.11. The

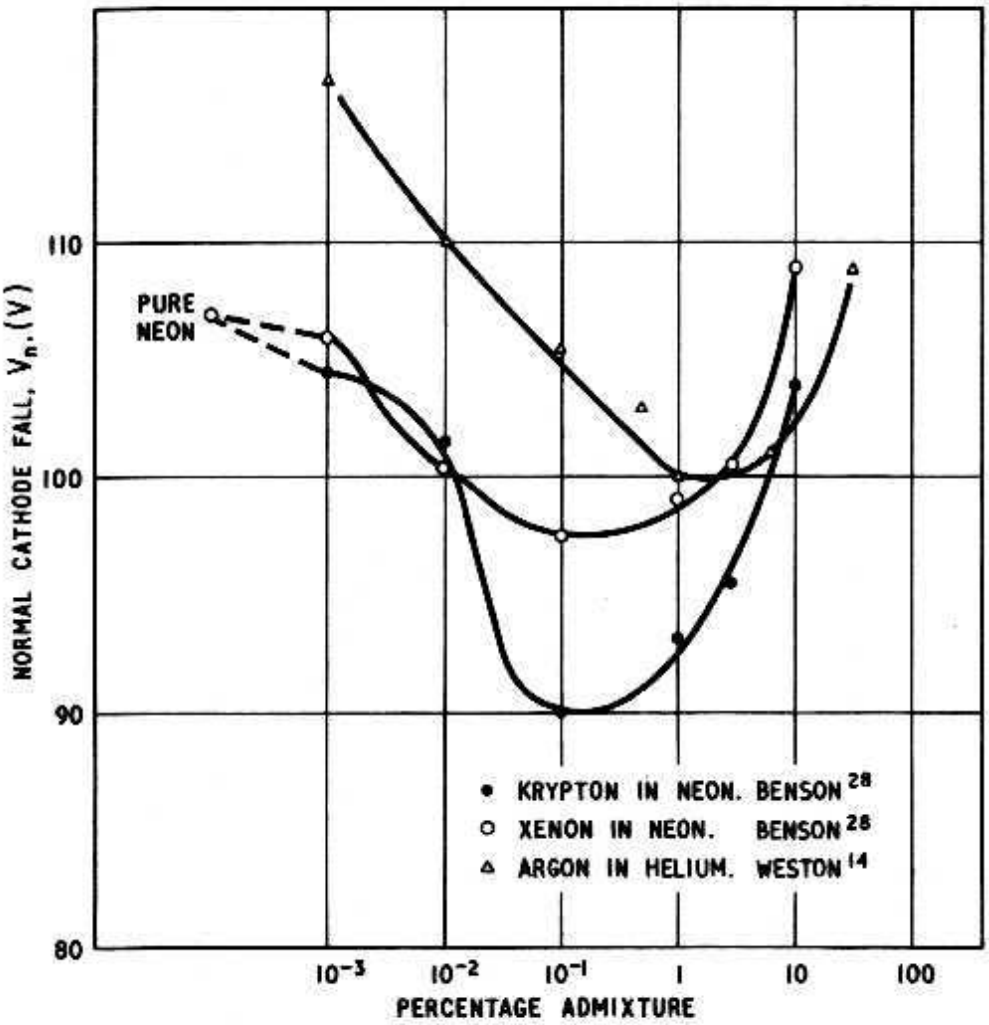


Fig. 3.11. Normal cathode fall for a molybdenum cathode in various Penning mixtures at a pressure of 40 torr as a function of the percentage of admixed gas

effect decreases as the difference between the metastable level of the neon and the ionisation potential of the added gas increases. This is expected as the probability of the Penning reaction shows a resonance when the two relevant potentials are equal. Also on Fig. 3.11 are values of V_n for helium-argon mixtures at 40 torr with a molybdenum cathode obtained by the Author.[14] Again the curve is similar although the minimum occurs at a slightly higher percentage of the admixed gas. Benson and Burdett[27] have

Table 3.5

Minimum maintaining potential of neon-argon-krypton and neon-argon-xenon at 40 torr with molybdenum cathode according to Benson and Burdett[27]

Gas	*Minimum maintaining potential* (V)
Neon + 0·3% argon	81·7 → 83·3
Neon + 0·3% argon plus 0·001% krypton	84·6 → 85·75
0·01% ,,	84·95 → 85·8
0·1% ,,	85·1 → 85·8
1·0% ,,	90·35 → 93·4
0·001% xenon	84·6 → 84·8
0·01% ,,	84·4 → 84·65
0·1% ,,	93·2 → 94·5
0% ,,	*98·0 → 101·9

*It is likely that an anode fall occurred in this mixture.

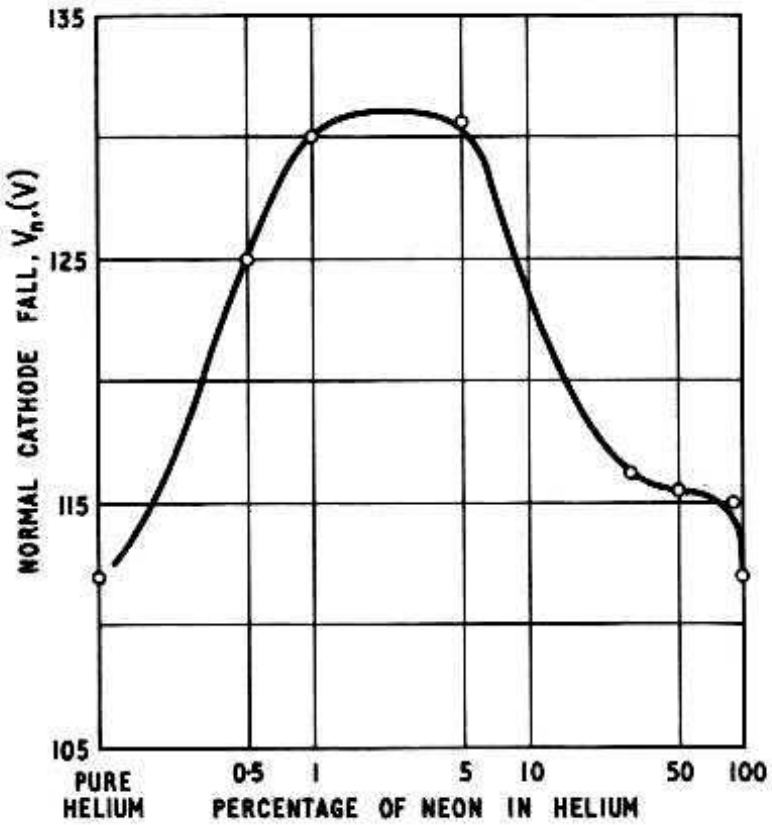


Fig. 3.12. The normal cathode fall as a function of the percentage of neon admixed in helium (molybdenum cathode)

also examined the effect of adding krypton and xenon to a neon-argon mixture. The results are given in Table 3.5.

The effect of adding the krypton and xenon is to reduce the effectiveness of the argon in the Penning reaction, although it is not clear exactly what takes place.

An increase in the normal cathode fall above that of the

constituent gases has been observed by the author[14] for helium-neon mixtures (see Fig. 3.12). The phenomenon is thought to be due to either a charge exchange process or the formation of a neon-helium ion (cf. Section 2.3). The pattern of behaviour is similar to the V_s curve against gas composition at the Paschen minimum (cf. Fig. 2.13). The fact that the normal cathode fall for gas mixtures

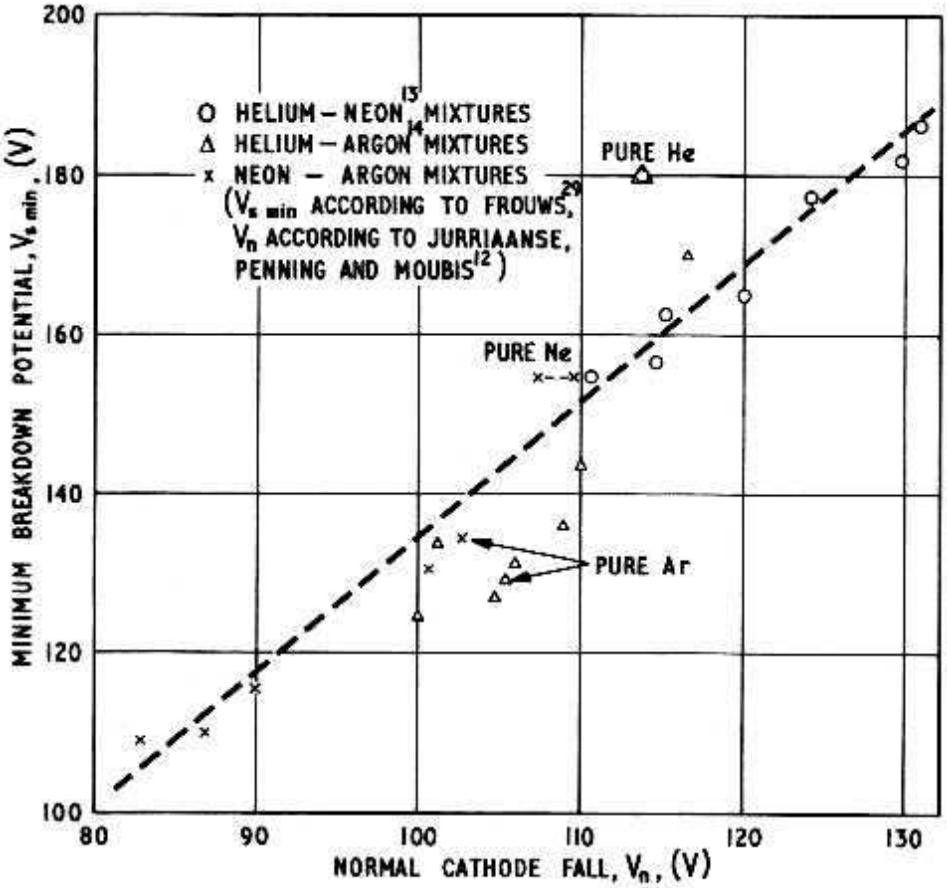


Fig. 3.13. The minimum breakdown potential plotted against the normal cathode fall for molybdenum cathodes in various inert gas mixtures

shows similar variations to the breakdown minimum suggests that the two parameters can be linearly related. In Fig. 3.13 V_n has been plotted against $V_{s\min}$ for helium-neon and helium-argon mixtures from the author's results, and V_n from Jurriaanse, Penning and Moubis[12] plotted against $V_{s\min}$ from Frouws[29] for neon-argon mixtures.

Although the helium-neon and neon-argon points lie close to the straight line

$$V_n = 0{\cdot}58V_{s\min} + 22 \tag{3.12}$$

where V_s and V_n are expressed in volts, the values for helium-argon deviate considerably. It is therefore concluded that a simple linear relation is unlikely, although Equation 3.12 gives a useful approximation for the design of glow discharge tubes.

3.3. The Positive Column and Anode Fall

The positive column normally appears only when there is a large separation between the electrodes, and is not present in the majority of practical cold-cathode devices. It can exist either as a uniform glowing column with its electrical properties constant along its length, as previously described, or as a series of striations with regions of bright light alternating with darker regions along its length. These striations can be stationary or progress along the column, usually from anode to cathode, at speeds of the order of 10^3–10^4 cm sec^{-1}. In inert gases, particularly pure gases, the uniform column is more common, although what often appears as a uniform column has sometimes been shown to be a series of striations moving too rapidly to be observed by eye. The theory of the positive column is based on balance equations relating (1) the ions lost to the walls with the production of ions, (2) the electron temperature required for ionisation with the number of ions produced, and (3) the energy dissipated with the input power. Several workers have developed such theories using empirical data or accepted assumptions, but in view of the fact that a positive column rarely occurs in cold cathode tubes the theories will not be outlined here and the reader is referred to the article by Francis[1] for more detail. The field in the positive column is found to be a function of the gas density and of the gas composition. In general, gases with the lowest ionisation potentials give the lowest fields, and this can be seen in Fig. 3.14 where the variation of E with p is given for a number of gases according to Klarfeld.[30] The current has little effect on the field except at very low current densities (10^{-6}A cm^{-2}) where it rises markedly as the current is reduced. The tube diameter, however, has a marked influence, the field decreasing as the diameter increases until the glow no longer fills the tube, when it becomes independent of the tube diameter. This will of course depend on the pressure. An interesting parameter of the positive column is the electron temperature which, since there is no thermal equilibrium with the ions or neutral particles, can be many times higher than the gas temperature. In fact it is of the order of 10^4–10^5 °K, and this is exploited in microwave circuits where the positive column of an inert gas discharge is employed as a noise source.

The anode fall is of more importance in cold cathode tubes since it is normally present in trigger tubes and counting tubes where a large difference is required between breakdown and maintaining potential. When the anode is situated in the positive column, electrons are attracted to it whilst positive ions are repelled, thus

setting up a negative space-charge sheath around the anode. The thickness of the sheath and the potential drop V_a are dictated by the need to produce the same number of positive ions in the sheath as flow out of the positive column at the other end into the Faraday dark space. Since this current is small, V_a need only be of the order of the ionisation potential, and this is true for large anodes where

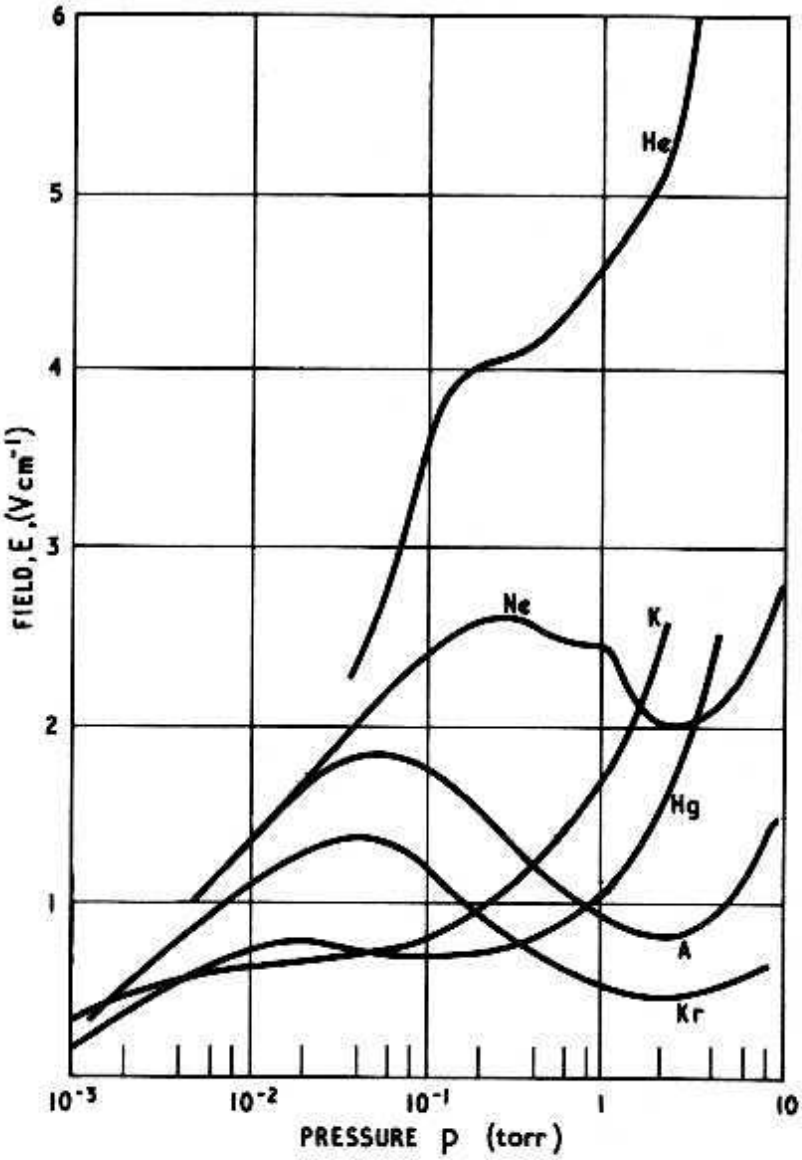


Fig. 3.14. The field in the positive column as a function of pressure for various gases, Klarfeld[30] *Tube diameter = 20 mm, i = 0·3 A*

most of the electrons from the positive column are collected. For small anodes there are greater diffusion losses and a higher potential is required to sustain the positive ion current. Similar considerations can also be applied to the current density where the potential drop increases as the current density rises. At low current densities the sheath appears as a uniform glowing layer about the anode surface, whereas at higher current densities the glow concentrates in small regions regularly spaced over the anode surface. When the anode is situated in the Faraday dark space near

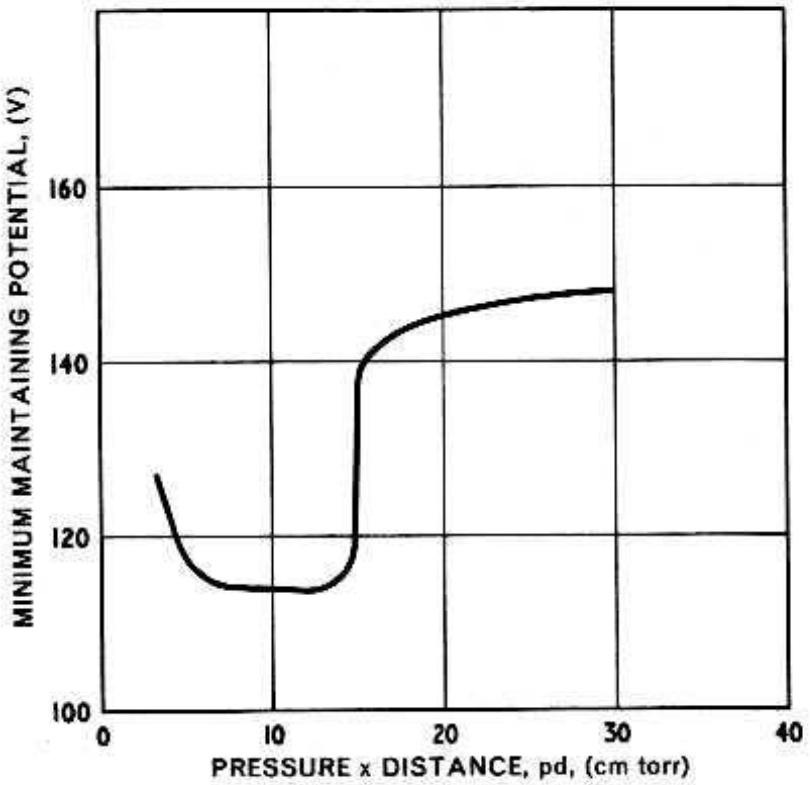


Fig. 3.15. Minimum maintaining potential as a function of pd for helium with molybdenum electrodes, showing the onset of anode fall

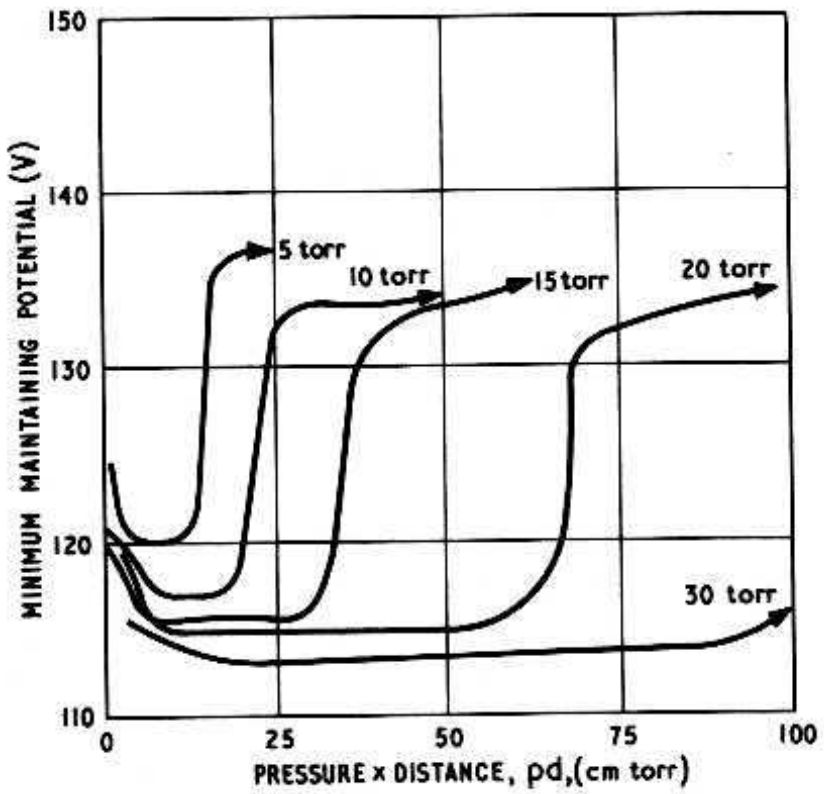


Fig. 3.16. Minimum maintaining potential as a function of pd for a mixture of 95% neon + 5% helium at various pressures, showing the deviation from similarity laws

the negative glow where there is excess of electrons, it is not necessary for any ionisation to occur since there is no flow of positive ions from the Faraday dark space into the negative glow. Thus an anode fall or glow does not develop and the current is maintained by the excess of electrons. Clearly, between the two anode positions there will be a transitional region as the anode is moved away from the cathode and there will be a critical distance d_c between the electrodes at which the anode fall sets in. In fact this occurs before the positive column develops, and depends on the pressure. Güntherschulze[31] observed that the product pd_c for onset of an

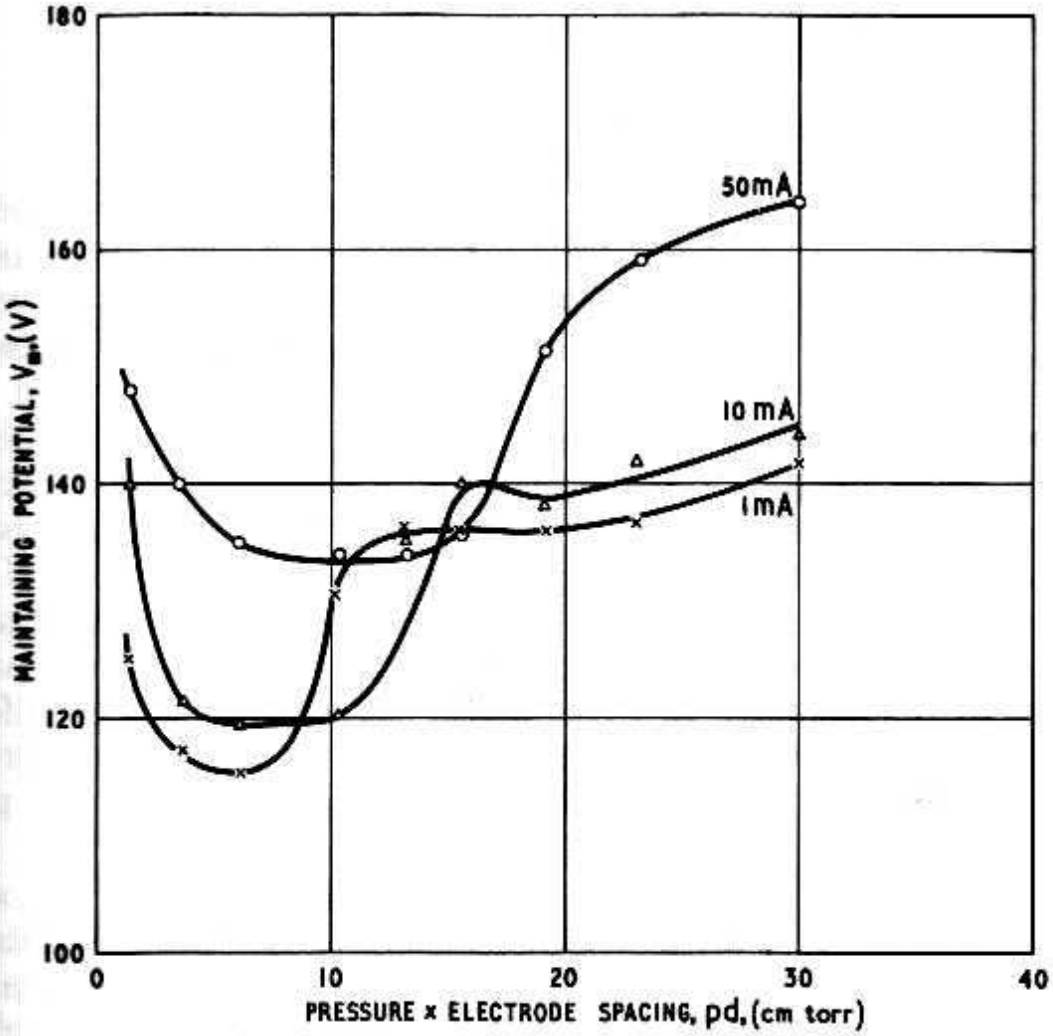


Fig. 3.17. Maintaining potential as a function of pd for a mixture of 70% helium + 30% neon at three different tube currents. Molybdenum electrodes

anode fall was constant for a given V_c and this was confirmed by the author[13] for helium, shown in Fig. 3.15 where the minimum maintaining potential has been plotted against *pd*. However in other gases it has been found that pd_c is not constant but depends on the pressure, for example see Fig. 3.16 where results are plotted for a 95% neon–5% helium mixture. The value of pd_c depends on the nature of the gas and also on the current density. The latter can

be seen in Fig. 3.17 which shows that the pd_c value increases with current. Thus for a given pd in a discharge tube an anode fall may occur at low currents but disappear when the current is increased.

The transition of the anode glow from one mode to another i.e. a uniform sheath to a series of spots, is also accompanied by a change in V_a, and further some modes are subject to relaxation oscillations whilst others are apparently stable. Unfortunately the anode fall has not been the subject of extensive study and although observations of the various modes have been reported, there are little data available linking the observations with measurable parameters.

3.4. Temperature Effects in a Glow Discharge

It is usual to take gas pressure as one of the dependent parameters in a gas discharge when attempting to explain observed phenomena, whereas it is the molecular concentration which should be used. However the use of pressure as a parameter is permissible on condition that there are no temperature gradients within the apparatus and that the reduced pressure is used, i.e. the pressure at 0°C which gives an identical gas density to the experimental conditions. In a glow discharge, however, the power is dissipated mainly in the cathode fall region, and the temperatures of the cathode and the gas in the immediate vicinity are considerably higher than that of the rest of the tube. Under these circumstances although the pressure is uniform the gas density is not uniform throughout the tube, and pressure cannot be used. This is important when considering the effect of envelope temperature on a glow discharge.

When the envelope temperature T_e of a glow discharge tube, such as a stabiliser, is increased at constant current the maintaining voltage V_m is usually observed to decrease. The temperature coefficient defined as the rise in V_m for 1°C increase in envelope temperature is therefore usually negative, although in certain tubes positive coefficients have been observed[32, 33]. Some typical curves of V_m versus T_e are given in Fig. 3.18 which shows that the relationship is not strictly linear and the coefficient, therefore, varies with temperature.

The effect of temperature was first studied by Jurriaanse[34] in 1946. His explanation was based on the earlier observation that, contrary to the similarity principle, the normal cathode fall decreased with increasing pressure. This effect of pressure is demonstrated in Fig. 3.19 where the V-i characteristics are given

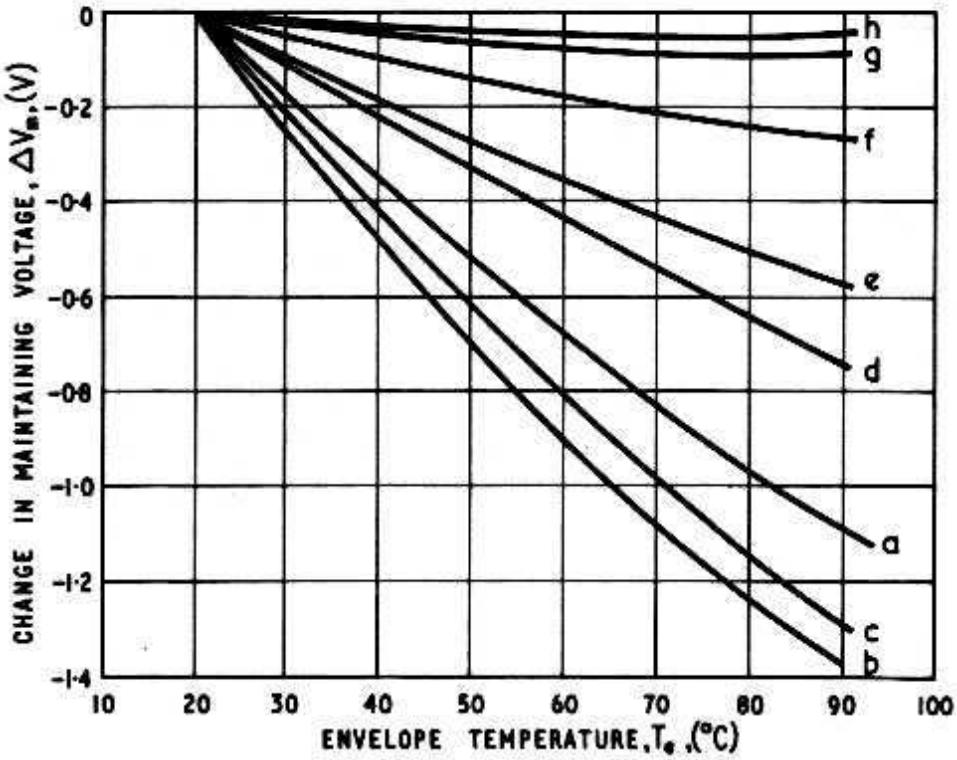


Fig. 3.18. Maintaining potential as a function of the envelope temperature for various neon-argon mixtures. Benson and Chalmers[33]

(a) *Pure neon*
(b) *Neon + 10^{-3}% argon*
(c) *Neon + 10^{-2}% argon*
(d) *Neon + 10^{-1}% argon*
(e) *Neon + 0·25% argon*
(f) *Neon + 0·5% argon*
(g) *Neon + 0·75% argon*
(h) *Neon + 1·0% argon*

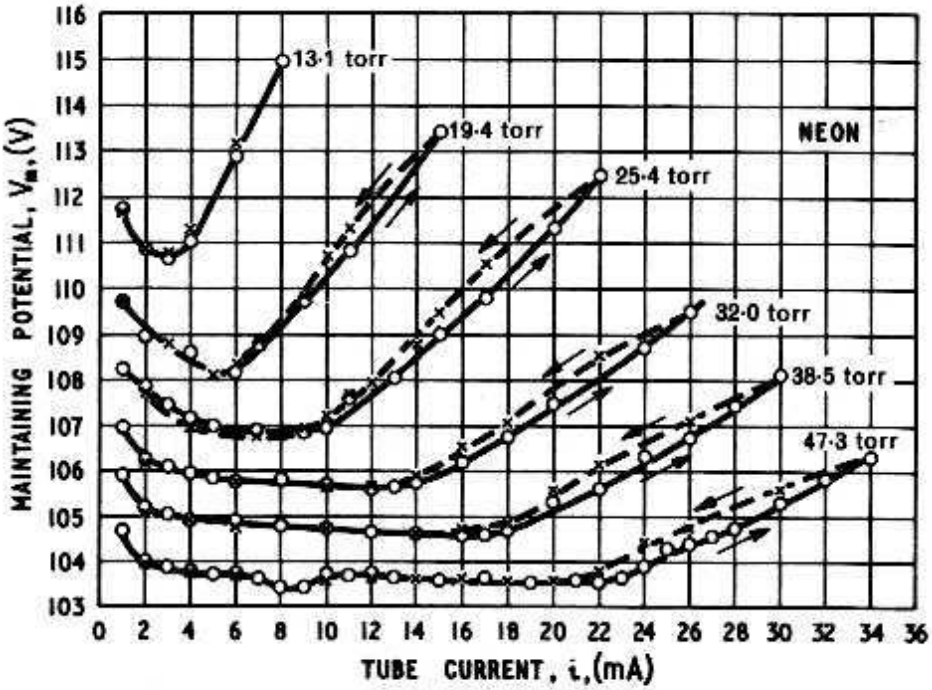


Fig. 3.19. Maintaining potential of a neon discharge tube with molybdenum electrodes as a function of current at various gas pressures. Jurriaanse[34]

for a glow discharge in neon at different pressures with molybdenum electrodes. By considering the discharge tube as consisting of two regions, the cathode region controlling the maintaining potential at temperature T_c and density ρ_c, and the rest of the tube at temperature T_e and density ρ_e, he showed that raising T_e increased the density ρ_c resulting in a reduction of the maintaining potential. Jurriaanse obtained reasonable agreement between experimentally determined and calculated results for a cylindrical electrode geometry although there is some doubt as to the correctness of the calculations due to typographical errors in some of the equations. A more rigorous treatment has been given by Smith[9]

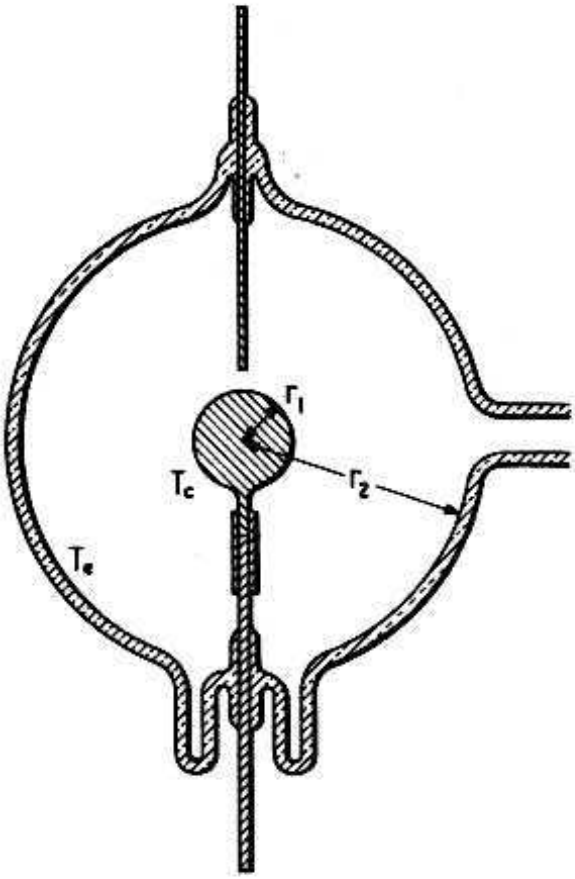


Fig. 3.20. Experimental spherical tube used by Smith[9] for measurement of the temperature coefficient of maintaining potential

using a spherical molybdenum cathode in a concentric spherical envelope (Fig. 3.20) to facilitate gas density calculations. It was assumed that the maintaining potential of a discharge with no anode fall present was dependent solely on the gas density adjacent to the cathode surface. This dependence was evaluated in a subsidiary experiment, by measuring V_m for various gas pressures at ambient envelope temperature (tube attached to the pump) and calculating the gas density at the cathode for each pressure. By calculating the gas density in the cathode area for various values of envelope temperature in sealed off tubes, the variation of V_m with wall temperature could be deduced. The agreement with experimental values is shown in Fig. 3.21. A similar although less exact analysis

was made for a cylindrical geometry by Benson and Chalmers[33] using pressure as the dependent parameter. Because of this, their calculations were incorrect but, as it was pointed out later by Benson and Rigg,[35] under certain conditions the use of pressure instead of density leads to only a small error in calculating the

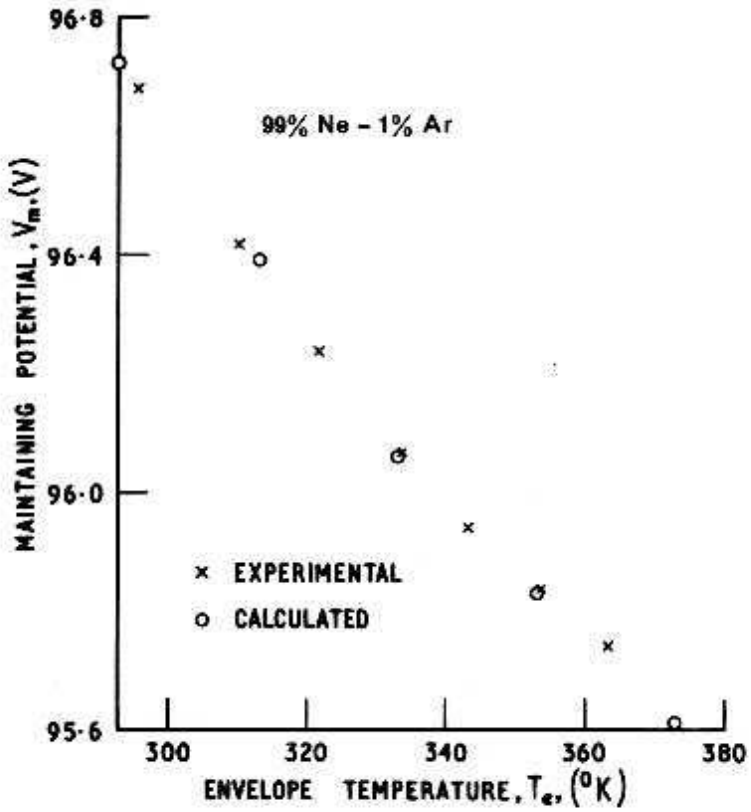


Fig. 3.21. Calculated and measured change of maintaining potential with change in envelope temperature for the spherical tube. Smith[9]

temperature coefficient. In this latter paper good agreement was obtained between the calculated and measured values of temperature coefficient in spite of the several approximations which had to be made because of the difficulty in evaluating the density distribution for the cylindrical geometry.

These theories which consider the cathode fall region only explain a negative temperature coefficient but not a positive one. It has been suggested that a positive coefficient might be caused by impurities, but it is now known to be caused by the formation of an anode fall in the tube. This was demonstrated by Smith[9] and is illustrated by Fig. 3.22. The top curves in this figure show the variation of density in the cathode and anode region ρ_c and ρ_a respectively as a function of the envelope temperature T_e, and the lower curves show the variation of cathode and anode falls as functions of the densities in the respective regions. If $\rho_{a_{crit}}$ is taken as the critical density in the anode region at which an anode fall commences then two situations can be considered (1) where ρ_a is

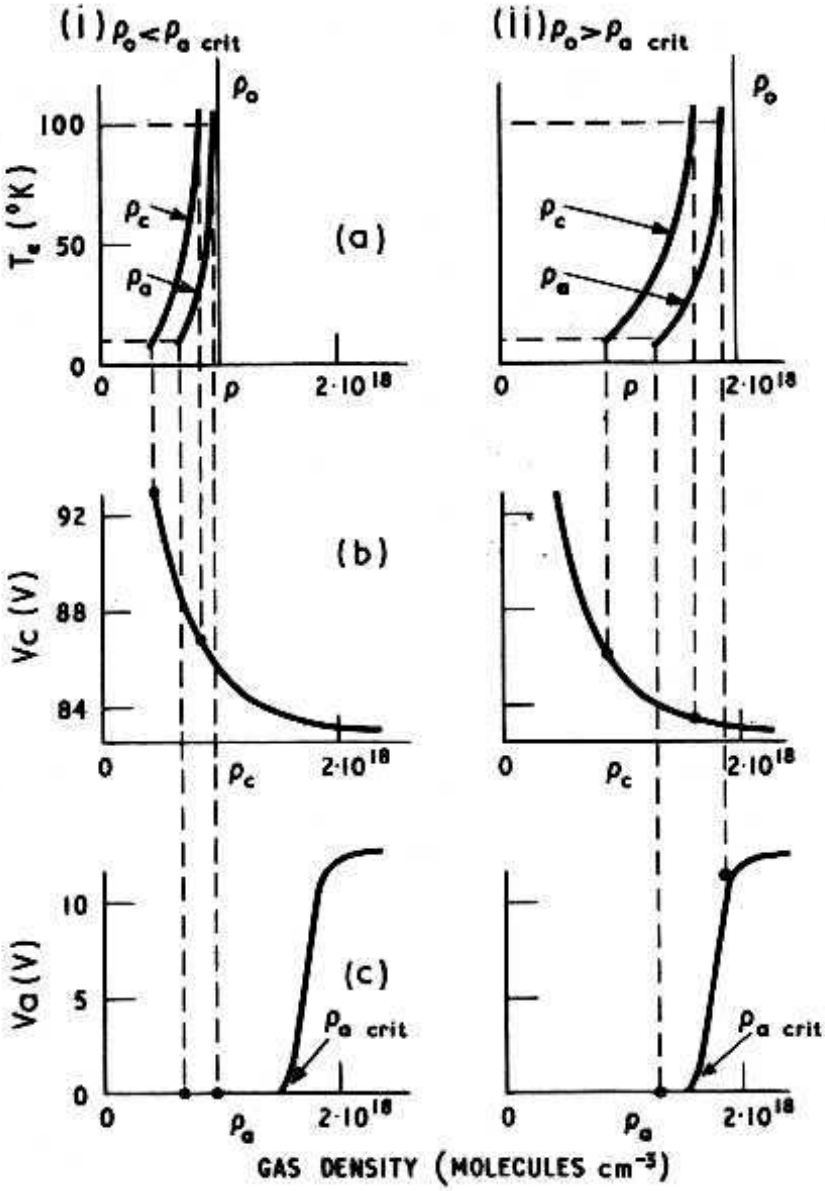


Fig. 3.22. Illustration of how the temperature coefficient can change from negative to positive due to anode fall formation. Smith[9]

always below $\rho_{a\text{crit}}$ and (2) where ρ_a goes above $\rho_{a\text{crit}}$. From Fig. 3.22 the average temperature coefficient for the range 10–100°K is therefore deduced as follows:

for $$\rho_a < \rho_{a\text{crit}}$$

$$\frac{(86{\cdot}6+0)-(93{\cdot}0+0)}{90} = -0{\cdot}07\text{V°C}^{-1}$$

for $$\rho_a > \rho_{a\text{crit}}$$

$$\frac{(83{\cdot}2+11{\cdot}5)-(85{\cdot}5+0{\cdot}0)}{90} = +0{\cdot}10\text{V°C}^{-1}$$

The low temperatures were chosen so that the broken vertical lines were well separated, nevertheless a similar result can be deduced for more practical temperature conditions. Smith con-

firmed the general deductions by measuring the temperature dependence of V_m in a tube in which the anode–cathode spacing could be varied. His resultant curves are shown in Fig. 3.23. Curves A and B are for the case where the anode–cathode distance, d, is small (equivalent to Case 1 of Fig. 3.22) and no anode fall occurs. Curve C is for the case where d is near the critical value for onset of an anode fall (equivalent to case 2 of Fig. 3.22) and D is where an anode fall is present.

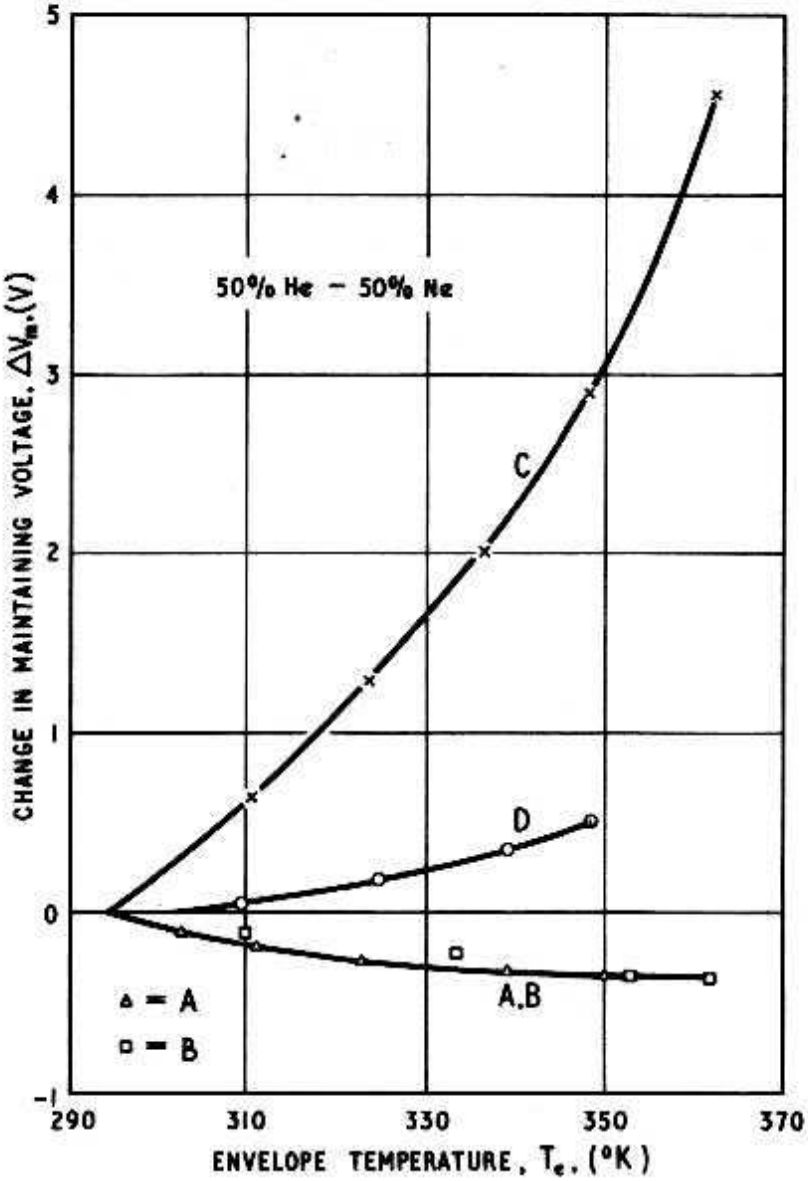


Fig. 3.23. Change in maintaining potential for 50% He + 50% Ne gas with molybdenum electrodes as a function of temperature for different gap settings. Curve A d = 0·1 cm, curve B d = 0·2 cm, curve C d = 0·4 cm and curve D d = 0·8 cm. Critical distance for onset of anode fall is about 0·4 cm

3.5. The Impedance of a Glow-discharge

In the normal glow region the voltage across a glow discharge does not vary appreciably with current, but it rises markedly in

the abnormal glow region. Therefore, under d.c. conditions a glow discharge can be represented as equivalent to a constant voltage in series with a resistor, R_g, the latter varying as the glow passes from normal to abnormal glow conditions. Although the V-i curve is not strictly linear in the abnormal glow region, over a reasonable working range the error introduced by assuming R_g is constant is small. Thus the voltage across the tube in this region can be expressed as

$$V_m = V_o + iR_g$$

where R_g depends on the gas composition and pressure, and on the cathode size and material and V_o is the intersection such a straight line would make with the voltage axis. Theoretically V_m depends on j/p^2 and thus it is expected that R_g will be inversely proportional to the cathode area $\times p^2$. In practice, the resistance is rather less dependent on these parameters due to the temperature effects discussed in the last section. For commercial tubes R_g is of the order of a few hundred ohms to a few kilohms.

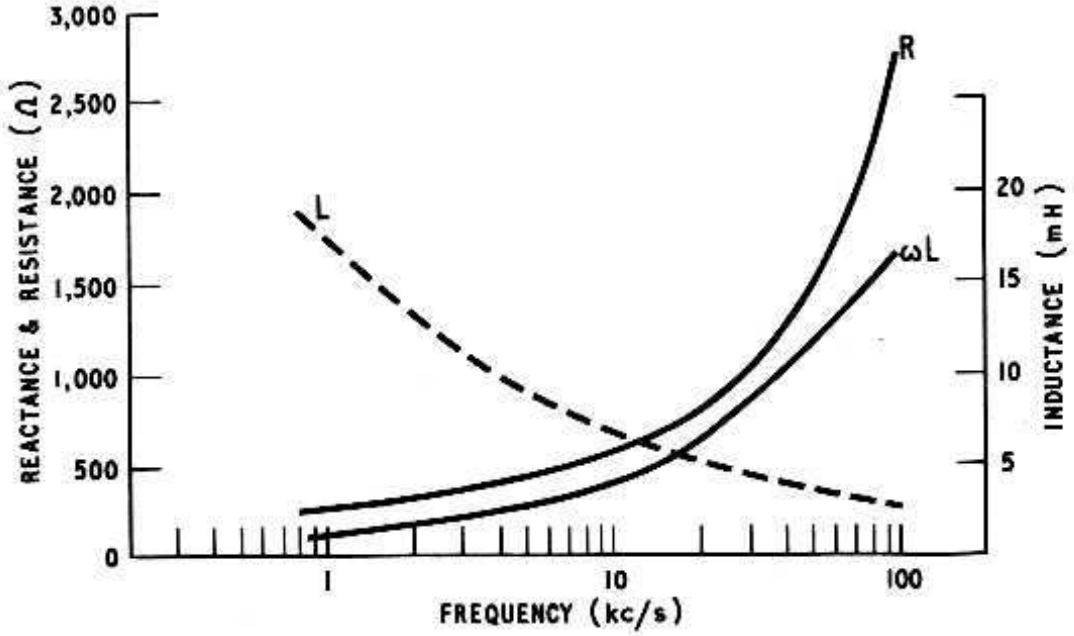


Fig. 3.24. Variation of resistive and inductive component of impedance with frequency for an 85A2 with a direct anode current of 10 mA

If an alternating signal is superimposed on the steady supply voltage, then, at frequencies up to about 1 megacycle, the a.c. impedance has an inductive component due to the inertia of the ionisation mechanism and of the ions themselves. The impedance at any one frequency can then be represented by a resistance and inductance in series. However both resistance and inductance vary with frequency as is shown in Fig. 3.24 where they are plotted for an

85A2 stabiliser. At frequencies above about 1 Mc/s the impedance becomes capacitive. The equivalent circuit valid over all frequencies is therefore complicated.

A number of results on commercial stabilisers have been published and attempts have been made to explain the data qualitatively in terms of the basic processes in a glow discharge.[36] An early theory by van Geel[37] was derived by considering small

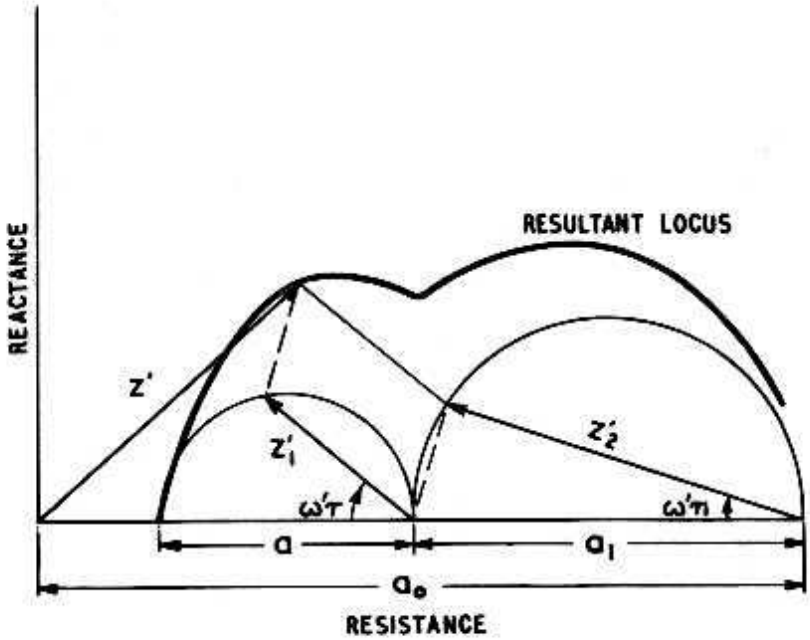


Fig. 3.25. Construction of an impedance locus on a complex diagram according to van Geel's theory for a discharge showing a delay mechanism. Z' is the impedance at frequency ω', which is the sum of the vectors Z'_1 and Z'_2

variations about the steady-state maintaining condition and by assuming that physical quantities in the discharge, such as space charge, could be regarded as functions of the instantaneous current and voltage alone. From this theory he deduced the impedance Z at low frequencies where self-capacity could be neglected as:

$$Z = \frac{R + j\omega L}{1 + j\omega\tau} \tag{3.13}$$

where R = the static resistance, L = the self-inductance, and τ = a time constant associated with the build-up of the discharge called the 'degression time'. The locus of the impedance against frequency according to Equation 3.13 can be represented on the complex diagram (reactance against resistance) by a semicircle on the real axis of diameter $a = L/\tau - R$. In a later paper van Geel[38] expanded the theory, and postulated the form of the impedance locus when a secondary mechanism having a delay time large

compared to τ was present in the discharge. The impedance then has the form

$$Z = a_o - \frac{a}{1+j\omega\tau} - \frac{a_1}{1+j\omega\tau_1} \tag{3.14}$$

where τ_1 is associated with the time constant of the after effect. The impedance locus in this case can be represented by a sum of two vectors whose individual loci are semicircles, with centres on the real axis, of diameter a and a_1 respectively. This is illustrated in Fig. 3.25 where the impedance at frequency $2\pi\omega'$ is constructed. Impedance loci of this type predicted by van Geel's theory have been verified qualitatively by the experimental work of Verhagen.[39] More recently Benson and Chalmers[40] have extended the measurements to higher frequencies, and have shown the effect of argon

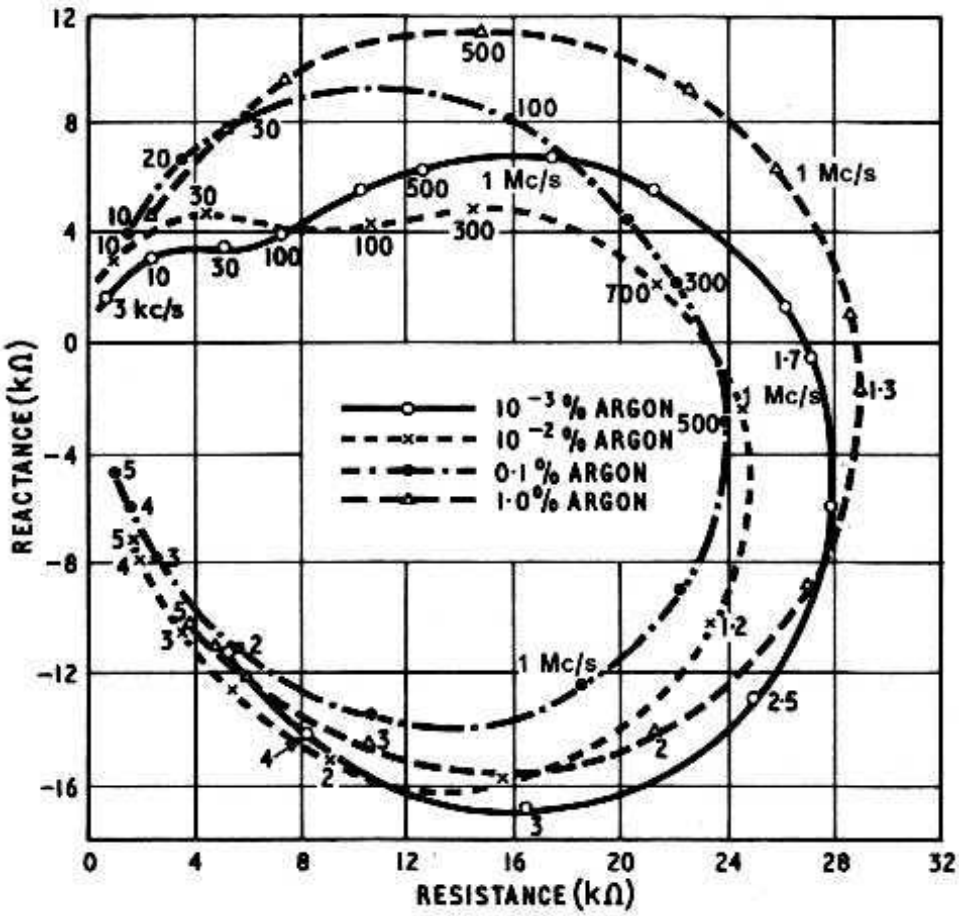


Fig. 3.26. Impedance loci for tubes containing neon-argon mixtures according to Benson and Chalmers[40]

admixtures in neon on the impedance diagram. Typical impedance loci for neon-argon mixtures are shown in Fig. 3.26 for a cylindrical geometry tube with no anode fall. The effective capacity C_d of the tube is not taken into account in van Geel's theory and thus to compare their experimental results Benson and Chalmers had to

subtract C_d from their measurements. Replotting the resultant locus for neon showed that the curve could be represented by a series of vectors whose individual loci are semicircles each having a different time constant τ, see Fig. 3.27. Thus the results are in accordance

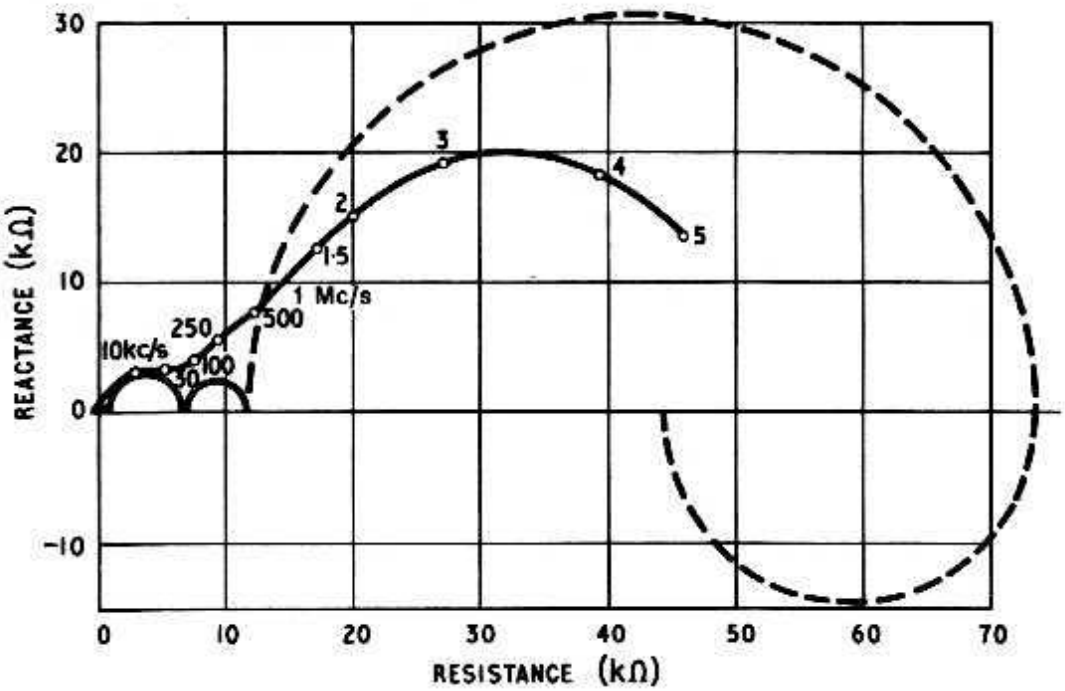


Fig. 3.27. Modified impedance locus for a neon-filled tube after correcting for the capacity C_d, showing the component semi-circular loci. Benson and Chalmers[40] $\tau = 300$ *μsec*, $\tau_1 = 13$ *μsec*, $\tau_2 = 0{\cdot}64$ *μsec*, $\tau_3 = 0{\cdot}04$ *μsec*, $\tau_4 = 0{\cdot}02$ *μsec*. $a = 1$ *kΩ*, $a_1 = 6$ *kΩ*, $a_2 = 5{\cdot}5$ *kΩ*, $a_3 = 62$ *kΩ and* $a_4 \simeq -30$ *kΩ*

with van Geel's theory if several delay effects are assumed. The impedance is then of the form

$$Z = a_o - \frac{a}{1+j\omega\tau} - \frac{a_1}{1+j\omega\tau_1} \cdots \frac{a_n}{1+j\omega\tau_n} \tag{3.15}$$

and the equivalent circuit is shown in Fig. 3.28. Reed and Dix[41], who made similar measurements but extended the frequency range to 10 Mc/s found that as f approaches ∞ the curve did not pass through the origin as inferred from the equivalent circuit of Fig. 3.28. They therefore concluded that the capacitor had an effective resistance in series which they suggested represented the dielectric loss due to the gas molecules.

Ahsmann and Oskam[42] also investigated the effect of gas mixtures, and found that the self-inductance defined by Equation 3.13 was related to the ion transit time across the cathode fall and negative glow. On the other hand Yeh[43] has related the ion transit time t_i to the frequency f_o at which the tube passes from inductive to capacitive, i.e. the intercept on the real axis in

Fig. 3.26, giving $t_i = 1/(2f_o)$. Evidence in support of this has been given by Ward and Schneekloth.[44]

The author[45] has investigated the effect of the d.c. current and the gas pressure on the impedance in the audio-frequency range for stabiliser tubes of the 85A2 type. Both the resistance R and reactance X decreased with increasing current at constant pressure

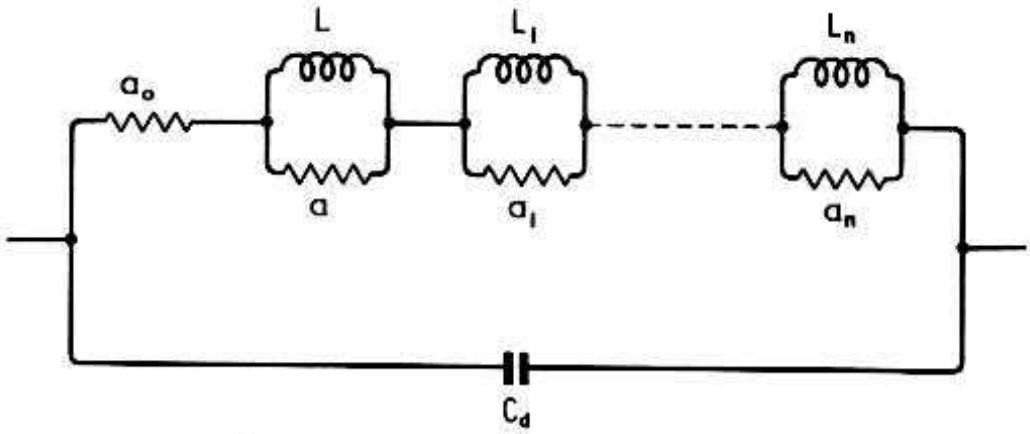


Fig. 3.28. Equivalent circuit of a glow discharge tube

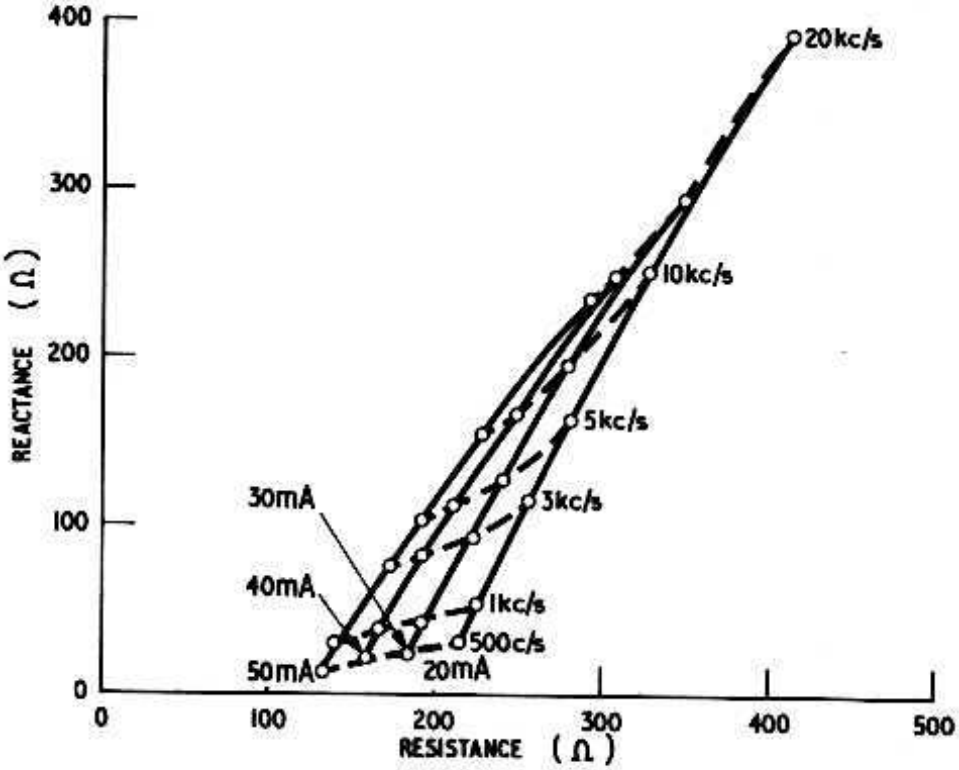


Fig. 3.29. Impedance loci for tubes filled with 99% neon + 1% argon at 30 torr pressure and with molybdenum cathodes, for different direct currents

as shown in the impedance loci diagram of Fig. 3.29. Similarly, R and X decreased with increasing pressure at constant current, Fig. 3.30. Provided the glow completely covered the cathode the reactive component at any one frequency varied inversely as the current, whereas the resistive component varied inversely as the

square root of the current. No such simple laws were found for the variation of R and X with pressure, although they both tended to depend on $1/p$ at high pressures and $1/p^2$ at lower pressures. However, examining the results in the light of van Geel's theory showed that if the impedance was expressed in terms of self-inductance and degression time (L and τ according to Equation 3.13) then L is proportional to $1/p$. This accords with the premise that L is related to the ion transit time, if a simplified theory such as that of von Engel and Steenbeck is assumed for the cathode fall. The effect of cathode material was also investigated using nickel, molybdenum and zirconium cathodes. The reactance was found to be almost independent of the cathode material showing that it is mainly a gas effect, whereas the resistance increased with

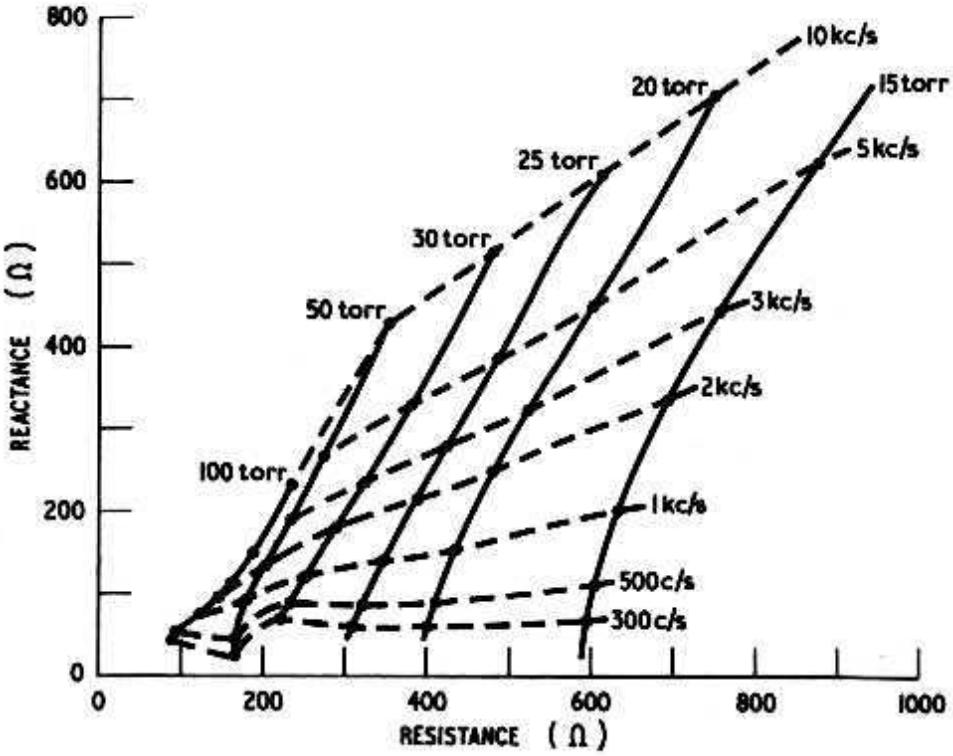


Fig. 3.30. Impedance loci for tubes filled with 99% neon + 1% argon and with molybdenum cathodes at different gas pressures. Tube current = 8 mA

increasing maintaining potential and was thus affected by the γ of the surface. The effect of pressure, current and cathode material on the impedance loci has also been examined by Benson and Bradshaw,[46] extending the frequency range up to 70 Mc/s. The loci were roughly circular, intersecting the real axis at a frequency of the order of 2 Mc/s similar to Fig. 3.26. They found that in general the diameter of the circles increased with decreasing pressure and current, and in the latter case the intercept at $X = 0$ (2 Mc/s) was approximately proportional to $1/i$. Analysis of the impedance loci gave a_n and L_n values which were also proportional to $1/i$.

Similar, though rather more complex equations to those obtained by van Geel, have been derived by Müller[47] from a more systematic calculation using the basic equations of the cathode fall region. Severin[48] has calculated the impedance of a discharge based on the motion of ions and electrons assuming a linear field and constant mean free path for the ions, and further assuming that when an electron attains a potential energy of eV_i it immediately ionises an atom. Nevertheless, with these assumptions he obtained an equivalent circuit which results in an impedance-frequency locus corresponding to the elongated circles obtained by Benson and his colleagues, i.e. Fig. 3.26. Since these two theories are relatively new a critical analysis in the light of experiment has yet to be made.

3.6. Noise in a Glow Discharge

Noise generated in a glow discharge can be associated with a number of causes such as movement of the glow, oscillations in the positive column, electro-magnetic pick-up of external signals, etc. In many cases, by suitable design, these can be avoided. However, in common with all electronic devices, there is a minimum noise which cannot be eliminated and which arises from the particle nature of the current flow. Two effects can be recognised (1) fluctuations of current due to the statistical random times of production of charged particles and (2) fluctuations in current due to their thermal agitation. The two effects, referred to as 'shot' noise and 'thermal' noise respectively, result in a background of spurious signals which are relatively uniformly spread over all frequencies.

Since 'shot' noise is normally associated with the random emission of electrons in a thermionic valve, it is common to express it in terms of the equivalent saturated diode current i_{eq}.

The r.m.s. noise current i_N is given by

$$i_N = \sqrt{(2ei_{eq}\Delta f)} \tag{3.16}$$

where e is the electronic charge and Δf the frequency bandwidth. To measure i_{eq} the device is connected in parallel with a tuned circuit. A noise diode (thermionic valve), also connected in parallel with the circuit, is then switched on, and the diode current adjusted such that the noise output power is doubled. The diode current is then i_{eq}. Thermal noise is normally associated with a conductor and is expressed in terms of the equivalent noise resistance R_N.

It is considered as an effective series generator with a mean square voltage V_N given by

$$V_N = \sqrt{(4kTR_N\Delta f)} \tag{3.17}$$

where k is Boltzmann's constant and T the absolute temperature of the device.

Both effects occur in a glow discharge, and in the case of shot noise both ion and electron production in the gas and electron production at the cathode will contribute. It is not possible to distinguish between the two effects, and thus the noise in a gas discharge can be expressed either in terms of a zero impedance noise source with an e.m.f. of V_N in series with the impedance Z, or as an infinite impedance with a short circuit current i_N in parallel with Z.

$$i_N = \frac{V_N}{Z} \tag{3.18}$$

A number of investigators have measured noise in stabiliser tubes,[49, 50] and both methods of expressing the noise have been used. In general the mean noise expressed as V_N is of the order of a few tens of microvolts and decreases with increasing current.

Normally shot and thermal noise is 'white', the amplitude being independent of the frequency. Van der Ziel and Chennette,[50] however, showed that in a gas discharge i_{eq} decreased with frequency over the range of 100 c/s to 10 Mc/s, which he explained in terms of the multiplication processes in the gas. Since these multiplication processes follow any change in the applied voltage, whether sinusoidal or random, i_{eq} due to shot noise would be expected to decrease as the frequency increased to values where some of the multiplication processes become inoperative. In support of this, van der Ziel and Chennette also measured the admittance Y and found that the curves of $(Y)^2$ and i_{eq} against frequency were very similar in shape for frequencies from 50 c/s to 100 kc/s but that there was no direct relationship above 1 Mc/s. At the higher frequencies, the admittance becomes capacitive and rises steeply due to the predominance of displacement current. Reed and Dix[41] pointed out that since the noise amplification was only concerned with convection current, one should only expect agreement between i_{eq} and the convection-current component of the admittance. At very high frequencies Reed and Dix found that the noise was fairly constant and corresponded closely with that from a noise diode of the same space current. They therefore concluded that the tubes gave full shot noise at the high frequencies which was amplified at the lower

frequencies due to the multiplication processes. From this they computed i_N in terms of the tube current, maintaining potential and convection-current component of impedance. The agreement of the computed curves with measured values is shown in Fig. 3.31 for an 85A2 reference tube. Barker and Benson,[51, 52] who have measured both i_N and V_N for neon and neon-argon mixtures over a frequency range of 1 kc/s to 8 Mc/s at various gas pressures, found that, although i_N versus frequency was congruent with the variation of the convection-current component of admittance over the range of 1 kc/s to 1 Mc/s, it did not agree so well above this frequency. They also found that the variation of noise with pressure was frequency dependent; at low frequencies the noise voltage decreased with increasing pressure, whereas at higher frequencies above 500 kc/s it increased. From their measurements they concluded that the theory of van der Ziel and Chennette, modified by Reed and Dix was substantiated, but that the computations of the latter workers could only be taken as approximately true.

In an effort to locate the site of the noise mechanism Galbraith and van der Ziel[53] have measured the spectral noise at a probe

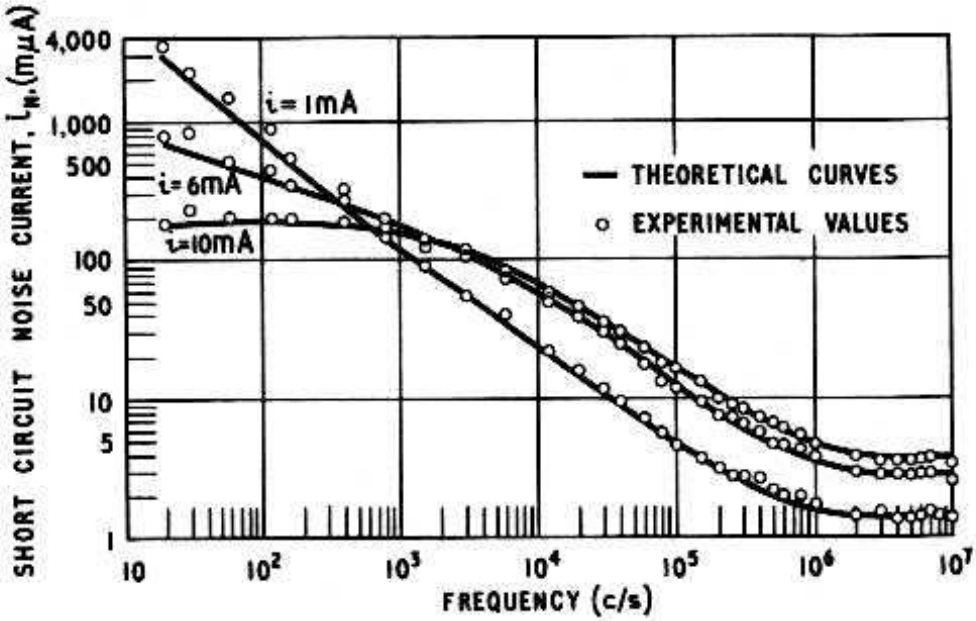


Fig. 3.31. R.m.s. noise current in a 6 kc/s bandwidth as a function of frequency. Reed and Dix[41]

inserted in a neon discharge at 2 torr pressure. They found that i_{eq} had a minimum at about 2 to 5 Mc/s which was dependent on the tube current; this is shown in Fig. 3.32. The results indicated that two noise mechanisms were operating in addition to the shot noise: a low frequency mechanism decreasing in intensity with increasing frequency, and a high frequency mechanism increasing in intensity with increasing frequency. By selected shielding of the discharge

with metal foil a.c. connected to earth, and by using more than one probe they were able to locate roughly these noise sources. The low frequency noise was predominantly located in the negative glow but also present in the Faraday dark space; this is probably noise due to the multiplication processes. The high frequency noise

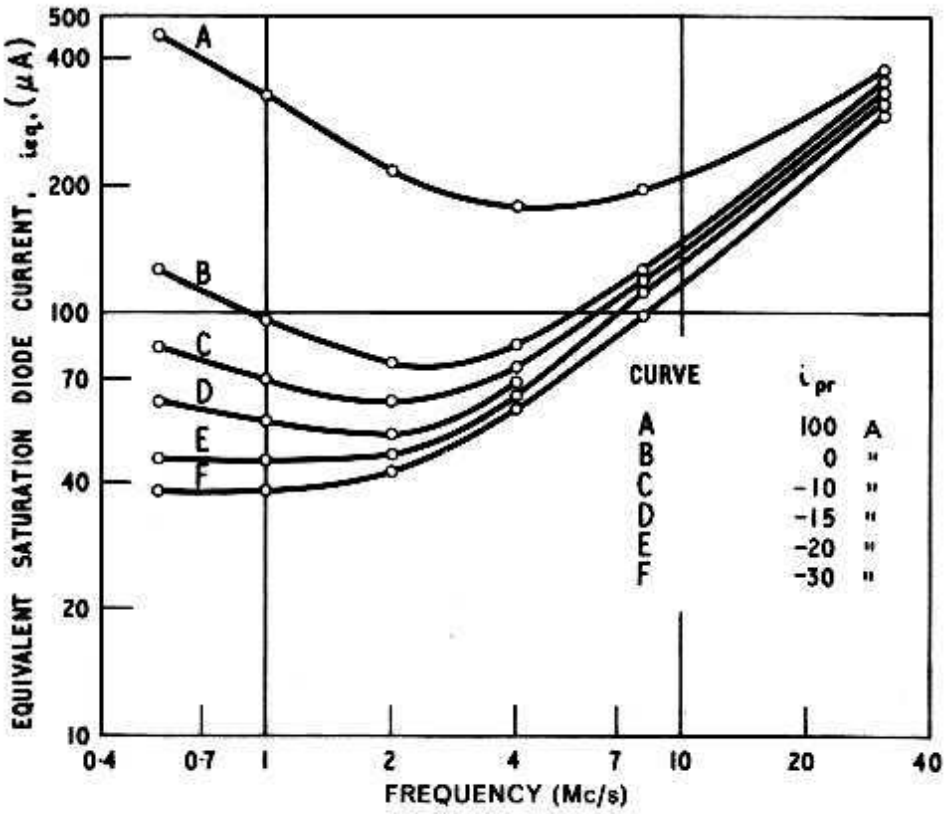


Fig. 3.32. Equivalent saturated diode current measured at a probe in the discharge as a function of frequency with varying probe currents. Galbraith and van der Ziel[53]

appeared mainly in the positive column and may possibly have been associated with the ion sheath region at the wall in the vicinity of the probe. Since the previous experiments reported were in discharges without a positive column, the high frequency noise would not have been observed.

3.7. The Hollow Cathode Discharge

The negative glow zone of a discharge is a source of positive ions, metastable atoms and photons, all of which are capable of releasing further electrons at the cathode. Although most of the ions drift to the cathode under the influence of the field, the uncharged metastables and photons travel in all directions and, probably in most tubes, less than half will impinge on the cathode. If the cathode is a plate and a second plate cathode is brought up close to it such that the two negative glows coalesce, the fraction

of metastables and photons impinging on the cathodes will approach unity and the γ will be enhanced. In addition each electron will effect more collisions, both ionising and exciting, before reaching the anode, thus increasing the efficiency of the discharge further. As a result the current density is orders of

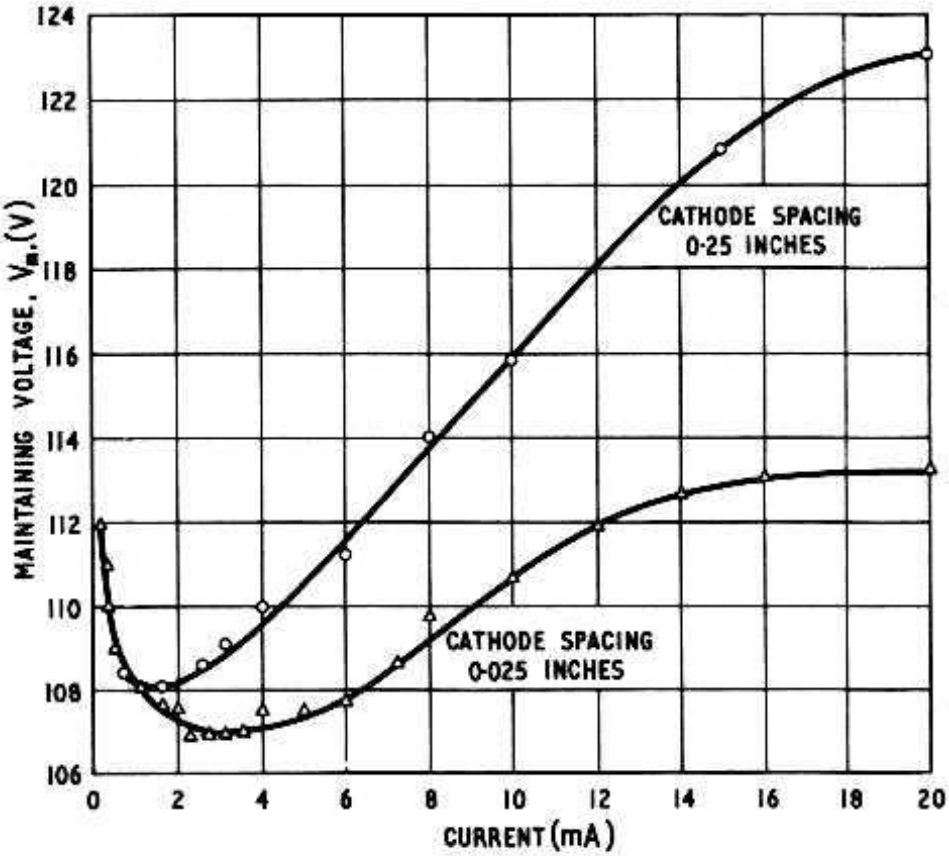


Fig. 3.33. Voltage-current characteristic of parallel-plate cathodes at two different spacings, showing the hollow cathode characteristics

magnitude larger than for a conventional discharge at the same potential, or at the same current the discharge is maintained at a lower voltage; this is shown in Fig. 3.33. The phenomenon is known as the hollow cathode effect and occurs with any cathode in which the negative glow from one part of the surface overlaps that from another, for example a U-shaped cathode or a hollow cylinder. Further, if a cathode has only a part of its surface which is hollow, the discharge will transfer into the hollow, leaving the rest of the cathode uncovered.

The efficiency of the hollow cathode effect depends on geometry, gas pressure, and gas composition, but although the effect is utilised in a number of applications the data on these parameters are limited. Badareu and his associates[54, 55] have examined the hollow cathode effect between two plate cathodes in air and mercury vapour, mainly to investigate the processes involved rather than to collect data for the two gases. More recently Ciobotaru[56] from the

same laboratory has extended the work to derive some simple empirical laws using xenon and krypton gas. The current as a function of the pressure and cathode spacing, a, is shown in Fig. 3.34 for a xenon discharge at 450V. At sufficiently high pressure and/or inter-cathode distance the hollow cathode discharge tends to degenerate into two conventional discharges as might be expected. Under these conditions for a given cathode fall j/p^2 = a constant which is in accordance with the general theories of a

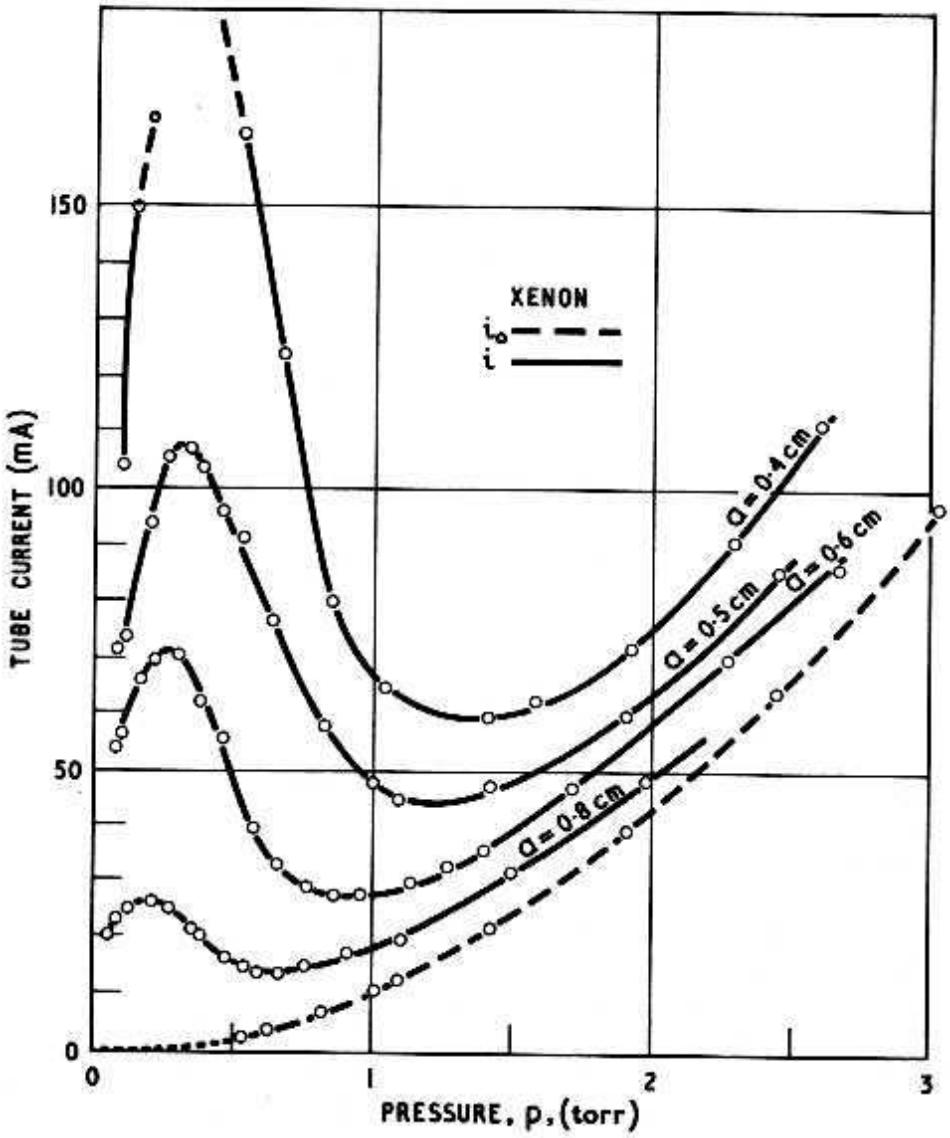


Fig. 3.34. Variation of current at a simple (i_o) and double (i) cathode as a function of the pressure of xenon for various inter-cathode spacings. Ciobotaru[56]

glow discharge (Section 3.2.1). On the other hand for sufficiently low values of pa the hollow cathode effect develops and under these conditions the current increases inversely as the square root of the pressure. This is in accordance with Little and von Engel's theory[4] (Section 3.2.1), since it has been found that, in the hollow cathode

region, the cathode dark space d_k remains constant with varying pressure and thus the equation

$$j/p^2 \propto \frac{V_c^{\frac{3}{2}}}{(pd_k^{\frac{5}{2}})}$$

reduces to $j \propto 1/p^{\frac{1}{2}}$ for constant V_c.

The dependence of i^2p on V_c and a for three gases is given in Table 3.6. From measurements of the amplification $q = i/i_o$ when the hollow cathode effect was fully developed, Ciobataru also found that the graph of log q against log p has a universal slope of value 2·5.

Table 3.6

i^2p in terms of V_c and a for krypton, xenon, and mercury in a hollow cathode discharge according to Ciobotaru[56]

Gas	*Cathode fall* V_c (V)	*Inter-cathode space a* (cm)	$i^2p \times 10^{-2}$ (mA^2 *torr*)
Krypton	350	0·6	26·4
Krypton	350	0·6	12·1
Krypton	400	0·6	106
Krypton	400	0·8	23·7
Xenon	400	0·4	20·5
Xenon	400	0·5	3·02
Xenon	400	0·6	1·32
Xenon	450	0·8	1·33
Mercury	700	0·45	2·87
Mercury	800	0·85	0·77

Similar measurements with a double-plate cathode have been carried out by Sturges and Oskam[57] in an attempt to discover whether one gas exhibited the hollow cathode effect more strongly than another. Using helium and neon, they found that although the maintaining potential for neon was lower than for helium at a given current the actual reduction of voltage from that obtained in a conventional discharge was greater for the helium. They also discovered that for any given pressure and discharge current, there existed two or more optimum values of cathode separation for which the tube voltage had a minimum value. The presence of more than one optimum distance is due to the reluctance to merge of the two negative glow regions. Although the extent of the hollow cathode effect as a function of the gas composition has not been established there is no doubt that mixtures of gases showing the Penning effect do not exhibit the hollow cathode effect nearly so

strongly as the pure gases.[58] This is probably due to the destruction of the metastables of the parent gas by the admixed atoms.

3.7.1. THE IMPEDANCE OF THE HOLLOW CATHODE DISCHARGE

A feature of the hollow cathode discharge is that by suitable choice of gas pressure, geometry, etc., a negative V-i characteristic can be obtained, which, unlike the transition from Townsend to glow discharge, is stable. A particular hollow cathode which exhibits this

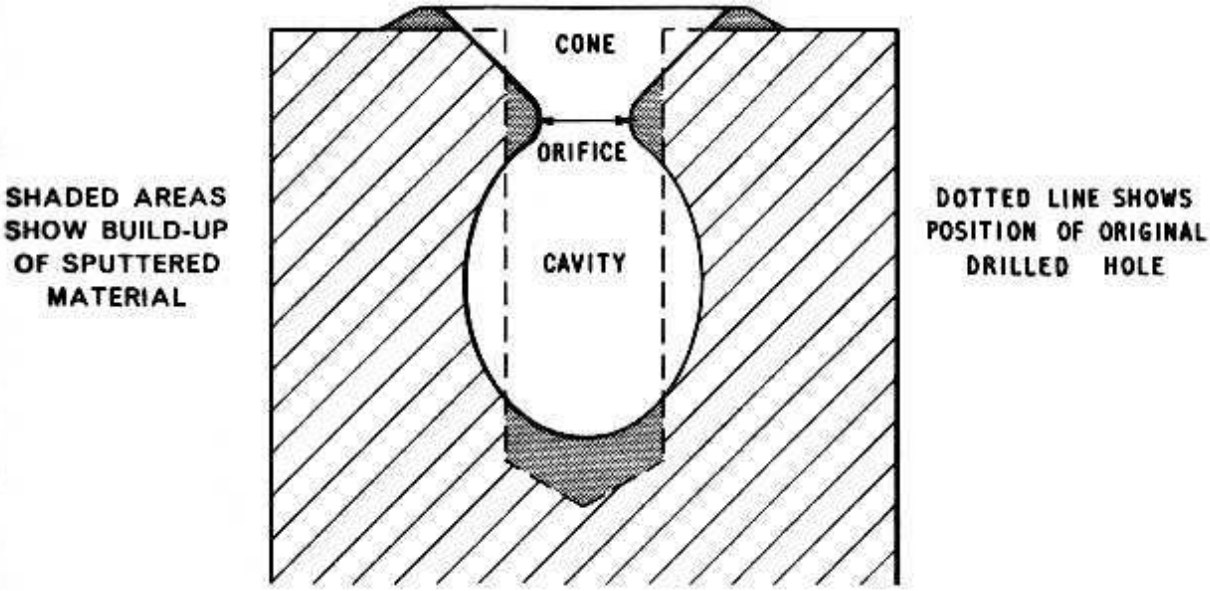


Fig. 3.35. Cross section of a cavity cathode

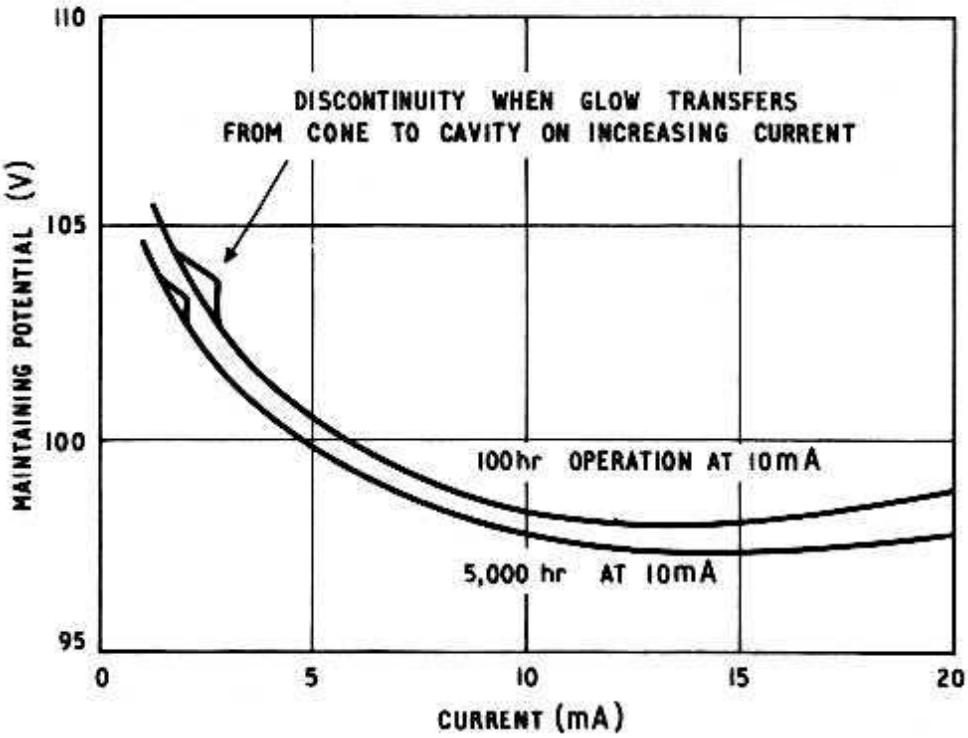


Fig. 3.36. Static voltage-current characteristic of a typical cavity cathode diode

effect rather markedly is one described by Townsend[59] as the 'cavity cathode'. Such a cathode can be formed from a cylindrical hole drilled in a piece of metal which is subsequently sputtered at a fairly high current. By this method the hole sputters into a spherical cavity with a restricted orifice, see Fig. 3.35. An equilibrium state is eventually reached when the glow completely covers the cavity wall and further sputtering only takes place within the cavity.

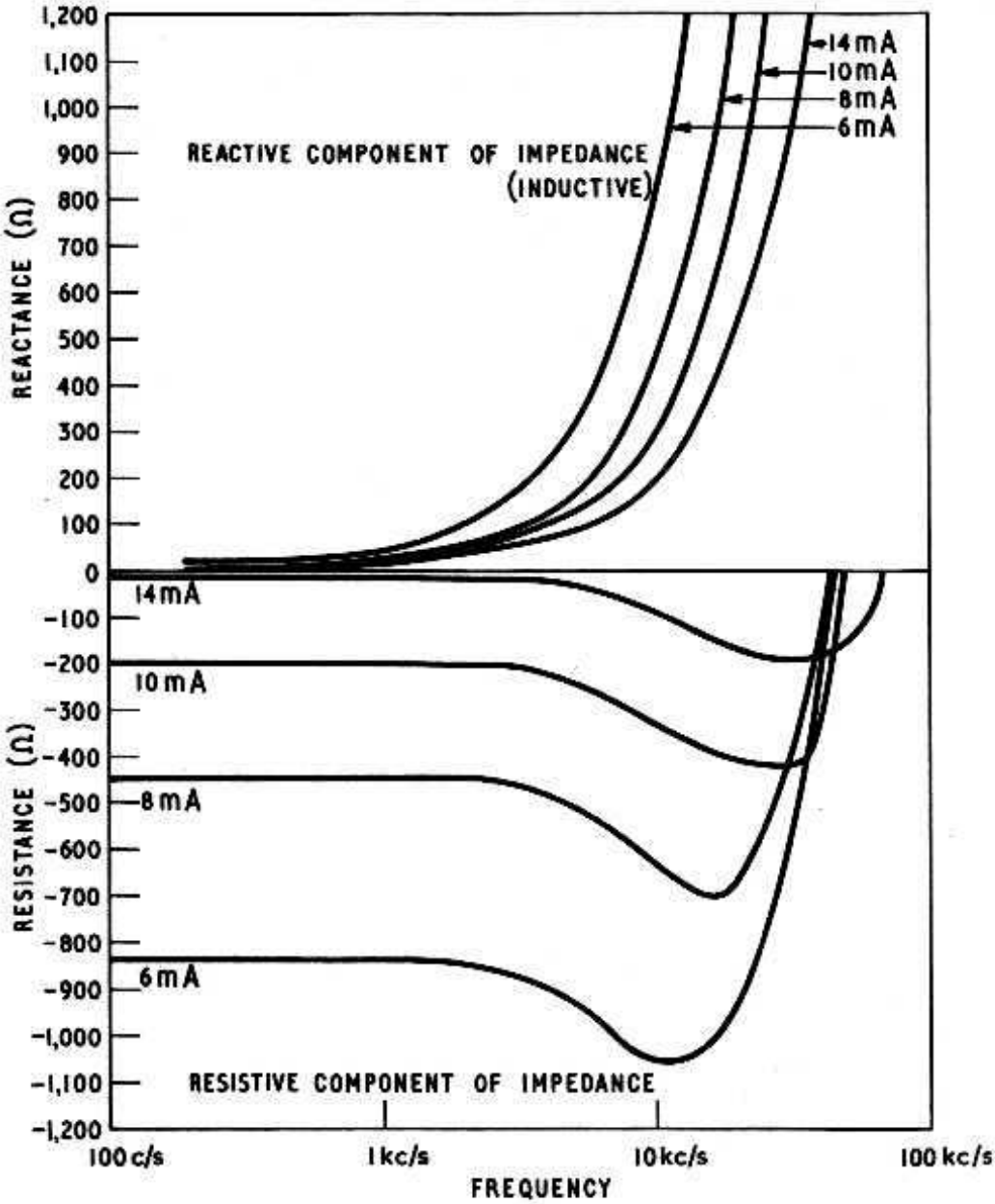


Fig. 3.37. Reactance and resistance of a typical cavity cathode tube as a function of frequency and for various direct currents through the tube

Since the orifice area is only 1/50th of the total surface area very little sputtered metal escapes. The cavity size for stable operation depends on the pressure and the running current. A typical static *V-i* characteristic is given in Fig. 3.36 showing that the negative resistance is quite high (> 1kΩ) at low currents, decreasing to become positive at about a few tens of ohms at higher currents.

For an alternating signal superimposed on the maintaining potential the resistive component remains negative over most of the audio-frequency range, although above 10 kc/s the reactive component becomes large enough to exceed the resistance; this is shown in Fig. 3.37. Thus for speech passage, a tube with such a cathode will give a gain, and is of interest as a possible switch in telephone circuits.

REFERENCES

1. FRANCIS, G., 'The glow discharge at low pressure', *Handbuch der Physik*, **22**, Springer-Verlag, Berlin, (1956)
2. VON ENGEL, A. and STEENBECK, M., *Electrische Gasentladungen*, Springer, Berlin, (1932)
3. DRUYVESTEYN, M. J., *Physica*, **5**, 875, (1938)
4. LITTLE, P. F., and VON ENGEL, A., *Proc. Roy. Soc.*, **A224**, 209, (1954)
5. WARD, A. L., *J. Appl. Phys.*, **33**, 2789, (1962)
6. LUCAS, J., *J. Electron Control*, **11**, 281, (1961)
7. MULLER, K. G. and TENSHERT, G., *Z. Physik*, **178**, 319, (1964)
8. SEVERIN, P. J. W., *Thesis*, Delft University, (1964)
9. SMITH, J., *Brit. J. Appl. Phys.*, **13**, 399, (1962)
10. BÄR, R., *Handbuch der Physik*, **14**, Springer, Berlin, p. 303, (1927)
11. PENNING, F. M. and MOUBIS, J. H. A., *Philips Res. Rep.*, **1**, 119, (1946)
12. JURRIAANSE, T., PENNING, F. M. and MOUBIS, J. H. A., *Philips Res. Rep.*, **1**, 225, (1946)
13. WESTON, G. F., *Brit. J. Appl. Phys.*, **10**, 523, (1959)
14. WESTON, G. F., *Brit. J. Appl. Phys.*, **11**, 305, (1960)
15. MILOVANOVA, R. A. and CHISTYAKOV, P. N., *Soviet Phys. Tech. Phys.*, **8**, 262, (1963)
16. PENNING, F. M. and MOUBIS, J. H. A., *Physica*, **15**, 721, (1949)
17. TAKEISHI, Y., *J. Phys. Soc. Japan*, **13**, 767, (1958)
18. TAKEISHI, Y., *J. Phys. Soc. Japan*, **11**, 676, (1956)
19. AHSMANN, G. J. M. and VAN GELDER, Z., *Le Vide*, **15**, 226, (1960)
20. BENSON, F. A. and RIGG, B., *Electron. Eng.*, **33**, 524, (1961)
21. BRAUN, A. and BUSCH, G., *Helv. Phys. Acta*, **20**, 33, (1947)
22. DILLON, J. A., *Bull. Am. Phys. Soc.*, **3**, 31, (1958)
23. LEMAIGRE-VOREAUX, M. P., *Bull. Soc. Franc. Elect.*, **2**, 331, (1961)
24. *Handbook of Chemistry and Physics*, Chemical Rubber Publishing Co., 46th Edition, Ohio, U.S.A. (1965–66)
25. HALL, R. F. and STOCKER, B. J., *Rad. Electron. Engr.*, **28**, 313, (1964)
26. MEISSNER, J., *Z. Physik*, **117**, 325, (1941)
27. BENSON, F. A. and BURDETT, G. P., *Proc. Inst. Elec. Engrs.*, **B108**, 501, (1961)
28. BENSON, F. A., *Proc. Intern. Conf. Ionisation Phenomena Gases*, 5th, Munich, North Holland Publishing Co., Amsterdam, 235, (1962)
29. FROUWS, S. M., *Proc. Intern. Conf. Ionisation Phenomena Gases*, 3rd, Venice, Soc. Italiana di Physica, 341, (1957)
30. KLARFELD, B., *J. Tech. Phys. U.S.S.R.*, **4**, 44, (1937), **5**, 725, (1938) (from ref. 1).
31. GÜNTHERSCHULZE, A. and KELLER, F., *Z. Physik*, **81**, 799, (1933)
32. SMITH, J., *Brit. J. Appl. Phys.*, **9**, 122, (1958)
33. BENSON, F. A. and CHALMERS, P. M., *Proc. Inst. Elec. Engrs.*, **C106**, 82, (1958)
34. JURRIAANSE, T., *Philips Res. Rep.*, **1**, 407, (1946)

35. BENSON, F. A. and RIGG, B., *Proc. Inst. Elec. Engrs.*, **110**, 29, (1963)
36. WILLIAMS, M. O., *Strowger J.*, **8**, 106, (1952)
37. VAN GEEL, C., *Physica*, **6**, 806, (1939)
38. VAN GEEL, C., *Thesis*, Delft University, (1955)
39. VERHAGEN, C. J. D. M., *Thesis*, Delft University, (1942)
40. BENSON, F. A. and CHALMERS, P. M., *Proc. Inst. Elec. Engrs.*, **B107**, 199, (1960)
41. REED, K. B. and DIX, J. F., *Electron. Technol.*, **39**, 31, (1962)
42. AHSMANN, G. J. M. and OSKAM, H. J., *J. Appl. Phys.*, **29**, 1768, (1958)
43. YEH, C., *J. Appl. Phys.*, **27**, 98, (1956)
44. WARD, A. L. and SCHNEEKLOTH, L. G., *Paper presented to 14th Annual Gaseous Electronics Conference*, Schenectady, U.S.A., (1961)
45. WESTON, G. F., *J. Electron. Control*, **12**, 379, (1962)
46. BENSON, F. A. and BRADSHAW, M. W., *Radio Electron. Engr.*, **29**, 313, (1965); *Proc. Inst. Elec. Engrs.*, **113**, 62, (1966)
47. MÜLLER, K. G., *Z. Physik*, **164**, 40, (1961)
48. SEVERIN, P. J., *J. Electron. Control*, **16**, 381, (1964)
49. BACHE, H. and BENSON, F. A., *Electron. Engr.*, **24**, 328, (1952)
50. VAN DER ZIEL, A. and CHENETTE, E. R., *Physica*, **23**, 943, (1957)
51. BARKER, K. and BENSON, F. A., *Proc. Inst. Elec. Engrs.*, **113**, 937, (1966)
52. BARKER, K. and BENSON, F. A., *Electronic Letters*, **2**, 193, (1966)
53. GALBRAITH, A. R. and VAN DER ZIEL, A., *Proc. Intern. Conf. Ionisation Phenomena Gases, 4th, Uppsala* N. Holland Publishing Co., Amsterdam, **1**, 297, (1959)
54. BADAREU, E. and POPESCU, I., *J. Electron. Control.*, **4**, 503, (1958)
55. BADAREU, E. and WÄCHTER, F., *J. Electron. Control.*, **4**, 539, (1958)
56. CIOBOTARU, D., *J. Electron. Control.*, **17**, 529, (1964)
57. STURGES, D. J. and OSKAM, H. J., *J. Appl. Phys.*, **35**, 2887, (1964)
58. ALLIS, W. P., *Proc. Intern. Conf. Ionisation Phenomena Gases, 3rd, Venice*, Societa Italiana di Fisica, 50, (1957)
59. TOWNSEND, M. A., *Bell System. Tech. J.*, **36**, 755, (1957)

CHAPTER 4

Cathodic Sputtering

4.1. Introduction

It is well known that during the life of a glow discharge tube, cathode material deposits on to the glass envelope. The effect is known as sputtering, and is due to the ejection of particles of the cathode material as a result of bombardment by gas ions from the discharge. Sputtering can be detrimental if light emission is required from the discharge tube, such as in indicator tubes and discharge lamps, but can be advantageous in some tubes as a method of cleaning the cathode and preventing contamination. In general the particles are uncharged atoms which pass through the gas by a diffusion process so that the amount of material deposited on a surface will be dependent on pressure as well as the total ion density and time.

In a glow discharge the ions can have any energy between zero and that corresponding to the cathode fall, and can impinge on the cathode at any angle. Further, at pressures above a few torr, the amount of material diffusing back to the cathode will be much greater than the amount escaping. Thus the sputtering process in a glow discharge is complex and measurements using such a discharge give little information about the actual mechanism occurring at the cathode.

For a better understanding of the sputtering process measurements must be made at lower pressures where back diffusion can be neglected, and where the ion energy and current can be varied independently. These conditions can be satisfied either by immersing the target as a negative probe in an arc discharge or by bombarding the target with a positive ion beam. A third method which has been used is to extend the glow discharge to lower pressures by introducing a magnetic field.

By these methods considerable information has been obtained on sputtering, and theories have been developed to explain the phenomenon. These are described in the first sections of this chapter. Of importance in this work are the velocity and angular distribution of the ejected atoms, and the sputtering yield, S, defined as the mean number of sputtered atoms per incident ion.

Unfortunately, it remains difficult to relate the results from low pressure measurements to the more complex conditions of the glow discharge, and experimental data directly applicable to cold cathode tubes are sparse. For a glow discharge the mass sputtered from the cathode per unit charge, as a function of the pressure and tube potential, is of more interest, and is usually quoted in the literature. The experiments and effects related to sputtering in the glow discharge are dealt with in the later sections of the chapter.

4.2. Sputtering Measurements at Low Pressures

Whatever method of investigating sputtering is employed, the sputtering yield must be determined either by measuring the material lost by the target or by measuring the material gained by an insulated collector. The former method, although giving the quantity sputtered directly, suffers from the disadvantage that surface impurity layers cannot be removed as a subsidiary process, whereas in the second method a screen can be inserted until the target is clean. However, the second method assumes that all the sputtered atoms hitting the collector adhere to it, and that the ratio of the amount depositing on the collector to that depositing elsewhere can be accurately deduced; these assumptions are not fully justified in practice. The amount of material lost or deposited can be determined by weighing, change in resistance, and various other methods. An important method of determining the sputtering yield is the change in light transmission of a transparent collector during deposition of the sputtered material, and this is the one which has been used most often in recent investigations.

In assessing the yield care must be taken to ensure that the measured ion current does not include the secondary electron current due to the secondary emission coefficient, γ, of the surface (cf. Chapter 1). Unfortunately γ cannot be determined in the arc discharge method, and since the measured current equals $i\,(1+\gamma)$, the measured yield, which is quoted, will be $S/(1+\gamma)$ where S is the true yield.

4.2.1. YIELD MEASUREMENTS IN THE GLOW DISCHARGE WITH A MAGNETIC FIELD

This method of observing sputtering was employed by Penning and Moubis[1] and enabled them to extend glow discharge measurements to lower pressures, where back diffusion was reduced and the mean ion energy was more equal to the voltage across the tube than in a conventional glow discharge. The effect of the magnetic field was to increase the length of the electron paths by causing the electrons to move in spirals, thereby increasing the ionisation. By varying the magnetic field, the discharge voltage, current, and pressure could be varied independently of each other. The sputtering yield was determined from the weight increase of small mica plates, which could be moved by a magnet and successively exposed to the sputtered particles without opening the tube.

Table 4.1

Sputtering yield in argon and hydrogen according to Penning and Moubis[1]

Metal	*Gas*	*S for 500V ions*	*Estimated accuracy*
Al	Ar	0·83	±25%
Cu	Ar	1·9	±25%
Ni	Ar	1·35	±50%
Ag	Ar	1·5	±50%
Ni	H_2	0·15	±50%
Ag	H_2	0·56	±50%

The most extensive measurements were for copper in argon over the pressure range 0·003–0·3 torr. After correction for γ, the yield, S, was found to be roughly proportional to the voltage, but to decrease slightly with increasing current. This latter result has not been confirmed by sputtering experiments using other methods, but might be explained by a decreasing probability for oblique angle impacts,[2] since at higher current densities the cathode fall distance decreases and less collisions with gas atoms occur in the cathode region. A summary of their results is given in Table 4.1. Although the accuracy of these measurements was poor, the experiment is interesting as the pressure range covered extends from that used in a glow discharge to pressures typical of arc discharge measurements described in the next section (cf. Fig. 4.10).

4.2.2. YIELD MEASUREMENTS IN AN ARC DISCHARGE

Although a glow discharge cannot be maintained at pressures low enough to overcome back diffusion (except in the presence of a magnetic field) an arc discharge of quite high current can be obtained with a thermionic or mercury-pool cathode. The target can then be immersed in the discharge plasma as a probe, and the ions drawn to it by applying a potential negative to that of the plasma.

Under these circumstances the target is surrounded by a space charge sheath the thickness of which decreases with increase of the ion current. The latter is given by

$$j_+ = \frac{N_+ e c_+}{4} \tag{4.1}$$

where N_+ is the ion density which is determined by the arc current, and c_+ the thermal velocity of the ions. Thus j_+ can be varied independently of the ion energy which is determined by the difference of potential between the target and the plasma.

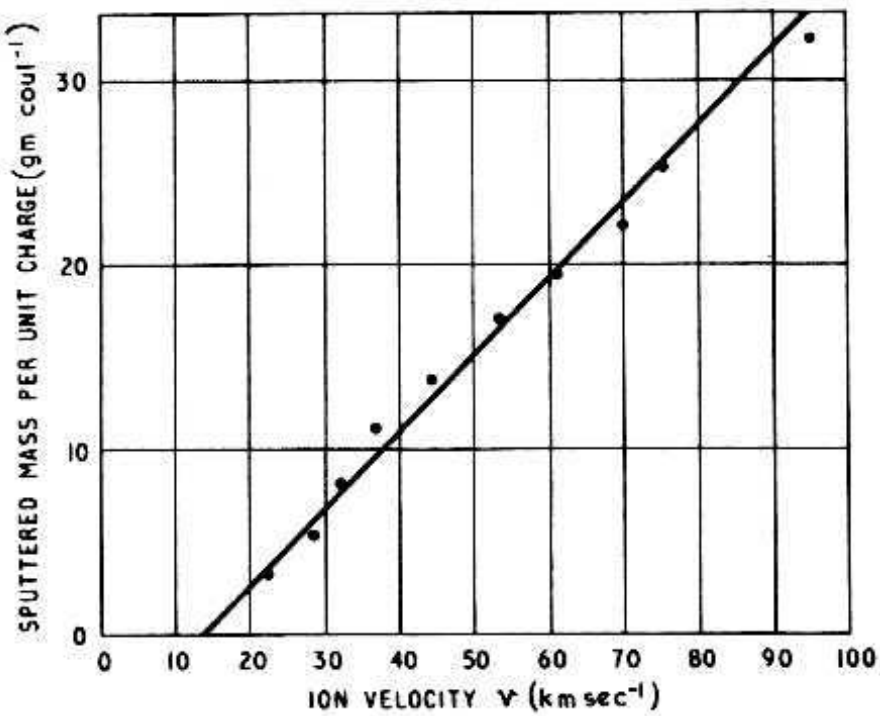


Fig 4.1. Sputtering yield (mass per unit charge) as a function of the ion velocity for tantalum sputtered by mercury ions. Güntherschulze[7]

Among the earliest yield measurements by this method were those of the laboratory staff of G.E.C.[3] and those of Kingdon and Langmuir[4] in 1923. Both these investigations employed wire targets, which meant that the angle of incidence varied, particularly as the ion density was low so that the sheath thickness was high,

of the order of the tube dimensions. Further, since the ion density was low, removal of surface contamination would have been difficult. Much more satisfactory conditions were present in the experiments of Güntherschulze and Meyer[5, 6] where arc currents of several amps were passed at pressures below 10^{-2} torr, and where the target was in the form of a disk. With such currents the sheath thickness was of the order of 5 mm so that with a disk of 5 cm diameter essentially normal bombardment of the target occurred. In the earlier experiments[5] sputtering measurements were made for copper and silver in helium, hydrogen, nitrogen, neon and argon, whilst in the later experiments measurements were made on sixteen metals in mercury vapour.

No sputtering was detected in helium. In other gases the mass sputtered per unit charge, Q, was plotted against ion energy; the effect of γ was not determined. At a later date Güntherschulze[7] found that the results could be related more satisfactorily to the square root of the energy (i.e. the velocity of the ions). In all cases the graphs of Q against the velocity were linear as for example that shown in Fig. 4.1 for tantalum in mercury vapour. Hence, assuming γ to be constant, the yield was related to velocity of the ions, v, according to an expression of the form

$$S = A(v - v_o) \tag{4.2}$$

where A and v_o are constants. The latter represents the threshold velocity of the ions for sputtering. For the eight metals Al, Si, Mn, Fe, Zr, Mo, Ta, and W in mercury vapour, the constants of Equation 4.2 were found to be the same, the equation being given by:

$$S = 0{\cdot}0629(v - 13{\cdot}3) \tag{4.3}$$

v being in km sec^{-1}.

With the advent of modern vacuum techniques giving clean surfaces free from contaminating layers, and more sensitive methods of determining the amount of material sputtered, a large number of results have been published over the last ten years giving reliable data and extending the range to lower ion energies. Of importance has been the work of Wehner and his co-workers who have produced considerable information on sputtering yield, and have contributed to the general understanding of the mechanism. In the earliest paper Wehner and Medicus[8] investigated the sputtering of platinum by xenon ions at low energies 50–200 eV in an attempt to establish whether or not a minimum ion energy was required for sputtering to occur, i.e. a sputtering threshold. The detection method was particularly sensitive as it was based on the change in

work function of a tungsten probe when the target metal was deposited on to it. Although a reliable calibration of the change in work function in terms of the yield could not be obtained, the technique did enable relative measurement at low ion energy. Contrary to Güntherschulze and Meyer they found that the yield was proportional to the square of the ion energy, i.e.

$$\frac{S}{1+\gamma} = C(V - V_o)^2 \tag{4.4}$$

where V_o = the threshold energy $\sim$ 40 eV.

Later Wehner[9, 10, 11, 12] used a pool-type mercury arc discharge as a source of ions with a grid inserted in the arc, a repeller electrode and a magnetic field to increase the ion density. Sputtered material from the target immersed in the plasma between anode and grid

Table 4.2

Threshold values for sputtering in mercury vapour, according to Wehner[10]

Metal	*Threshold energy exp.* (eV)	*Threshold Energy calculated according to Equation 4.5* (eV)
Al	120–140	136
Si	60–70	61
Ti	110–130	123
V	120–130	128
Cr	60– 80	64
Fe	60– 70	75
Co	80–100	96
Ni	70– 90	92
Cu	50– 70	74
Ge	40– 50	57
Zr	120–130	120
Cb	120–130	142
Mo	80–100	72
Rh	70– 80	62
Pd	50– 80	50
Ag	40– 50	43
Sb	30– 40	23
Hf	150–180	163
Ta	120–140	138
W	80–100	87
Re	70– 80	69
Pt	70– 90	67
Au	40– 50	46
Pb	20– 40	25
Th	120–140	133
U	70– 80	78

passed through the slit in a cylindrical metal screen to deposit on to a surrounding cylindrical glass collector. The collector could be rotated to expose several portions of the glass in turn to the discharge, and the yield was determined by the variation of the light transmission through it. The initial measurements were approximate determinations of the threshold energies for sputtering of 26 metals,[9, 10] determined as the maximum energy for no visible deposit. The values obtained are given in Table 4.2. It was found that the product of the momentum transferred to the target atoms at threshold and the sound velocity in the metal was proportional to the heat of sublimation H; in terms of threshold energy expressed in eV:

$$V_o = \left[\frac{1{\cdot}7 \times 10^5 (m+M)H}{m^{\frac{1}{2}} M v_s}\right]^2 \tag{4.5}$$

where m = mass of the gas ion
M = mass of the metal atom
v_s = the velocity of sound in the metal in cm/sec.
H = is in kcal./mol.

Values of threshold energy calculated by Equation 4.5 are also given in Table 4.2. Measurements of the sputtering yield as a function of the ion energy were given for platinum and are shown in Fig. 4.2. The yield was proportional to the ion energy:

$$\frac{S}{1+\gamma} = C'(V - V_o') \tag{4.6}$$

where C' and V_o' are constants. However, at low ion energies the yield was proportional to $(V - V_o)^2$ as for xenon. Wehner also carried out a number of qualitative experiments which showed that the yield was increased and the threshold reduced for ions incident at angles less than 90°, thus confirming earlier results of Fetz[13] for wires of varying diameters. More direct information on the effect of the angle of incidence on sputtering yields was obtained by Wehner[14] in 1959. Ion beams were extracted from a low pressure (2×10^{-4} torr) mercury plasma through small holes in a graphite plate, held negative with respect to the plasma. Small polycrystalline metal spheres, with diameters less than the beam diameter, were mounted opposite the holes in the graphite plate and held at the same potential as the plate. The angle effect of the sputtering yield was studied by measuring the thickness of material removed in different zones using shadow micrographs of the spheres. The ion energy range was 125–800 eV, with a beam current density of 1 mA cm^{-2} at 150 eV. The metals studied could be divided into

three groups: (1) gold, silver, copper and platinum which showed only a slight variation with angle, although giving high yields at normal incidence; (2) nickel and tungsten with a moderate angle effect; and (3) iron, tantalum, and molybdenum with a very pronounced angle effect. In most cases the yield showed a maximum for an angle of incidence of 30° to 40° to the surface. Fig. 4.3 shows the results obtained for tungsten. Similar curves have been obtained using ion beams at higher energies (1–50 keV) in vacuum systems. The increase in yield as the incident angle is increased from the

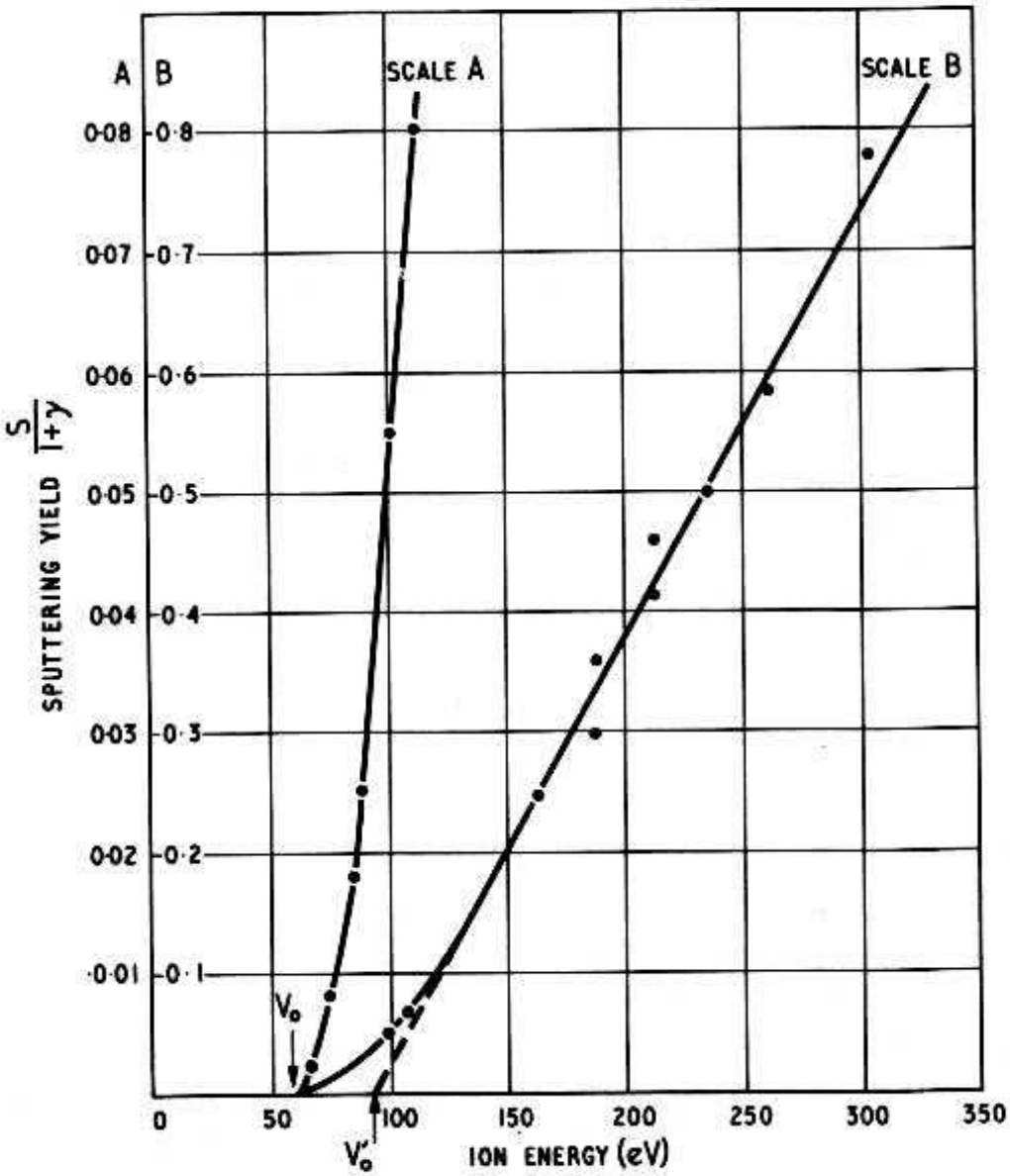


Fig. 4.2. Sputtering yield as a function of ion energy for platinum sputtered by mercury ions. Wehner[10]

normal (90°) is considered to be due to the fact that the penetration depth of the ions is reduced, and therefore the metal atoms excited by the impact can escape more readily. The subsequent decline at low angles of incidence is probably due to the incident ions being reflected before giving up their energy to the surface atoms.

In the further papers of Wehner[11, 12] the yield measurements were extended to 25 further metals using similar apparatus to that used in his earlier work[9, 10]. In each case the yield varied linearly with ion energy, except near threshold. Values taken from Wehner's curves have been inserted in Table 4.3. Wehner[11] has stated that

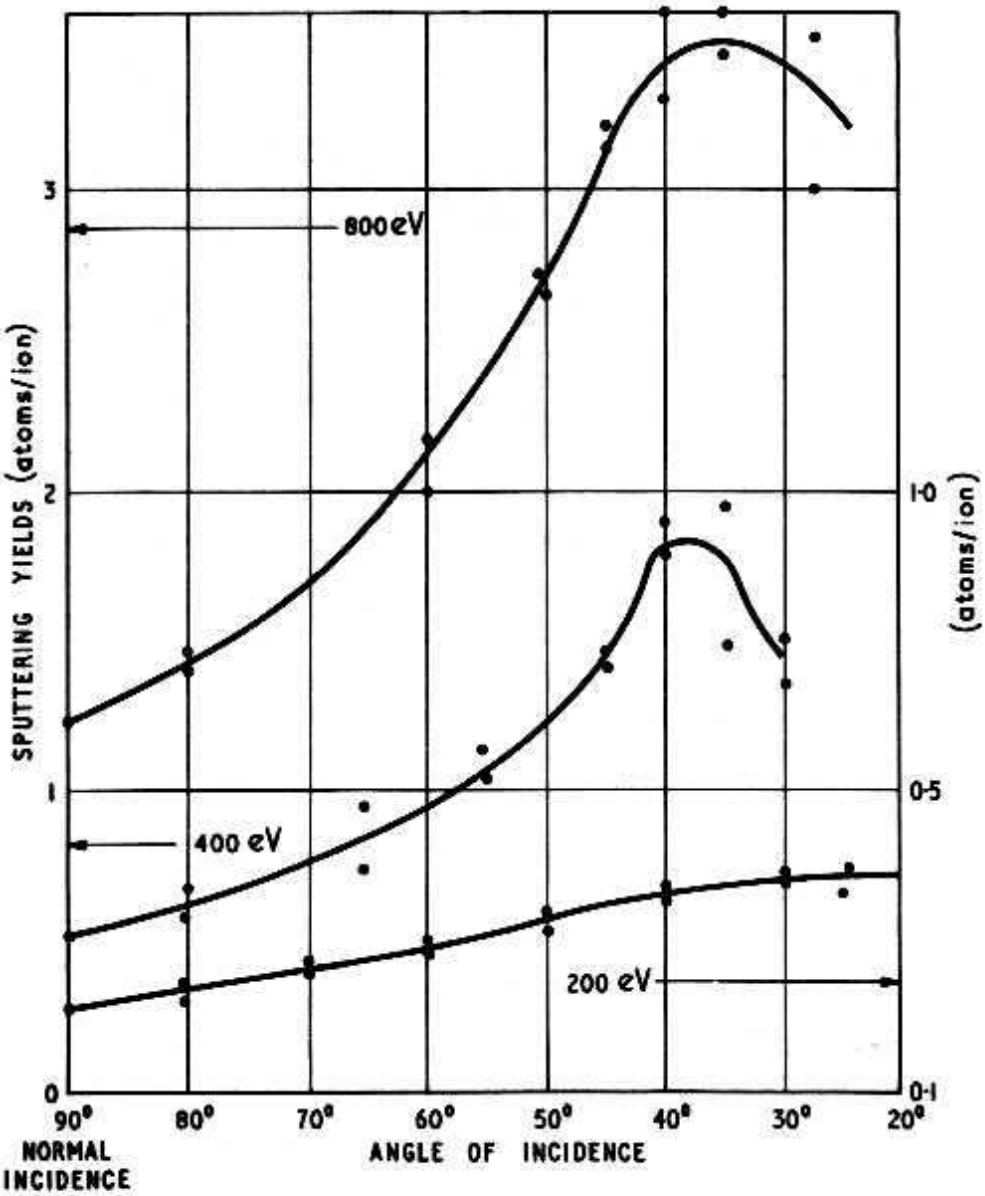


Fig. 4.3. Sputtering yield of tungsten as a function of the angle of incidence of mercury ions, at three different ion energies. Wehner[14]

the earlier approximate determinations of the threshold energies can be roughly identified with the cut-in energies, V_o' of Equation 4.6. However, Morgulis and Tishchenko[15] using radio-active tracer techniques obtained threshold energies of the order of 3–21 eV, much lower than obtained by Wehner. Their results are given in graphical form in Fig. 4.4 as a function of the heat of sublimation. However, as Stuart and Wehner[16] point out, these low thresholds may have been due to doubly-charged ions in the plasma, although it seems likely that there are always some loosely

bound atoms on the surface which can be sputtered by very low energy ions. Thus the more sensitive the detector the lower the threshold will appear to be.

More recently Laegreid and Wehner[17] and Rosenberg and

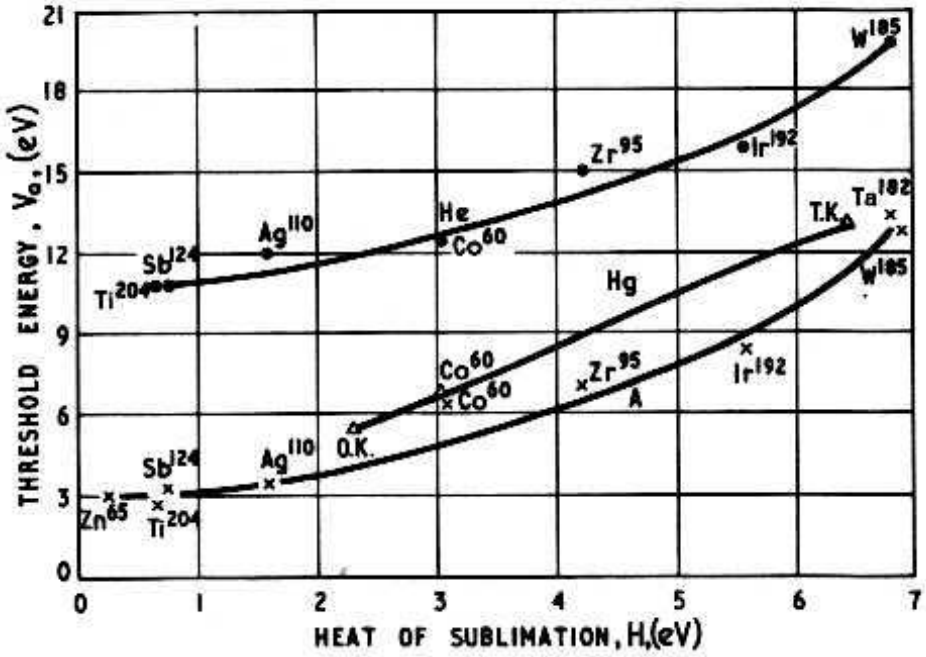


Fig. 4.4. The threshold energy for sputtering by helium, argon and mercury ions as a function of the heat of sublimation of the target metal. Morgulis and Tishchenko.[15] *The points marked* T.K. *and* O.K. *are for thoria on thoriated tungsten and barium on barium oxide cathodes respectively obtained by 'change in work function' methods*

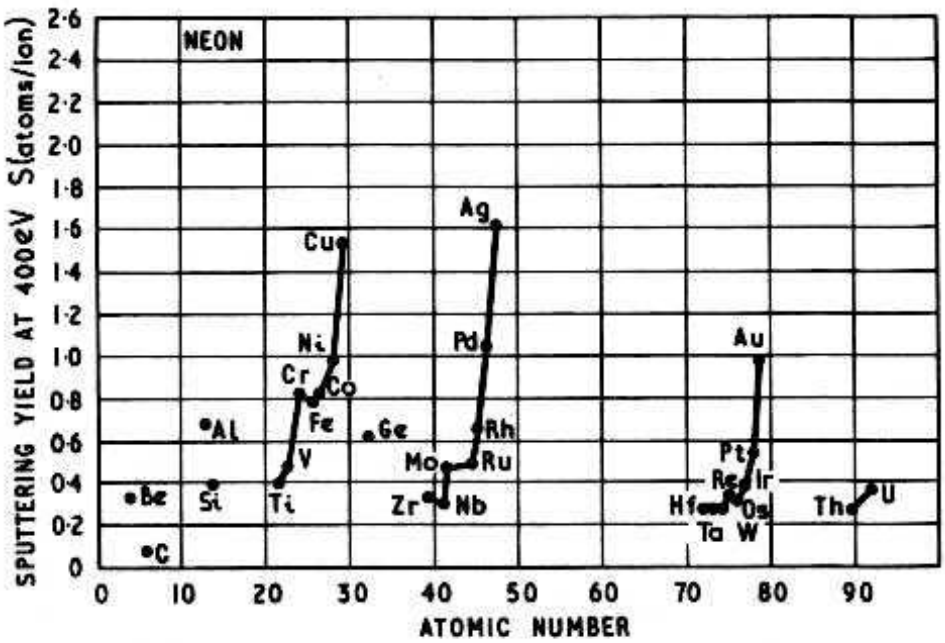


Fig. 4.5. Sputtering yields for 400 eV neon ions as a function of the atomic number of the target element. Laegreid and Wehner[17]

Wehner[18] have extended the yield measurements to 30 metals and semiconductors sputtered by ions of all the inert gases. The values are given in Table 4.3 and represent the most comprehensive and reliable data for sputtering at low ion energies. Unlike the mercury-vapour measurements the yield curves showed only limited linear portions, the slope decreasing steadily with increasing ion energy. Wehner showed that if the yield at the same ion energy for a given ion is plotted against the atomic number of the target element, then the yield increases in any one period of the periodic table with increased filling of the d-shells (see Section 4.3). This is shown in Fig. 4.5 for neon ions at 400 eV. The highest yields were for copper, silver and gold.

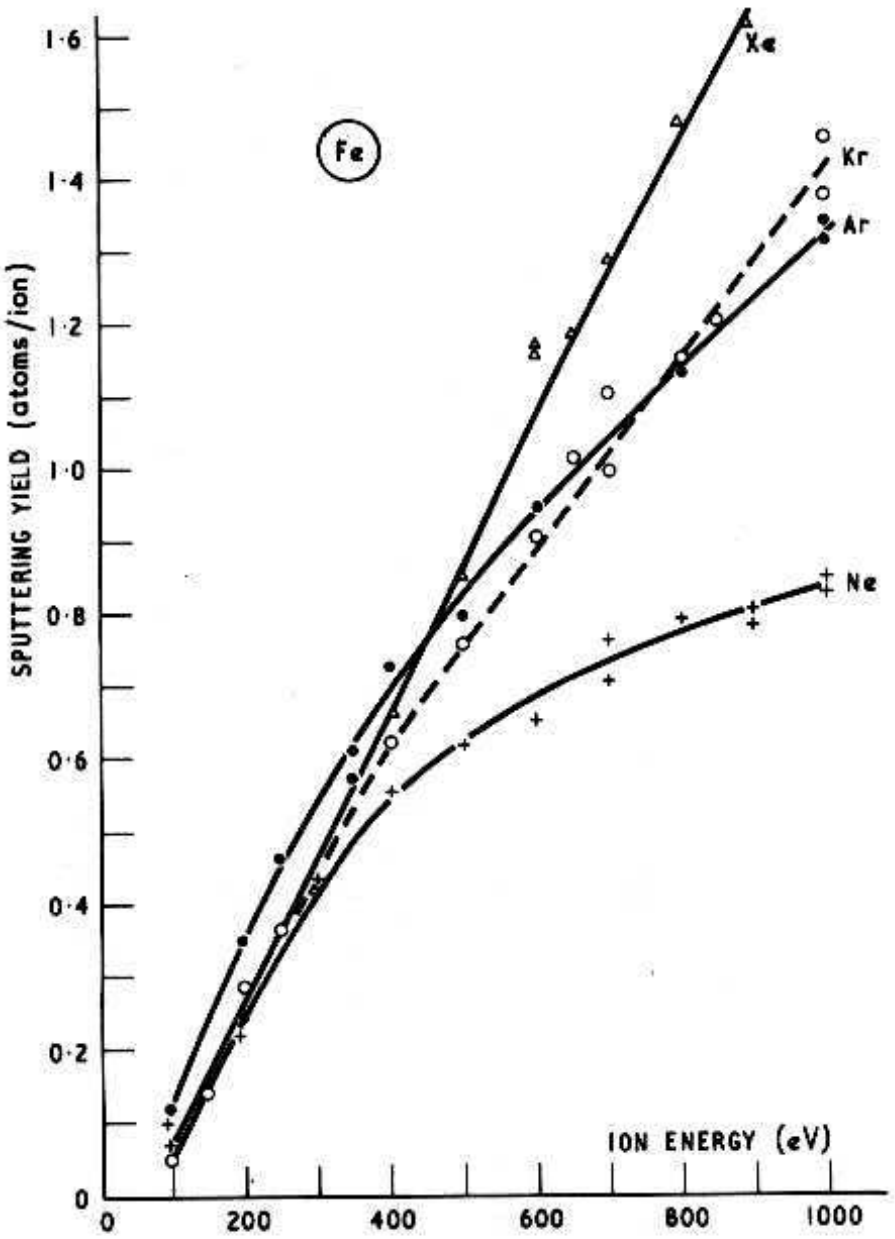
Other yield measurements in arc discharges have been made by Strachan and Harris,[19] Wolsky *et al*[20, 21] and Koedam *et al.*[22, 23] Strachan and Harris used targets of carbon and iron sputtered by mercury and argon ions, determining the yield by the direct weight loss method. The yield, uncorrected for secondary emission, varied linearly with ion energy up to a few kilovolts where the variation changed to proportionality with the square root of the energy, i.e. confirming Equation 4.2. They determined the secondary emission by observation of the space-charge sheath and then found that, above a few hundred volts energy, S was given by

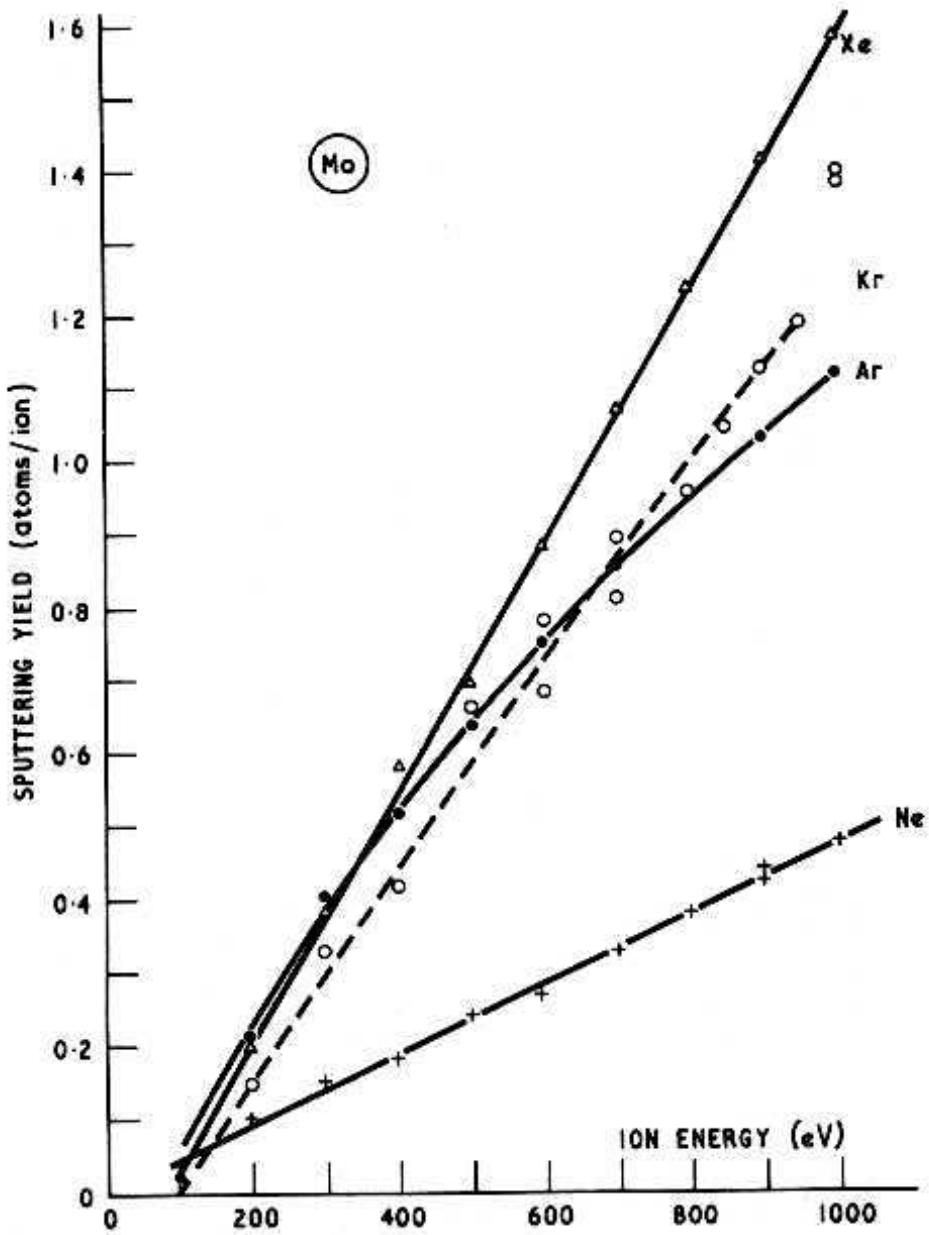
$$S = \mathrm{C} \log_{\mathrm{e}} \frac{V}{V_2} \tag{4.7}$$

where C and V_2 are constants.

Koedam carried out yield measurements on polycrystalline silver and single crystals of copper and nickel in neon, argon and krypton. The targets were immersed in a plasma created by an electron beam which was prevented from spreading by an axial magnetic field. The yields were determined from observations of the optical transmission of the deposits on a glass plate. The yield curves were linear over most of the energy range, but in neon and argon the values for silver were lower than those of Laegreid and Wehner, e.g. at 200 eV the values were 0·7 atom/ion for neon and 1·13 atoms/ion for argon compared with Laegreid and Wehner's values of 1·00 and 1·58 respectively.

Weijsenfeld, Hoogendoorn and Koedam[23] have extended the measurements to copper, nickel, iron and molybdenum in neon, argon, krypton and xenon. The values for iron and molybdenum were rather lower than Laegreid and Wehner, but for copper and nickel good agreement was obtained. Their curves are reproduced in Fig. 4.6.


Fe
Xe
Kr
Ar
Ne
SPUTTERING YIELD (atoms/ion)
ION ENERGY (eV)
0
0·2
0·4
0·6
0·8
1·0
1·2
1·4
1·6
200
400
600
800
1000


Mo
Xe
Kr
Ar
Ne
SPUTTERING YIELD (atoms/ion)
ION ENERGY (eV)
0
0·2
0·4
0·6
0·8
1·0
1·2
1·4
1·6
200
400
600
800
1000

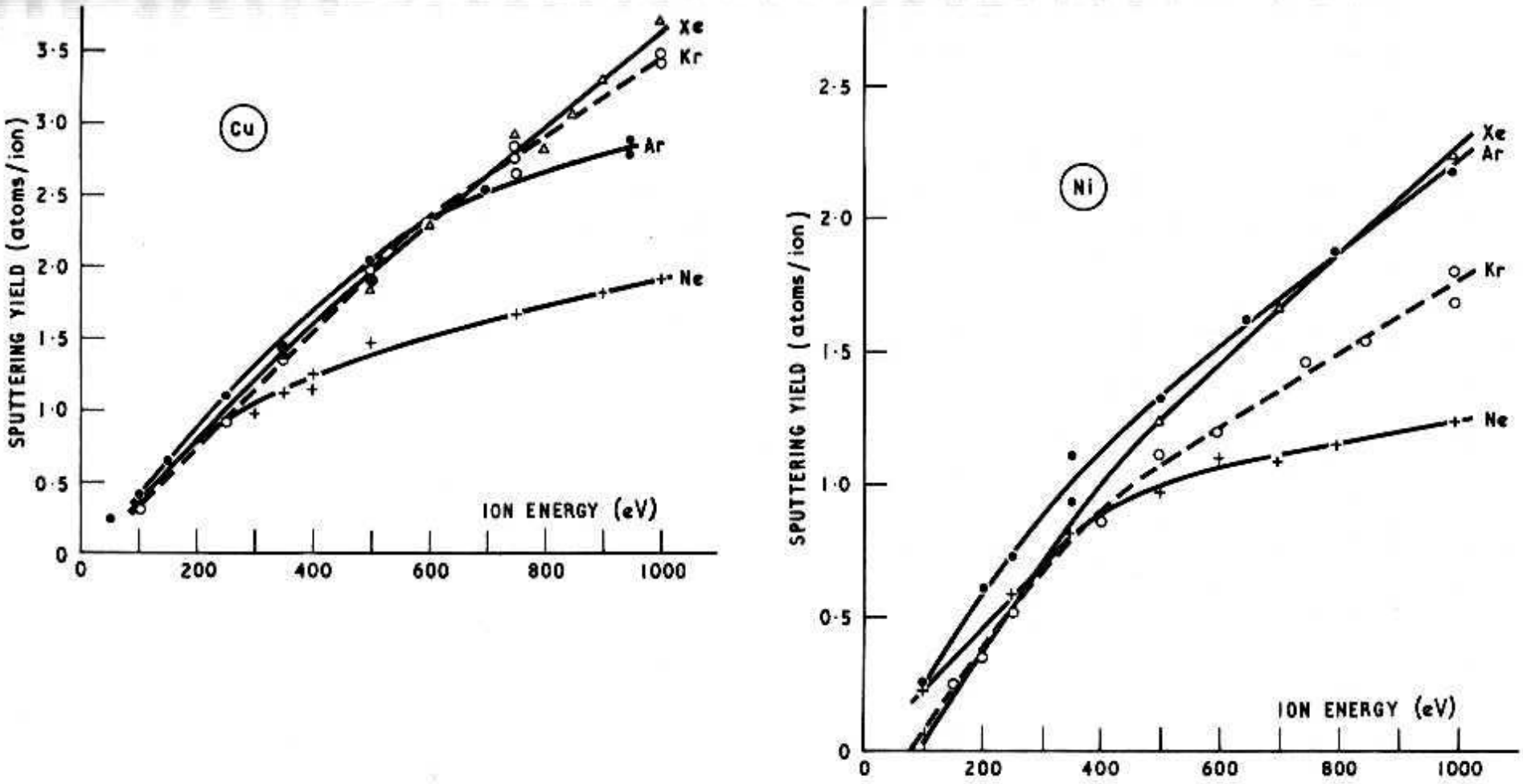


Figure 4.6. Sputtering yields as a function of ion energy for copper, nickel, iron and molybdenum sputtered by neon, argon, krypton, and xenon ions. Weijsenfeld, Hoogendoorn and Koedam[23]

4.2.3. YIELD MEASUREMENTS USING ION BEAMS

One of the first attempts to use an ion beam for sputtering measurements was that of Timoshenko[24] using an argon ion beam from an arc discharge incident on a silver target. The curve obtained for yield against ion energy up to 6 keV, shown in Fig. 4.7, was unusual and has not been repeated by later workers.

More extensive measurements were carried out by Keywell[25, 26] some ten years later using an ion beam from a Philips ion-gauge discharge (P.I.G. discharge). The ion current density was of the order of 50 $\mu A\,cm^{-2}$ at a pressure of $1{\cdot}2 \times 10^{-3}$ torr. The first experiments were with silver using an argon ion beam, but the experiments were later extended to silver with krypton, neon and helium ions, copper with krypton and argon ions, and lead with argon and helium ions. For argon sputtering of silver the yield was proportional to the logarithm of the ion energy, i.e. agreeing with Strachan and Harris, Equation 4.7. The equation to the curve

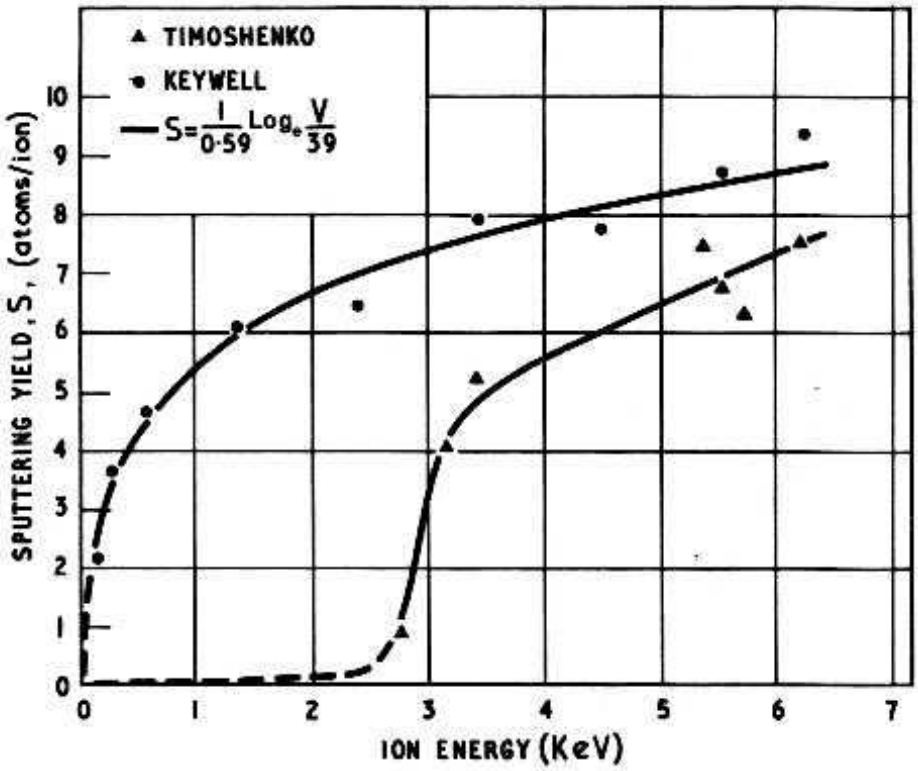


Fig. 4.7. Sputtering yields as a function of ion energy for silver sputtered by argon ions, showing results of Timoshenko[24] and of Keywell[25] to which the neutron cooling theory is applied

is shown on Fig. 4.7. The spreads, however, were large, which as pointed out by Wehner[2] is likely to occur with a P.I.G. discharge where ions with multiple charges result. At about the same time Bradley[27] reported measurements of the sputtering rate for sodium and potassium using helium, neon, argon and xenon ions. He used

a Finkelstein-type ion source giving singly-charged ions of 0–1·8 keV, and kept his background pressure down to 5×10^{-8} torr. He measured the sputtering yield by ionising the sputtered atoms at the surface of a hot platinum electrode and collecting the resulting ions at a negative collector. This is a particularly sensitive method of measuring sputtering yield, but unfortunately can only be applied to the alkali metals, for which the ionisation potential is lower than the work function of the platinum. He found that for argon and xenon the yields were nearly linear with the momentum of the ion (square root of the energy), but that for helium and neon ions the yields varied linearly with the logarithm of the energy.

A novel method of measuring the yield from ion beam sputtering has been applied by Wehner and Rosenberg[28] for mercury ions and later by Ken Knight and Wehner[29] for hydrogen ions. The sputtering yields were calculated from the time required for the beam to penetrate the target in the form of thin foil. In the case of mercury ions the yield increased steadily with increasing energy up to 15 keV, and for the various target materials the pattern within the periodic tables was similar, for both gas ions, to that obtained at lower energies (Fig. 4.5). Even more interesting is a method developed by McKeown,[30] where the target material was plated on to a quartz crystal connected to an oscillator, and the sputtering observed as a change in frequency of the crystal. The method gives the yield absolutely and continuously, and is particularly sensitive. McKeown measured sputtering of gold by argon ions of energies below 100 eV. His results agreed with those of Wehner, but the threshold obtained was much lower, being $\sim$ 10 eV.

A number of workers have made yield measurements at energies up to 200 keV normally using electromagnetic isotope separators as ion sources, in particular Yonts, Normand and Harrison;[31] Grønlund and Moore,[32] Rol, Fluit and Kistemaker,[33] Almen and Bruce,[34] Perovic and Cobic[35] and Ramer *et al.*[36] Although such energies would not occur in normal glow discharges, it is interesting that at such energies it is found that the yield passes through a maximum or tends to a constant value as the ion energy increases.

Before leaving ion beam measurements, mention must be made of yield values obtained with insulating materials as targets. Hines and Wallor[37] measured the yield from vitreous silica bombarded by 20–60 keV xenon ions by observing interference fringes between the sputtered surface and an optical flat. He found the yield increased from 0·71 at 20 keV to a maximum of 1·78 at 40 keV reducing at 60 keV to 1·31. Akishin, Vasilev, and Isaev[38] have measured the sputtering yield of mica, fused quartz, and glass

bombarded by krypton ions of 4–30 keV energies. They found the yield rose to a steady value at ion energies above 10 keV and were of the same order for the three materials, between 1·5 and 2.

The only measurements of the sputtering of insulating materials at lower energies were made by Jorgenson and Wehner[39] on quartz and Pyrex glass. The method was based on recording the changes in a Langmuir probe characteristic when a thin film of insulating material was deposited on the probe by r.f. sputtering and was resputtered by d.c. The *i-V* characteristic of such a probe is extremely sensitive to minute deposits of foreign materials, particularly insulating material, and deposits of less than a monolayer can be detected. The deposit rate from the r.f. discharge was determined from longer sputtering runs in which a measurable step height was built up on flat substrates mounted in the same position as the probe. The rate of removal of the deposited layers of known thickness was then used for collecting yield data. Analysis of the deposit sputtered from a quartz target in a 10^{-2} torr argon discharge shows it to be $SiO_{1.6}$. The sputtering yield was $\sim 0{\cdot}1$ molecules/ion for 100 eV ions with a sputtering threshold of ~ 16 eV. The yield and sputtering threshold for Pyrex 7740 were found to be similar.

4.2.4. SPUTTERING OF SINGLE CRYSTALS

In view of the etch patterns that have been observed on the sputtering of polycrystalline targets, and the momentum transfer results, it was expected that the yield from single crystals would depend on the crystal faces. The first study of the sputtering of single crystals, however, was not carried out until the 1950's when Wehner[10] examined the sputtering from a tungsten single crystal ball of 1 mm diameter. He found that the deposit on a glass collector placed one centimetre away was not uniform but exhibited a pattern resembling electron field-emission microscope patterns. In better adapted experiments with single-crystal rods of copper and silver, he discovered that the deposited patches corresponded to the ⟨110⟩ axes of the crystal, which for a f.c.c. crystal is the direction of closest packing. This was confirmed by single crystals with plane surfaces oriented with the crystal lattice. For the (111) surface three-fold symmetrical patterns were obtained, whereas for the (100) plane four-fold symmetry resulted and for the (110) plane there was a central deposit surrounded by four spots in two-fold symmetry. For b.c.c. crystals such as tungsten and molybdenum

the directions of preferred ejection were also along the directions of closest packing, in this case ⟨111⟩.

The patterns, however, were dependent on the ion energy, in general more patches being observed as the ion energy increased due to atoms also being ejected in directions of second nearest neighbours as well as the directions of closest packing. Similar results were obtained by Koedam[40] and are shown in Plate 4.1a

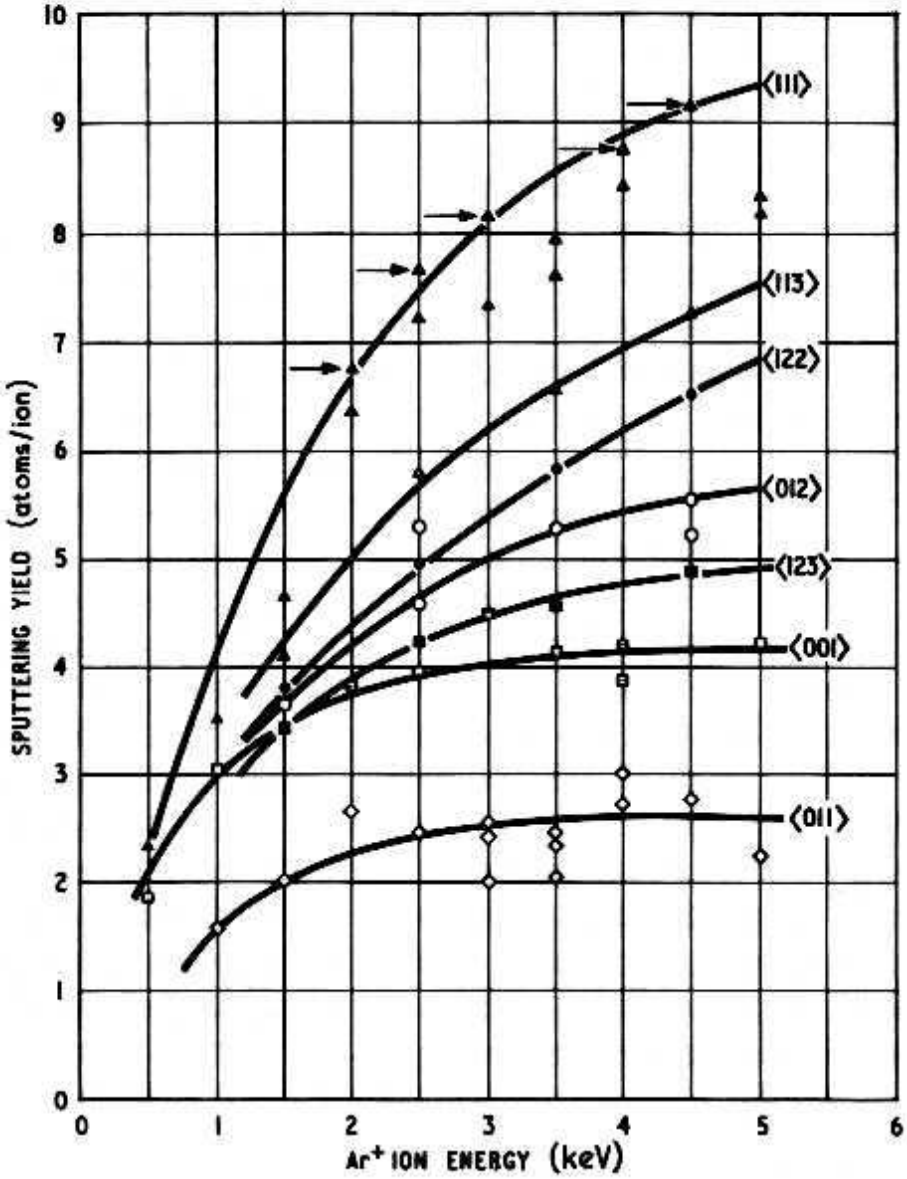


Fig. 4.8. Sputtering yields as a function of ion energy for copper single crystals sputtered by normally incident argon ions for the various crystal faces. The arrows indicate points obtained on freshly electropolished (111) surfaces. Southern, Willis and Robinson[44]

where the sputter patterns are given for a copper crystal with a (100) surface bombarded by krypton ions of various energies. At low energies sputtering only occurs along the directions of closest packing, as the energy increases sputtering occurs normal to the

surface (the ⟨001⟩ direction) and at higher energies still the sputtering in this direction predominates. Koedam also compared the ejection patterns with various inert gas ions of the same energy (500 eV) and showed that as the ion mass increased the sputtering in the ⟨001⟩ direction decreased; this is demonstrated in Plate 4.1b. A more extensive study of sputtering of single crystals has been carried out by Anderson and Wehner[41] for mercury ions incident on copper, nickel, gold, iron and germanium crystals. Contrary to Koedam, who found no preferential ejection of sputtered atoms from germanium unless the temperature was raised to 500°C when a vague pattern occurred, Anderson and Wehner obtained fairly pronounced patterns. However, the patterns resembled those from a b.c.c. crystal rather than those expected from the diamond structure of germanium, due presumably to the formation of interstitial mercury or germanium ions near the surface. This work has been extended, by Anderson to neon and argon ions,[42] and to several b.c.c. crystals with neon, argon, and xenon ions.[43]

Deposit patterns can still be observed at much higher energies. For example, Rol *et al*[33] observed the pattern produced by 20 keV argon ions incident on the (110) and (100) surfaces of copper and obtained similar results to those at lower energies.

The yield from the various crystal plane surfaces has been measured by Southern, Willis and Robinson[44] for copper monocrystals bombarded by argon ions of 1–5 keV. The yields were deduced from the weight loss of the target and are shown in Fig. 4.8. Similar results were obtained by Magnuson and Carlston[45] for (100) (110) and (111) planes of copper with argon ion energies from 1–10 keV. The fact that the yields varied considerably from the different crystal faces, particularly at the higher ion energies, may well explain the discrepancies found in the literature for polycrystalline copper, since the different samples used were likely to have had different crystal structures.

4.3. Theories of Sputtering

Sputtering theories fall into three classes: theories in which the momentum of the bombarding ion is considered to transfer to the atoms of the target in a small number of separate two-body collisions between the ion and an atom of the metal and between two further metal atoms, in up to three collisions; theories in which a large number of collisions are considered; and the theory of evaporation of atoms from localised regions raised to a high temperature by impact of the ions.

The only evidence in support of the evaporation theory is the

cosine angular distribution of sputtered material from polycrystalline targets at high ion energies, found for example by Grønlund and Moore.[32] It cannot explain the angular distribution at low ion energies or the deposit patterns from single crystals, which are explained by collision theories. It has been used, however, by Townes[46] who has been the only worker to attempt a theory of sputtering in the glow discharge, and will be given brief mention in Section 4.4.1.

4.3.1. COLLISION THEORIES INVOLVING FEW COLLISIONS

The simplest collision theory has been given by Stark[47] who considered the ejection of an atom by transfer of energy in a single elastic collision. He assumed that for sputtering to occur the energy transferred by a normally incident ion to a target atom must be at least equal to the binding energy of the atom. Hence the threshold energy for normal incidence is

$$V_o = \frac{(m+M)^2}{4mM} . H \tag{4.8}$$

where m = the mass of the ion, M = the mass of the target atom, and H = the heat of sublimation.

Although this expression is similar to the empirical formula of Wehner (Equation 4.5) it does not agree with experimental threshold values. A more extensive theory was developed by Henschke[48] in 1957 in which more than one collision was considered between the ion and the target atoms. He considered the ions and atoms as hard spheres undergoing elastic collisions, with energy losses due to the excitation of Debye waves in the lattice during the collisions. For oblique incidence, he considered a single collision mechanism, whereas for normal incidence he considered a double and a triple collision mechanism. The triple mechanism can only occur after some of the lower surface atoms have been removed causing furrows in the lattice. The single and double collision mechanisms are shown in Fig. 4.9 (a) and (b). For oblique incidence, mechanism (a), he deduced a threshold energy

$$V_o = \frac{(m+M)^2}{mM} H_{hkl} \frac{1}{(1+\delta)^2 \cos^2\Phi \cos^2\theta} \tag{4.9}$$

where H_{hkl} = the heat of sublimation appropriate to the surface crystal plane (hkl)

Φ = angle of the line joining the centres of ion and atom at collision with the surface normal. (This will also be the angle of ejection)

θ = the angle of incidence also with reference to the line joining the centres of ion and atom at impact

δ = the Debye wave dissipation coefficient

For normal incidence where two impacts occur, mechanism (b),

$$V_o = \frac{(m+M)^2}{mM} H_{hkl} \frac{1}{(1+\delta)^2 \cos^2\Phi \cos^2\theta_2 [1-(1-\delta^2)\cos^2\theta_1]} \quad (4.10)$$

where θ_1 and θ_2 are the angles of incidence for the first and second collisions.

The quantities H_{hkl} have not been measured and Henschke had to estimate them from average values of H for polycrystalline materials. The values of δ were found to be about 0·5–0·6 for a b.c.c. metal and 0·4–0·5 for a f.c.c. metal when the theoretical values

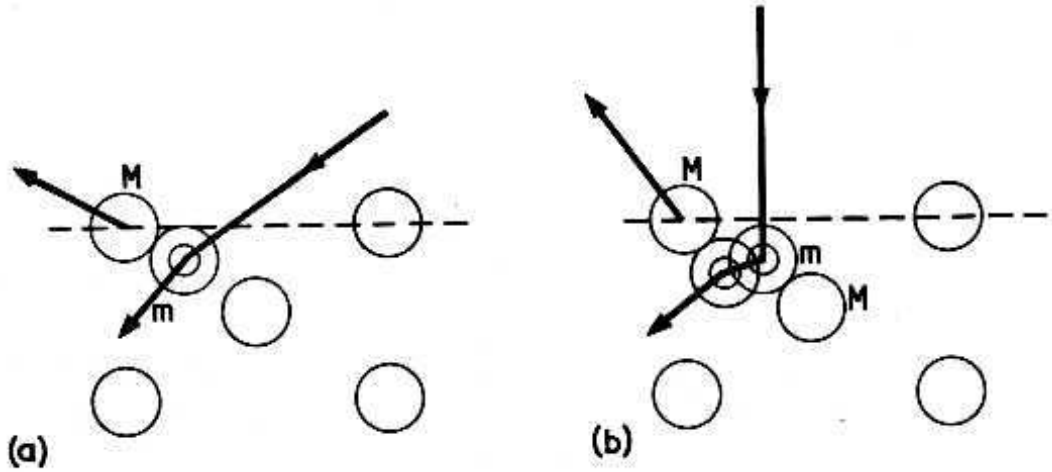


Fig. 4.9. The single (a) and double (b) collision mechanisms considered by Henschke[48]

of V_o from Equation 4.10 were compared with Wehner's values, Table 4.2. Henschke[49] was able to explain the deposit patterns from a single crystal by means of the theory, allowing for the attenuating influence of the different electron densities in the ion paths within the lattice. He also obtained a sputtering yield formula for mechanism (b) as follows[48]

$$S - S_m = \frac{C(V-V_o)^2}{2V_o} \text{ for } V \geqslant V_o \quad (4.11)$$

where S_m is the yield at the threshold energy V_o defined as the lowest energy at which sputtering is detected and C is a constant for any one crystal plane but could vary with energy when averaged over all planes.

Many other simple collision mechanisms are possible. In

particular it is likely that the momentum an atom has received from the incident ion is transferred to a neighbouring atom which is ejected rather than the ion hitting two atoms. Such a process has been considered by Langberg.[50] The second collision between two surface atoms, was considered not as a collision between hard elastic spheres, but in terms of the potential energy between the atoms as a result of a Morse potential interaction. He deduced the threshold energy for a lattice atom tied by N bonds as

$$V_o = \frac{10(m+M)^2}{4mM} H\left(\frac{1{\cdot}1N}{n}+0{\cdot}2\right)\left[1-\frac{1}{ad}\log_e(1+0{\cdot}83N^{\frac{1}{2}})\right]^{-2} \quad (4.12)$$

where n = the number of nearest neighbours
a = a lattice parameter related to compressibility
d = the equilibrium separation of the atoms

For a polycrystalline material N varies between a minimum, N_{min} and maximum N_{max} depending on the position of the atom relative to its neighbouring surface atoms. Similarly there will be a minimum and maximum threshold value $V_{o\,min}$ and $V_{o\,max}$ corresponding to N_{min} and N_{max}.

Langberg also deduced yield values for particles tied by N bonds and showed that the total yield for all values of N was given by

$$S = 0 \text{ for } V \leqslant V_{o\,min}$$

$$S = \tfrac{1}{2}\mathrm{b}(V-V_{o\,min})^2 \text{ for } V_{o\,min} \leqslant V \leqslant V_{o\,max}$$

$$S = \mathrm{b}(V_{o\,max}-V_{o\,min})\left[V-\tfrac{1}{2}(V_{o\,max}+V_{o\,min})\right] \text{ for } V \geqslant V_{o\,max} \quad (4.13)$$

where b is a constant. Equations 4.13 are of similar form to Equations 4.4 and 4.6 where $V_{o\,min}$ is the threshold energy and $\frac{1}{2}(V_{o\,max}+V_{o\,min})$ is the 'cut in' energy. The calculations were in good agreement with Wehner's results when suitable values of b were chosen, and it would appear that Langberg's theory is the most satisfying of those involving only a few collisions. However it is possible that Henschke's collision mechanism (*b*) might be more appropriate for the lowest energies.

At high ion energies, an explanation of the dependence of sputtering yield on the crystal plane for monocrystals has been sought in terms of the penetration of the ions into the crystal lattice. The higher the transparency, the deeper the penetration and thus the lower the sputtering yield. Such a mechanism was first proposed by Rol *et al*[33] and was extended by Southern *et al*[44] and Magnuson and Carlston[45] who obtained good agreement with experimental data.

4.3.2. COLLISION THEORIES INVOLVING MANY COLLISIONS

The single- or two-collision theory does not explain satisfactorily the deposit patterns from single crystals. Wehner[10] suggested that a shock wave is initiated from the point of impact along the close-packed directions in the lattice until, as a result of reflections, a small part of the energy reaches a surface atom. If this energy is sufficient to overcome the binding energy of the surface atom the latter may be sputtered. This transfer of energy along chains of close-packed atoms is similar to the well-known experiment of allowing a billiard ball to move along the axis of a row of billiard balls in contact and to hit the first ball in the row. Momentum is transferred from ball to ball down the row but only the last ball shows appreciable displacement. This description by Wehner was strengthened by the theoretical study of Silsbee[51] who showed from considerations of the potential energy of atoms in the crystal, that if an incoming particle was incident at an angle β to the axis of a chain of atoms, then the angles at which successive atoms moved would decrease until the motion was entirely along the chain; i.e. the momentum would be focused along the chain, provided

$$\cos\beta > \frac{d}{2r}$$

where d is the separation of the centres of the atoms and r is their diameter. This focusing of the momentum transfer along the line of centres of the atoms behaves like a particle moving through the crystal with a certain loss mechanism and has been termed a 'focuson'.[42, 52]

A rather different approach, which also entails consideration of multiple collisions, has been taken by Harrison.[53, 54, 55] The theory is based on that of neutron diffusion, in which all particles not in thermal equilibrium within the lattice were described by a gas-like model in which the atoms moved with constant mean free path. Collisions were assumed to occur only with 'fixed' lattice atoms, but binding energies were neglected. The expression for yield, which was difficult to apply in practice and requires a long proof is

$$S = \mathrm{R_o} \sum_{n=1}^{n=4} \mathrm{R_n} \left[\left(\frac{V}{V_o}\right)^{\mathrm{t_n}} - 1\right] \tag{4.14}$$

where $\mathrm{R_o}$, $\mathrm{R_n}$, and $\mathrm{t_n}$ are constants which are in principle calculable.

The theory has been fitted to Keywell's results with reasonable success[53, 54] (See Fig. 4.7). Recently Harrison[56] has adapted his theory to include the Silsbee chain mechanism.

In summing up, it seems evident that a rigorous theory of sputtering will probably require the Silsbee chain mechanism. However, the Langberg or Henschke theories appear adequate at low energies, and give fair agreement with experimental results.

4.4. Sputtering in a Glow Discharge

In the preceding sections the phenomenon of sputtering at a surface has been discussed and experimental data given for the condition where the mean free path is large compared with the dimensions of the apparatus, so that the sputtered particles travel unhindered to the collecting surface. As the pressure is increased collisions between gas atoms, ions, and sputtered atoms become

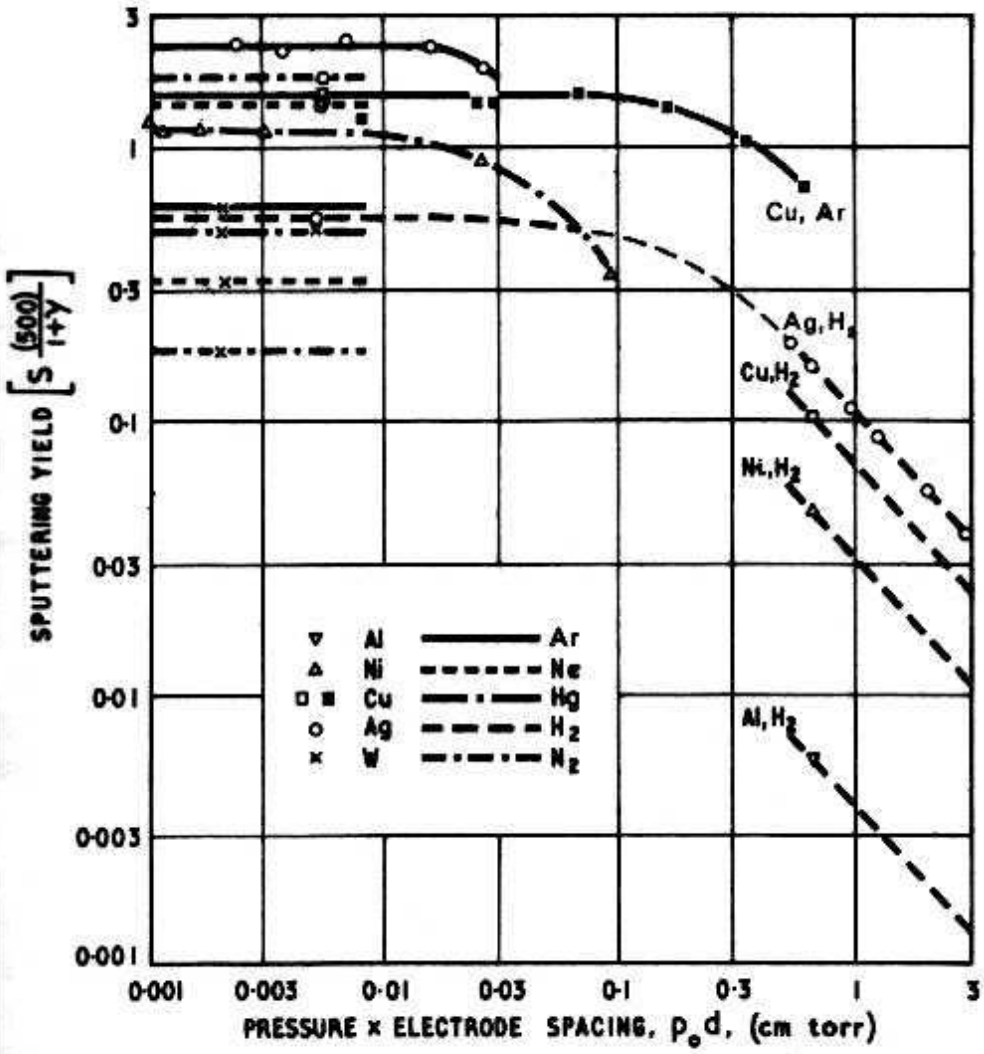


Fig. 4.10. Sputtering yield at 500 eV energy as a function of the product of pressure and electrode separation p_0d—compiled by Penning and Moubis[1]

more frequent, and some of the latter will be scattered back on to the cathode. As a result the apparent sputtering yield decreases. This is demonstrated in Fig. 4.10 which was compiled by Penning and Moubis[1] and combines early arc and glow discharge measurements. The decrease in yield commences at 0·01–0·1 torr cms, which is at a higher pressure × distance than expected if the effect were due to the collisons with gas atoms alone. This is probably accounted for by the increasing probability of bombardment at oblique angles of incidence as the pressure increases.

At the higher pressures (above 10 torr) employed in glow discharge tubes, the mean free path is very short compared with the electrode separation and the number of particles diffusing back to the cathode far exceeds the number escaping, so that the sputtering yield is more than an order lower than obtained from the low pressure experiments. Under these circumstances sputtering yields pertinent to low pressures cannot be directly applied, and measurements at the higher pressures must be made to obtain the necessary data.

4.4.1. MEASUREMENTS OF SPUTTERING YIELD IN A GLOW DISCHARGE

Early measurements on sputtering yields were carried out in discharges at pressures around 1 torr. Most workers found that the mass, Q, sputtered per unit charge passing through the discharge varied with the cathode fall V_C according to a relation:

$$Q = \mathrm{A}(V_C - V_o) \tag{4.15}$$

where A and V_o are constants for any particular gas-metal combination.

The most detailed measurements in this pressure range have been made by Güntherschulze[57] using a silver cathode with hydrogen gas in a discharge tube with plane-parallel electrodes. By varying the pressure, the inter-electrode spacing, and the discharge current he obtained a relation of the form

$$Q = \frac{\mathrm{B}V_C}{pd} \tag{4.16}$$

where B is a constant, as well as Equation 4.15.

Equation 4.15 but not 4.16 was also verified for a number of other cathode materials in hydrogen and oxygen. Güntherschulze also investigated the sputtering of a wire cathode of various

diameters, and found the sputtering rate decreased with increasing diameter, which he attributed to an increase in the amount of silver intercepted by the wire.

Güntherschulze and Tollmien[58] showed that for certain target materials the sputtered metal which returned to the cathode was deposited as microscopic cones with apex angle characteristic of the metal, this was particularly marked for Mg, Cd, Zn, Al, Sn,

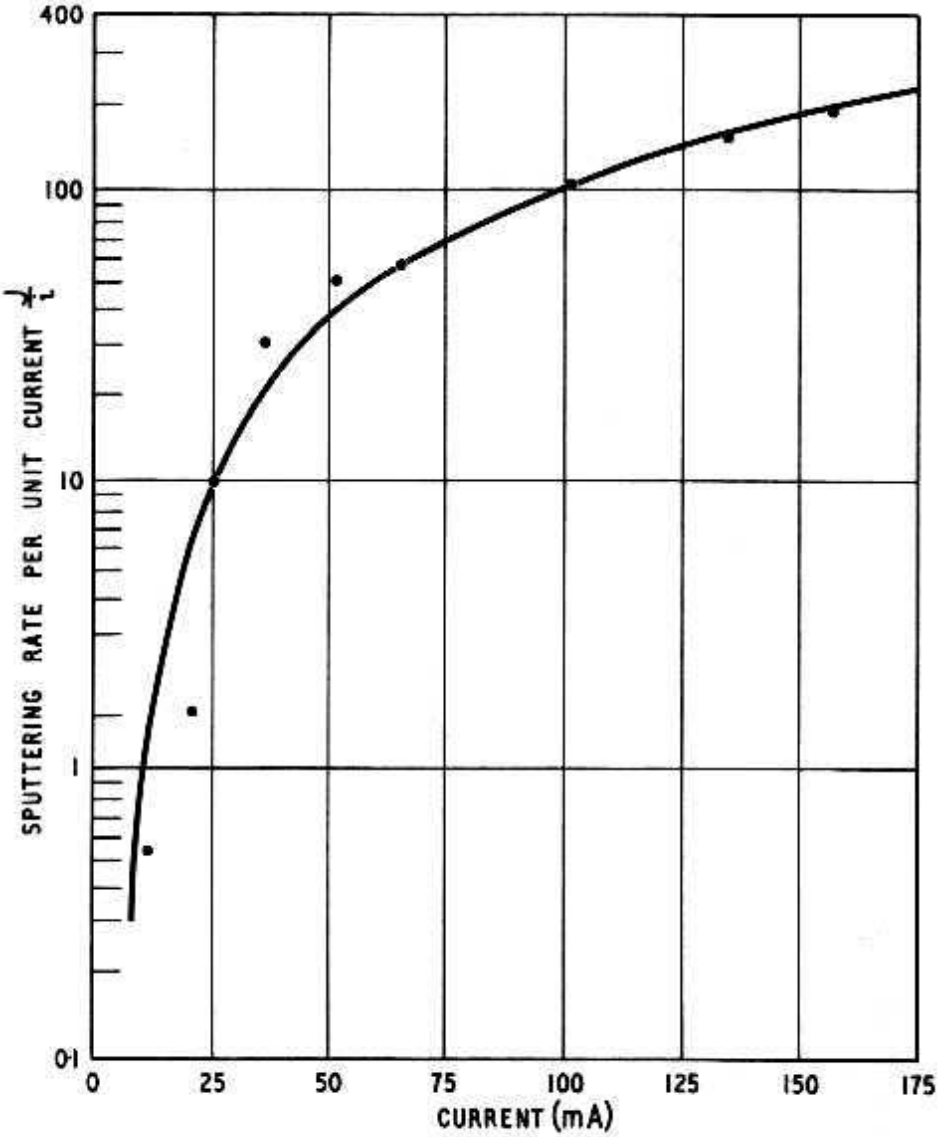


Fig. 4.11. Rate of sputtering per unit current as a function of the current for 99% neon – 1% argon. Experimental points according to Rockwood[59] curve obtained by Townes[46] from Equation 4.17. J measured as 1,000/tube life in hours

Pb, Sb and Bi. The presence of the cones increased the probability of sputtered atoms being trapped and hence the apparent rate of sputtering decreased with time.

At higher pressures, above 10 torr, the only measurements that appear to have been made before 1960 are those of Rockwood[59] who calculated sputtering rates from the life times of commercial cold cathode tubes having cathodes of barium and strontium and

their oxides on nickel substrates. In an attempt to formulate a theory for sputtering in a glow discharge at such pressures, Townes[46] combined the evaporation theory of sputtering, a theory for ion energies in a glow discharge and formula for the diffusion of the sputtered atoms. He obtained a simplified equation of the form

$$\frac{J}{i} = a(7 + b/i^{2/5}) \exp(-b/i^{2/5}) \tag{4.17}$$

where J is the number of atoms deposited per second on the collector, i is the current, and a and b are adjustable constants. He applied the formula to Rockwood's results and the curve is shown in Fig. 4.11 where the theoretical line is drawn through

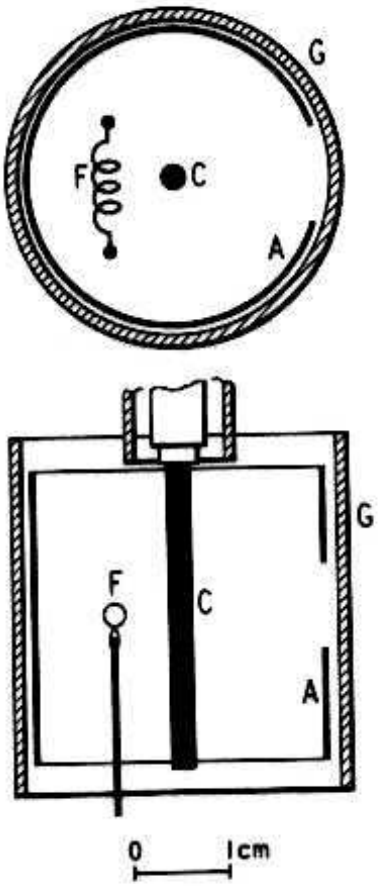


Fig. 4.12. Schematic diagram of experimental tube used by Stocker[60] showing the position of anode A, *cathode* C, *Filament* F, *and glass collector* G

experimental points. However, the line does not curve sufficiently to fit the points adequately, and too many approximations had to be made in order to obtain a reasonable expression.

Recently Stocker[60] reported measurements of the rate of deposition of sputtered molybdenum films in inert gases over the pressure range of 3–21 torr. The deposition of material on a glass collector was observed by measuring the optical transmission.

A preliminary experiment established that the optical density was linearly related to the mass deposited, and thus the variation of the optical density gave a relative measurement of the sputtering rate. The experimental tube is shown diagrammatically in Fig. 4.12, and consisted of a central molybdenum cathode rod C, surrounded by a concentric cylindrical molybdenum anode A in which was

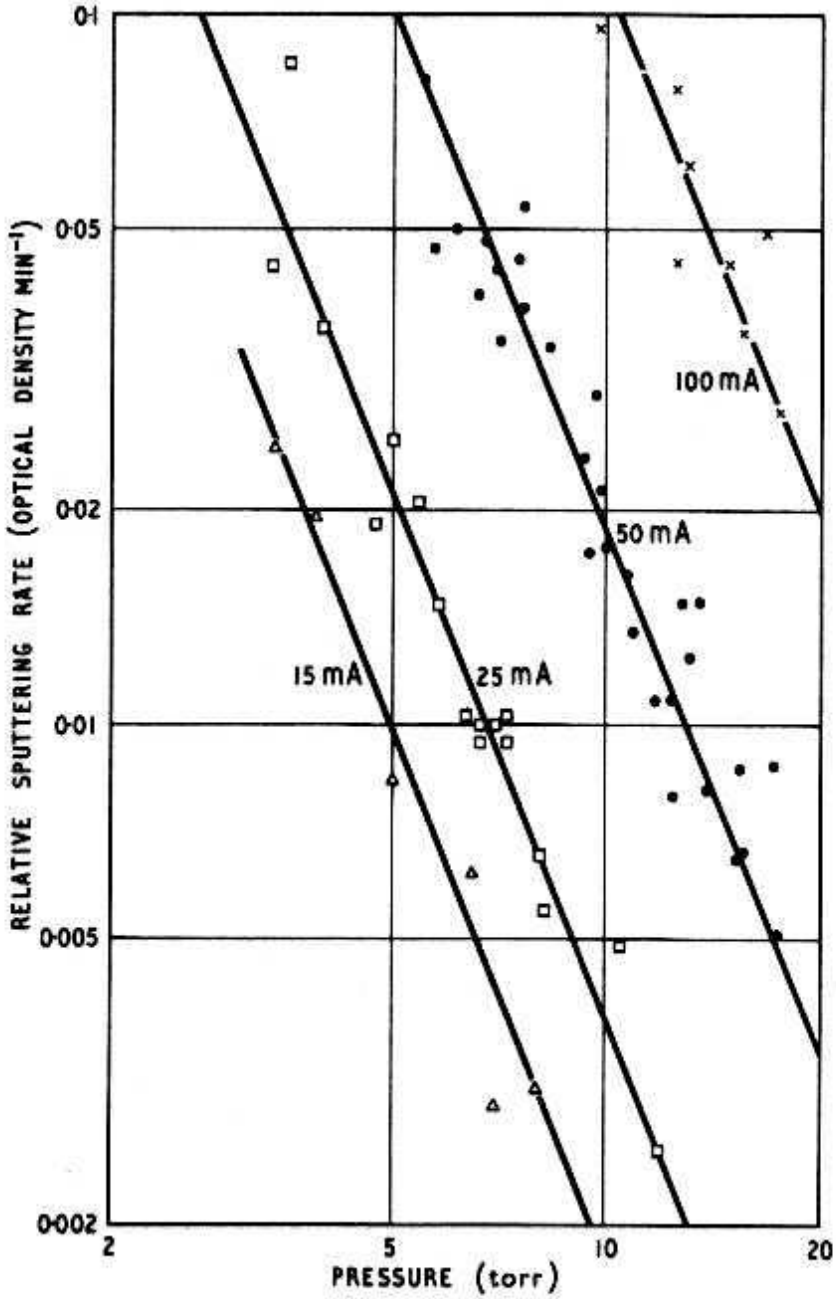


Fig. 4.13. Relative sputtering rate of molybdenum as a function of pressure for 99% neon + 1% argon. Stocker [60]

cut a square aperture through which the sputtered particles could pass. A glass cylindrical collector G round the outside of the anode could be rotated magnetically, so that a series of deposits could be made in one tube. A tungsten filament, F, mounted inside the anode was used as the light source, and the transmission was monitored

by a photometer calibrated in terms of optical density. A series of graphs of optical density against time was obtained at various currents and pressures for neon and for 99% neon + 1% argon. These graphs were normally straight except for slight initial curvature, and had slopes proportional to the rate of deposition. It was found that the sputtering rate $M = Qi$ could not be expressed adequately in terms of the voltage, current, and pressure by means

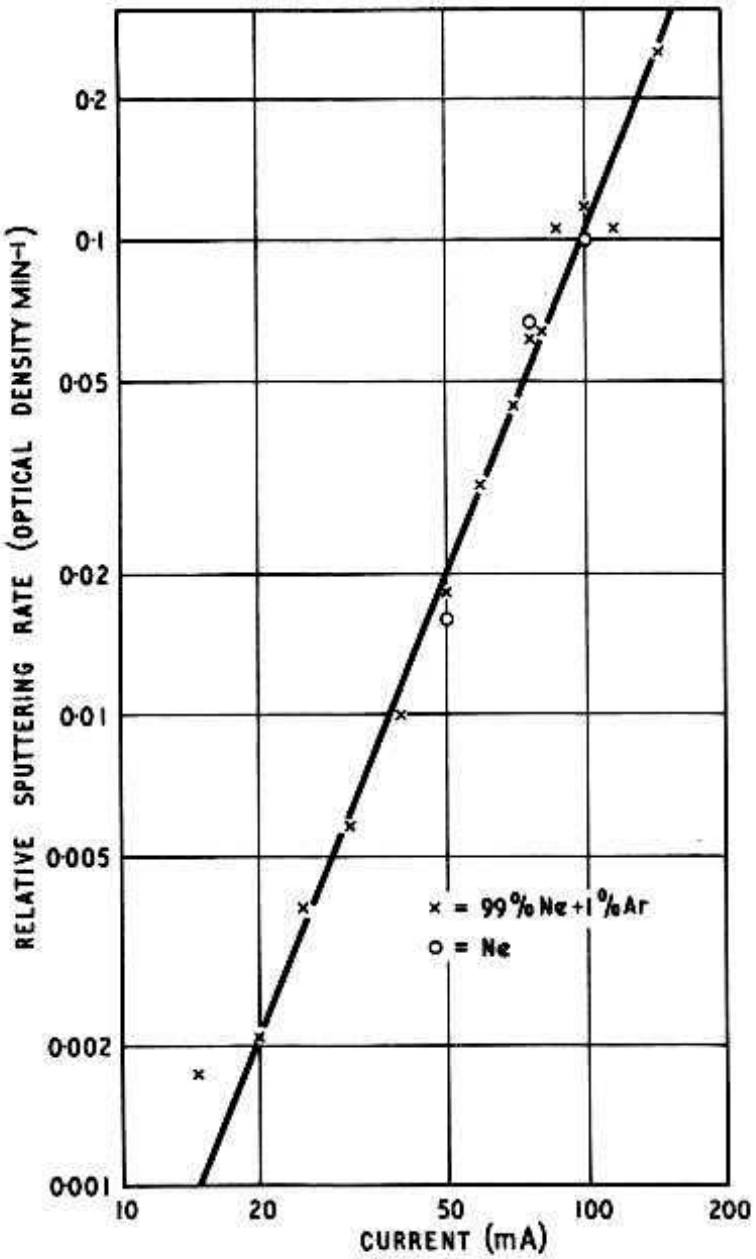


Fig. 4.14. Relative sputtering rate of molybdenum as a function of current at a pressure of 10 torr. Stocker[60]

of Equation 4.15 or 4.16 but could be related to the current and pressure alone. Fig. 4.13 shows a series of curves of relative sputtering rate as a function of pressure in 99% neon + 1% argon on a logarithmic scale for different currents. As can be seen the negative slope for all currents is similar, of magnitude 2·5. Similar curves

were obtained for pure neon. In Fig. 4.14 the rate of sputtering at 10 torr is plotted against current; for both gases the points lie on one straight line with positive slope of approximately 2·5. Hence the rate of sputtering can be related to the current i and pressure p by an empirical expression of the form

$$M = Qi = C\left(\frac{i}{p}\right)^{2\cdot 5} \tag{4.18}$$

where C is a constant which is approximately the same for both neon and 99% neon + 1% argon. Stocker found almost negligible sputtering of molybdenum in helium under similar conditions, but if 0·085% neon was introduced the sputtering rate increased to a value comparable with that in pure neon.

Takatsu and Toda[61] also found the sputtering rate from a molybdenum cathode in neon and in argon over a similar pressure range to depend on some power of the current, i.e.

$$M \propto i^n \tag{4.19}$$

where n was near 3

They measured the amount of sputtered material by the change of frequency of a quartz oscillator as a collector. Takatsu and Toda attempted to explain their results theoretically by inserting the average ion energy in the cathode fall from Townes' theory into Langberg's equation (Equation 4.13). Although they derived an expression of the same form as Equation 4.19 this can only be considered as fortuitous, since their treatment does not take account of the diffusion of the sputtered particles and there is some confusion in their use of temperature scales. The author and Sethna[62] extended the work of Stocker to sputtering under pulsed current conditions using tubes of the same construction. With square wave pulses of peak current i_p, and mean current i_m, the fraction of the time during which the pulse is applied is equal to i_m/i_p; thus the sputtering rate from Equation 4.18 should be given by

$$M = C\left(\frac{i_p}{p}\right)^{2\cdot 5} \frac{i_m}{i_p} \tag{4.20}$$

so that for constant mean current

$$M \propto i_p^{1\cdot 5} \tag{4.21}$$

The sputtering rates found experimentally, however, were lower than given by Equation 4.21 and in fact, from the curves shown in Fig. 4.15, gave a power law of approximately 0·4. This was

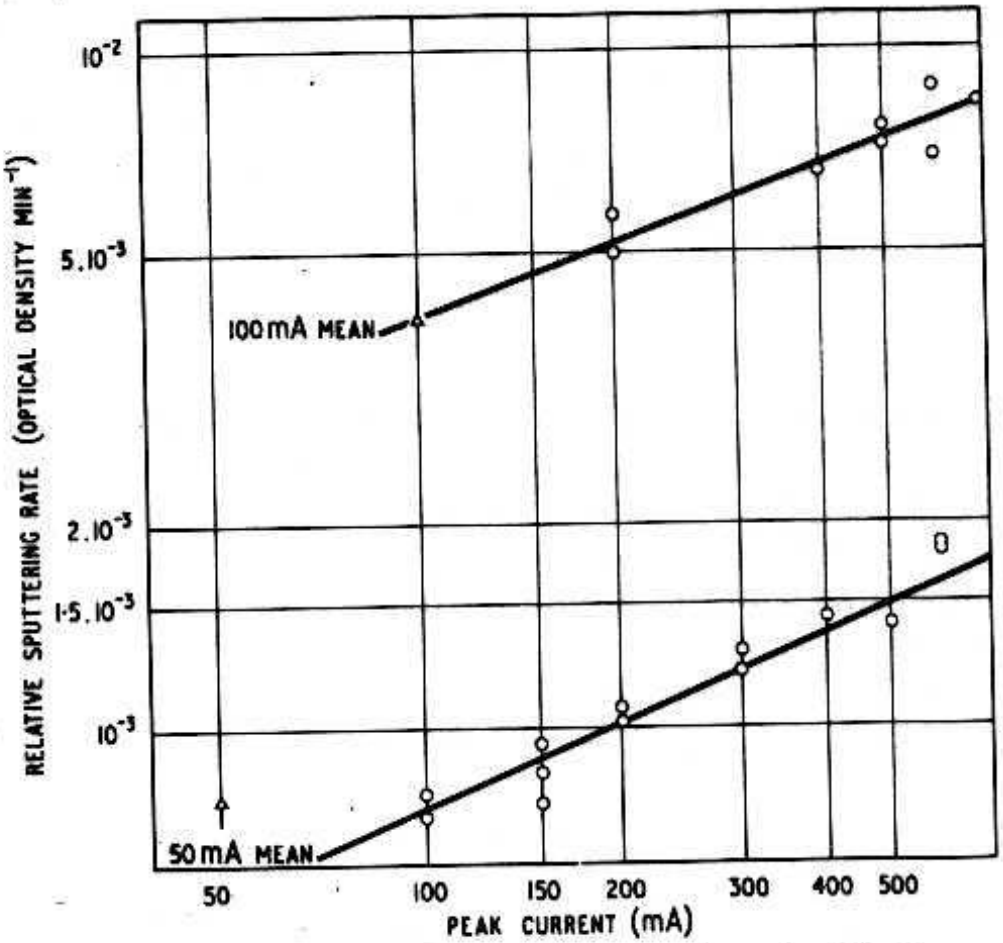


Fig. 4.15. Relative sputtering rate of molybdenum under pulsed discharge conditions as a function of the peak current. Weston and Sethna[62]

explained by the fact that in the case of d.c. sputtering the density in the cathode region decreases with increasing current due to the rise in cathode temperature, whereas under pulse conditions with a constant mean current, the cathode temperature and thus the gas density remains constant as the peak current is varied. Hence the variation of sputtering rate with direct current is partly due to the density changes in the cathode region, and since this depends on the electrode and tube geometry (cf. Section 3.4) it is likely that the power law relating sputtering rate to current will vary over a small range for tubes of different geometry.

4.4.2. THE CLEANING OF ELECTRODES BY SPUTTERING

The cleaning of metal cathodes in a discharge tube by sputtering to provide reproducible glow discharge characteristics (breakdown and maintaining potentials) was reported by Penning and Moubis[63] in 1946. Since then commercial production of close tolerance cold cathode tubes has been based on their findings. Essentially the cathode is extensively sputtered, firstly to remove any oxide or other contaminating layer from the cathode surface, and secondly

to deposit a metal layer on to the glass envelope in order to prevent gases being evolved from the glass. Further, as the metal diffuses through the gas any active component is 'gettered' and thus the gas is also cleaned. The effectiveness of the sputtered metal film as a 'getter' has been demonstrated by Robinson[64] who has used a mass spectrometer for comparing the gases evolved from glass bulbs having various metal films sputtered on to them. Some of his results are given in Table 4.4 for the relative output of mass 18

Table 4.4

Mass 18 output on reheating of glass surface with a sputtered film as a fraction of the output on first stoving without the film according to Robinson[63]

Sputtered film	*50°C*	*100°C*	*200°C*	*400°C*
None	0·5	0·13	0·20	0·28
Al	0·009	0·0028	0·0016	0·006
Mo		—all very small—		0·01
Zr	0·012	0·032	0·011	0·007
Ni	0·18	0·028	0·032	0·013
Ta	0·25	0·13	0·089	0·020
Cu	0·89	0·63	0·79	0·56

(water vapour) at different bulb temperatures. The bulbs had previously been baked at 450°C for 1 hr. For example, for aluminium the amount of water vapour coming from the film is 0·0016–0·009 of that coming from the glass heated for the first time, whereas glass on second heating only showed a decrease to a value between 0·13 and 0·5 over the same temperature range. A sputtered film of molybdenum gives very little output, and at certain temperatures acts as a very strong getter reducing the water vapour content below the level of residual gases in the mass spectrometer system itself. It is noticeable that copper gave higher values than the glass bulb without a metal film, and may explain the difficulty in obtaining reproducible values with copper electrodes (cf. Section 3.2.2).

The rate of sputtering during cleaning must be limited by the temperature that the bulb reaches, if this is too high the gases evolved counteract the cleaning effect of sputtering. The processing schedule is, therefore, often prolonged particularly if the tube is filled to a high pressure. By using pulse sputtering with high pulse currents and low duty cycles the processing schedule can be considerably reduced, even though the sputtering rate is not as high as predicted. This is common practice particularly for high pressure tubes or those with cathodes of large surface area.

4.4.3. GAS CLEAN-UP IN GLOW DISCHARGES

If a discharge is passed in a low pressure of inert gas $\sim$ 1 torr it is frequently found that the gas pressure decreases slowly with time. It had long been suspected that this phenomenon, known as gas clean-up, was connected with sputtering but it was not until recently that the mechanism has been explained. This was mainly as a result of the work of Blodgett and Vanderslice[65] who succeeded in relating the gas clean-up of inert gases to the sputtering rate of a metal electrode in the discharge, and, in another experiment, to the ion-current density variation over the surface of an ion collector. Clean-up occurs at any surface on which sputtered material is being deposited, by trapping excited gas atoms under the deposit. If a small negative potential is applied to the surface, clean-up is assisted by attracting positive ions which are then trapped. However if the surface is sufficiently negative and the ion current density great enough to cause resputtering of the material, gas is released again, and an equilibrium state is soon reached after which no more gas is trapped. Hence gas clean-up does not occur at the cathode of a discharge as long as the ion current density is uniform (unless sputtered material is received from another electrode at a faster rate than it is being sputtered off). Using a different approach the general findings were confirmed qualitatively by Hall[66] who measured gas clean-up in hot cathode tubes. In a situation where sputtering is not significant, for example at much lower pressures such as experienced in ion gauges, or in helium, clean-up occurs by penetration of the atoms into the lattice of the target.[67]

In glow discharge tubes where the pressure is greater than 10 torr, gas clean-up is not normally a serious problem. However in certain tubes where the current density is high, i.e. a hollow cathode tube, it can be a contributory factor to the end of life of the tube.

4.4.4. SCREENING OF THE ENVELOPE FROM A DEPOSITED LAYER

In applications where light output is required from the glow discharge, the deposition of cathode material on to the tube envelope is detrimental. However, it is possible to screen part of the envelope by interposing a suitable obstacle between it and the cathode. When the mean free path of the sputtered particles is small compared with the dimensions of the tube, they will travel by a diffusion process, taking a zig-zag path as they collide with the gas atoms. As a result an obstacle placed in the path of the sputtered

atoms does not leave a completely clear area corresponding to the optical shadow, but an area of less dense coating which extends beyond the optical shadow. Braun and Oskam[68] have calculated this screening effect for certain geometries, in particular a cylindrical object placed in a cylindrical tube, using the linear diffusion equation:

$$\frac{dn}{dt} = D\Delta n + \varepsilon$$

Where D is the diffusion coefficient, n the density of sputtered particles and ε the initial source density. They considered the

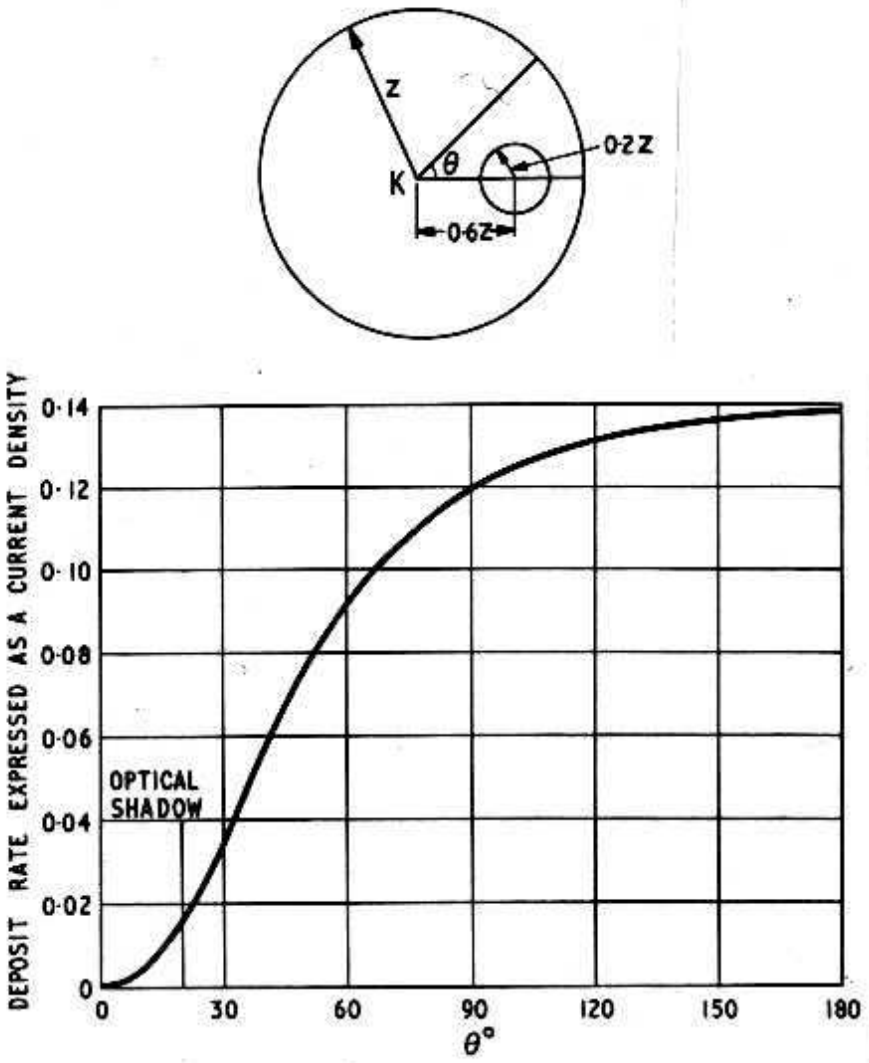


Fig. 4.16. The distribution of sputtered material on the walls of a cylindrical bulb from a co-axial cathode in the presence of a cylindrical obstacle, for the dimensions shown. Braun and Oskam[68]

problem as essentially a stationary diffusion phenomenon ($dn/dt = 0$) which could be treated by potential theory. The solution for the case of a line source (cathode) mounted centrally in the envelope, and for particular dimensions of the tube and the obstacle, is

illustrated in Fig. 4.16 where the deposit rate expressed as a current density is plotted as a function of the angle made with the line joining the cathode to the centre of the cylindrical object. It is seen that even at an angle double that equivalent to the optical shadow the sputtering rate is less than half that to the opposite wall. Qualitative agreement with this theory was obtained by the author and Schagen[69] in the design of an indicator tube (see Section 9.5.2) in which a window along the length of a cylindrical tube was provided by a rod anode. If a second cylindrical obstacle is mounted at the side of the first with a narrow gap between them say of the order of their radii, then from considerations of Fig. 4.16 the amount of sputtered material reaching the bulb in the area in front of the gap would be quite small. A slit in a metal electrode will also provide a window, the amount of sputtered material reaching the envelope depending on the slit width, the thickness of the said metal electrode, the gas pressure and the total sputtering rate. For example the author has carried out comparative tests in which a rod cathode was surrounded by a cylindrical anode of 15 mm diameter and 0·15 mm wall thickness mounted in a 20 mm diameter bulb. The anode had various size slits cut along its length of width 0·5–4 mm. The cathode rod was sputtered, and the thickness of the film deposited on the envelope with time determined by its optical transmission, for different gas pressures. From the results it was seen that the thickness of the film increased markedly with slit width, and that for a given slit width the deposit increased with pressure. At 50 torr in 99% neon+1% argon a heavy sputtered layer was deposited in front of a 2 mm slit after 24 hrs, reducing the transmission by 50%. For the same time and current at 20 torr, however, the change in transmission in front of the 2 mm slit was not measurable. Thus a slit width which is a hundred times the mean free path presents an effective screen to prevent sputtered atoms passing through it.

Clearly a series of slits or better still a mesh of suitable hole size can be employed to keep a large area of the bulb free from a sputtered deposit. The mesh can be of reasonably high optical transmission. For example an 80 in^2 mesh having an 80% optical transmission was used by the author in an experimental tube and compared with a similar tube without the mesh. After a 300 hr sputtering period when the tube without the mesh had an opaque layer of deposit on the bulb, the optical transmission of the glass opposite the mesh had not changed measurably. The use of grids or meshes for this purpose is now firmly established and many display tubes incorporate the principle in their design (see Chapter 9).

REFERENCES

1. PENNING, F. M. and MOUBIS, J. H. A., *Proc. Koninkl.Ned. Akad. Wetenschap.*, **43**, 41, (1940)
2. WEHNER, G. K., *Advances in Electronics and Electron Physics*, Academic Press Inc., New York, *Vol.* **7**, 239, (1955)
3. RESEARCH STAFF OF G.E.C., *Phil. Mag.*, **45**, 98, (1923)
4. KINGDON, K. H. and LANGMUIR, I., *Phys. Rev.*, **22**, 148, (1923)
5. GÜNTHERSCHULZE, A. and MEYER, K., *Z. Physik*, **62**, 607, (1930)
6. MEYER, K. and GÜNTHERSCHULZE, A., *Z. Physik*, **71**, 279, (1931)
7. GÜNTHERSCHULZE, A., *Vacuum*, **3**, 360, (1953)
8. WEHNER, G. K. and MEDICUS, G., *J. Appl. Phys.*, **25**, 698, (1954)
9. WEHNER, G. K., *Phys. Rev.*, **93**, 633, (1954)
10. WEHNER, G. K., *Phys. Rev.*, **102**, 690, (1956)
11. WEHNER, G. K., *Phys. Rev.*, **108**, 35, (1957)
12. WEHNER, G. K., *Phys. Rev.*, **112**, 1120, (1958)
13. FETZ, H., *Z. Physik*, **119**, 590, (1942)
14. WEHNER, G. K., *J. Appl. Phys.*, **30**, 1762, (1959)
15. MORGULIS, N. D. and TISHCHENKO, V. D., *Soviet Phys. JETP*, **3**, 52, (1956) and *Bull. Acad. Sci. USSR, Phys. Ser.*, **20**, 1082, (1956)
16. STUART, R. V. and WEHNER, G. K., *J. Appl. Phys.*, **33**, 2345, (1962)
17. LAEGREID, N. and WEHNER, G. K., *J. Appl. Phys.*, **32**, 365, (1961)
18. ROSENBERG, D. and WEHNER, G. K., *J. Appl. Phys.*, **33**, 1842, (1962)
19. STRACHAN, J. F. and HARRIS, N. L., *Proc. Phys. Soc.*, **B69**, 1148, (1956)
20. WOLSKY, S. P., *Phys. Rev.*, **108**, 1131, (1957)
21. WOLSKY, S. P. and ZDANUK, E. J., *Phys. Rev.*, **121**, 374, (1961)
22. KOEDAM, M., *Philips Res. Rep.*, **16**, 101, (1961)
23. WEIJSENFELD, C. H., HOOGENDOORN, A. and KOEDAM, M., *Physica*, **27**, 763, (1961)
24. TIMOSHENKO, G., *J. Appl. Phys.*, **12**, 69, (1941)
25. KEYWELL, F., *Phys. Rev.*, **87**, 160, (1952)
26. KEYWELL, F., *Phys. Rev.*, **97**, 1611, (1955)
27. BRADLEY, R. C., *Phys. Rev.*, **93**, 719, (1954)
28. WEHNER, G. K. and ROSENBERG, D., *J. Appl. Phys.*, **32**, 887, (1961)
29. KENKNIGHT, C. E. and WEHNER, G. K., *J. Appl. Phys.*, **35**, 322, (1964)
30. MCKEOWN, D., *Rev. Sci. Instr.*, **32**, 133, (1961)
31. YONTS, O. C., NORMAND, C. E. and HARRISON, D. E., *J. Appl. Phys.*, **31**, 447, (1960)
32. GRØNLUND, F. and MOORE, W. J., *J. Chem. Phys.*, **32**, 1540, (1960)
33. ROL, P. K., FLUIT, J. M. and KISTEMAKER, J., *Physica*, **26**, 1000, (1960); *J. Appl. Phys.*, **34**, 690, (1963)
34. ALMEN, O. and BRUCE, G., *Nucl. Instr. and Methods*, **11**, 257 and 279, (1961)
35. PEROVIC, B. and COBIC, C., *Proc. Intern. Conf. Ionisation Phenomena Gases, 5th, Munich*, North Holland Publishing Co., 1165, (1962)
36. RAMER, C. E., NARASIMHAM, M. A., REYNOLDS, H. K. and ALLRED, J. C., *J. Appl. Phys.*, **35**, 1673, (1964)
37. HINES, R. L. and WALLOR, R., *J. Appl. Phys.*, **32**, 202, (1961)
38. AKISHIN, A. I., VASILEV, S. S. and ISAEV, L. N., Bull Acad. Sci. U.S.S.R. Phys. Ser. (USA), **26**, 1379, (1964)
39. JORGENSON, G. V. and WEHNER, G. K., *J. Appl. Phys.*, **36**, 2672, (1965)
40. KOEDAM, M., *Philips Res. Rep.*, **16**, 101 and 266, (1961)
41. ANDERSON, G. S. and WEHNER, G. K., *J. Appl. Phys.*, **31**, 2305, (1960)
42. ANDERSON, G. S., *J. Appl. Phys.*, **33**, 2017, (1962)
43. ANDERSON, G. S., *J. Appl. Phys.*, **34**, 659, (1963)
44. SOUTHERN, A. L., WILLIS, W. R. and ROBINSON, M. T., *J. Appl. Phys.*, **34**, 153, (1963)

45. MAGNUSON, G. D. and CARLSTON, C. E., *J. Appl. Phys.*, **34**, 3267, (1963)
46. TOWNES, C. H., *Phys. Rev.*, **65**, 319, (1944)
47. STARK, J., *Z. Elektrochem*, **15**, 509, (1909)
48. HENSCHKE, E. B., *Phys. Rev.*, **106**, 737, (1957)
49. HENSCHKE, E. B., *J. Appl. Phys.*, **28**, 411, (1957)
50. LANGBERG, E., *Phys. Rev.*, **111**, 91, (1958)
51. SILSBEE, R. H., *J. Appl. Phys.*, **28**, 1246, (1957)
52. ANDERSON, G. S., *J. Appl. Phys.*, **34**, 659, (1963)
53. HARRISON, D. E., *Phys. Rev.*, **102**, 1473, (1956)
54. HARRISON, D. E., *Phys. Rev.*, **105**, 1202, (1957)
55. HARRISON, D. E., *J. Chem. Phys.*, **32**, 1336, (1960)
56. HARRISON, D. E., *J. Appl. Phys.*, **32**, 924, (1961)
57. GÜNTHERSCHULZE, A., *Z. Physik*, **38**, 575, (1926)
58. GÜNTHERSCHULZE, A. and TOLLMIEN, W., *Z. Physik*, **119**, 685, (1942)
59. ROCKWOOD, G. H., *Trans. Amer. Inst. Elect. Engrs.*, **60**, 901, (1941)
60. STOCKER, B. J., *Brit. J. Appl. Phys.*, **12**, 465, (1961)
61. TAKATSU, K. and TODA, T., *Proc. Intern. Conf. Ionisation Phenomena Gases, 5th, Munich*, North Holland Publishing Co., Amsterdam, 96, (1962)
62. WESTON, G. F. and SETHNA, D. R., *Brit. J. Appl. Phys.*, **14**, 723, (1963)
63. PENNING, F. M. and MOUBIS, J. H. A., *Philips Res. Rep.*, **1**, 119, (1946)
64. ROBINSON, N. W., *Internal Report of Mullard Research Lab.*, **No. V.P.L.** 54, (1954)
65. BLODGETT, K. B. and VANDERSLICE T. A., *J. Appl. Phys.*, **31**, 1017, (1960)
66. HALL, R. F., *J. Electron. Control*, **11**, 101, (1961)
67. CARMICHAEL, J. H. and TRENDELENBURG, E. A., *J. Appl. Phys.*, **29**, 1570, (1958)
68. BRAUN, G. and OSKAM, H. J., *Internal Report No. 8137 of the Natuurkundig Laboratorium*. Philips Eindhoven
69. WESTON, G. F. and SCHAGEN, P., Electron Eng., **36**, 21, (1964)

CHAPTER 5

The Temporal Growth and Decay of a Discharge

5.1. Classification of Delay Times

In the preceding chapters the gas discharge has been considered in an equilibrium state. If, however, the voltage to an electrode system in a gas is applied suddenly, a finite time will elapse before the equilibrium state is reached. This time lapse in the case of the breakdown and establishment of a glow discharge is known as the time lag or delay, and consists of two parts, the statistical lag and the formative lag. The former, a random process, is the time taken for an initiatory electron to be produced at a suitable position in the interelectrode space by the incident radiation. The formative lag is the time taken for the ionising and secondary emission processes to increase the current from the initiatory value to the equilibrium value set by the circuit impedance. Similarly, on removing the applied voltage the ions and electrons take a finite time to recombine or diffuse out of the interelectrode space i.e. for the gas to revert to the non-ionised state. This is known as the de-ionisation time.

This temporal growth and decay of a discharge is of considerable importance in the design and use of cold cathode tubes. The statistical lag can be of the order of seconds, although in most practical devices it is reduced by the introduction of a priming agent such as a radioactive source. The formative lag is relatively short, of the order of microseconds, and dictates the pulse length required for firing a tube. Since the presence of ionised gas depresses the breakdown potential, reapplication of the potential to a tube after extinction must be delayed by a period governed by the de-ionisation time. In a practical device this will be very much longer than the other delay times, of the order of milliseconds for inert gases, and it is this time factor which normally limits the

frequency at which a cold cathode tube can be switched. Although shorter delays are obtainable with gases such as hydrogen, the time effects are inherent in a cold cathode tube and always introduce a limitation to their applications.

5.2. The Statistical Lag

5.2.1. THEORY OF THE STATISTICAL LAG

As the name implies, the time taken for the production of an electron or electrons which will be sufficiently multiplied to lead to breakdown is a random quantity, and any theory dealing with it must be based on statistical considerations. Such a theory is that of von Laue[1] which, although published in 1925, is equally valid today. Von Laue calculated the statistical lag in terms of the number N_o of primary electrons formed per second by say, irradiation (equivalent to the initiatory current i_o, see Chapter 1), and the probability P that a primary electron will give rise to an avalanche of sufficient multiplication to cause breakdown. He showed that if the number of lags n, out of n_o, that exceeded t seconds are plotted against t, then the curve of n/n_o against t is exponential, of the form

$$\frac{n}{n_o} = \exp(-N_oPt) \tag{5.1}$$

i.e. purely statistical in time.

Hence the probability that a discharge will occur within a time dt after a time t from the instant of applying the voltage is

$$N_oP\exp(-N_oPt)dt \tag{5.2}$$

and the average or mean statistical lag t_s is given by

$$t_s = \frac{\int_o^\infty N_oPt\exp(-N_oPt)dt}{\int_o^\infty N_oP\exp(-N_oPt)dt} = \frac{1}{N_oP} \tag{5.3}$$

(For a more detailed study see Acton and Swift[2]).

The value of P in terms of the multiplication factor $M = \gamma(\exp\eta V - 1)$ has been calculated by Braunbeck[3] and by Hertz.[4] The calculations are based on the fact that although the average value of M need not be a whole number, for each of the N_o primary electrons M must be equal to 0, 1, 2, etc. Every time $M = 0$ an avalanche process is stopped, and by considering the

various possibilities, the fraction of the primary electrons which give rise to avalanches which are stopped can be derived. Wijsman[5] extended this principle to include fluctuations of γ and $\exp \eta V$ about average values, and for the case of primary electrons being produced at the cathode derived the expression,

$$P = 1 - \frac{1}{M} \tag{5.4}$$

where M is the average value of the multiplication factor. This equation results in a curve of P against M, which starts at $P = 0$ for $M = 1$ and rises asymptotically to 1 at high values of M.

Since M increases with increasing applied potential a similar curve can be plotted for P against voltage or overvoltage (the value of applied voltage in excess of the static breakdown potential V_s). Sample curves of P against percentage overvoltage are shown in Fig. 5.1, derived by Fuller[6] for particular tubes using empirical data.

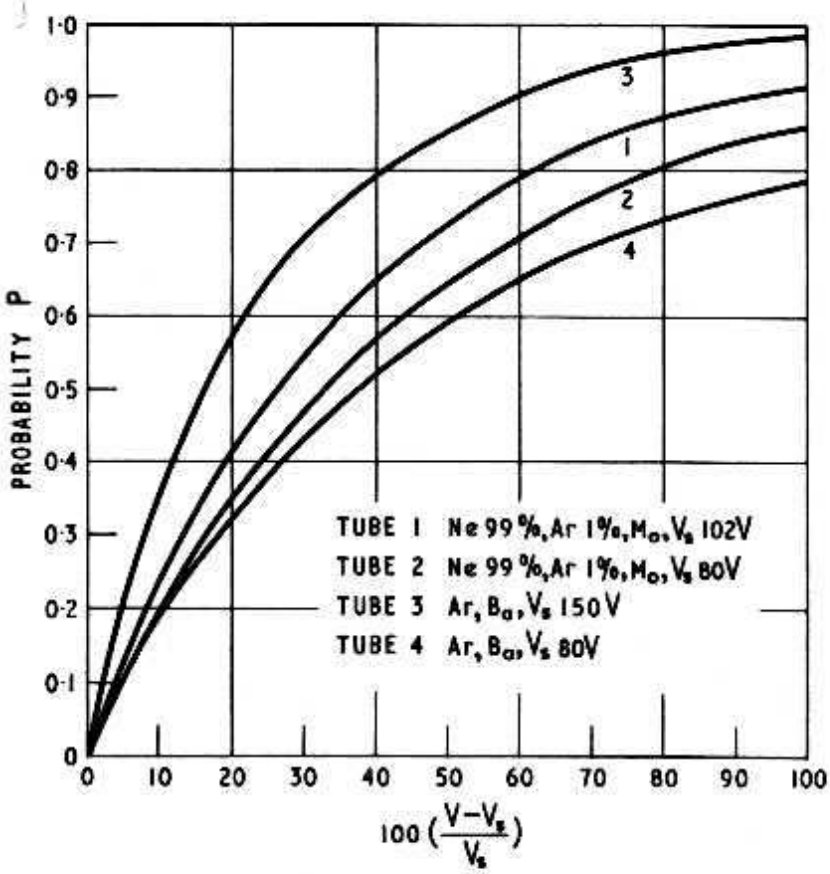


Fig. 5.1. Probability P of an electron producing breakdown as a function of the percentage overvoltage, $100\,(V-V_s)/V_s$. Fuller[6]

Thus from Equation 5.3 it is seen that the mean statistical lag is infinite at the static breakdown potential V_s, but as the applied voltage is raised, t_s decreases rapidly to approach a limiting value at high over-voltages, which depends on the initiatory current i_o. If this initiatory current is provided by cosmic radiation, where the

number of electrons per second in the gap N_o is less than 5 electrons, then t_s will be greater than 0·2 seconds even at high overvoltages, and to reduce this to the order of microseconds requires a priming current of the order 10^{-13} A i.e. $N_o = 10^6$. Although t_s can be numerically defined in any circumstance, individual times can have any magnitude because the statistical lag is a random process. Nevertheless t_s is a useful parameter in comparing tubes, and also for defining the probability of a tube firing when a given pulse is applied. For example, for a pulse duration of $2t_s$ the probability of a tube firing is approximately 90% whereas for $5t_s$ it is 99%.

5.2.2. METHOD OF MEASURING THE MEAN STATISTICAL LAG

Two methods can be used to determine the mean statistical lag: (1) by measuring the number of breakdowns in time t and from the data plotting $\log_e (n/n_o)$ against t, which gives t_s from the slope (Equation 5.1), or (2) by measuring each delay from n_o pulses and obtaining t_s from the distribution. The latter method has been used by Llewellyn Jones,[7] who employed an attenuated replica of the overvoltage pulse to gate and terminate a 1 Mc/s scaler to which the output of a 100 Kc/s oscillator is fed for the duration of each time lag. The length of each time lag is thus designated in terms of the number of cycles m_n of oscillation on the scaler unit. From the results of n_o pulses,

$$t_s = \frac{1}{n_o}\sum_{1}^{n_o} m_n/f$$

where $f =$ the frequency of the oscillator. In the other method a train of rectangular pulses of suitable amplitude is applied to the tube, and counters are employed to count firstly the number of applied pulses n_o and secondly the number of times the tube breaks down n_1. These counts are then related to the mean statistical lag by Equation 5.1 which can be written as

$$1 - \frac{n_1}{n_o} = \exp\left(-\frac{t}{t_s}\right) \tag{5.5}$$

where t is the pulse duration.

In both the methods the number of pulses n_o used must be large (of the order of a thousand). Also unless the formative lag is negligible

in comparison with the statistical lag it must be taken into account when calculating t_s.

5.2.3. EXPERIMENTAL INVESTIGATION

Many studies have been made of the time delays involved in breakdown, but often these have combined statistical and formative lags without separating the two components.

Early measurements on statistical lags were aimed at confirming the general theory, and the data obtained were limited by the techniques available at that time. Such measurements were those of von Laue[1] and Zuber,[8] Strigel[9] and Tilles.[10] Of interest was the fact that the last named author found an exponential relationship between t_s and overvoltage,

$$t_s = \frac{3{\cdot}7 \times 10^{-3}}{i_o^{0{\cdot}76}} \exp\left[\frac{-39(V - V_s)}{V_s}\right] \tag{5.6}$$

which does not directly result from the theory.

The introduction of pulse circuit techniques and scaler counters has allowed simpler and more accurate measurement of the mean statistical lag. However, most of the experiments reported have been aimed at the determination of N_o and little information is available on delays and overvoltage requirements which is applicable to cold cathode tubes. For example Llewellyn Jones and his co-workers[7, 11, 12] have been interested in the initiating current, in the absence of external illumination, in parallel-plate tubes filled to pressures of the order of an atmosphere. They found quite large rates of electron production ($N_o \simeq 10^4$ electrons $\sec^{-1}$) which they assumed were emitted from the cathode. The emission could be obtained for comparatively modest electric fields of the order of 10^5 V cm^{-1}, particularly if a layer of fine metal dust or an oxide were present. In the former case the metal particles enhanced the field. Removal of the dust by an air blast was found, in a particular experiment, to reduce the emission by a factor of 10^3. For oxide or tarnish layers the field also is considered to be enhanced by positive ions on the surface; the emission from a carefully polished clean surface could be increased by as much as an order by exposure to the atmosphere for a few hours.

Although the emission mechanism appeared to be complex, over a limited range the relation between the electric field and the emission current was analytically the same as the Fowler-Nordheim

field emission equation. Applying this equation gave effective work functions more than an order lower than generally accepted values and with emitting areas of atomic dimensions, suggesting that the emission takes place from a few sites of low work function on the metal surface. They found that the time lags were independent of pressure, and of the gas composition in this range.

Measurements of statistical lag in glow discharge tubes have been made in the author's laboratory. In parallel-plate tubes using molybdenum electrodes and neon gas the author found delays at low overvoltages which were measurable with a stop watch if the tube was in complete darkness. However, a considerable reduction occurred in the presence of ambient light. The average statistical lag, plotted against the percentage overvoltage, for one of these experimental tubes is shown in Fig. 5.2. A similar reduction in the

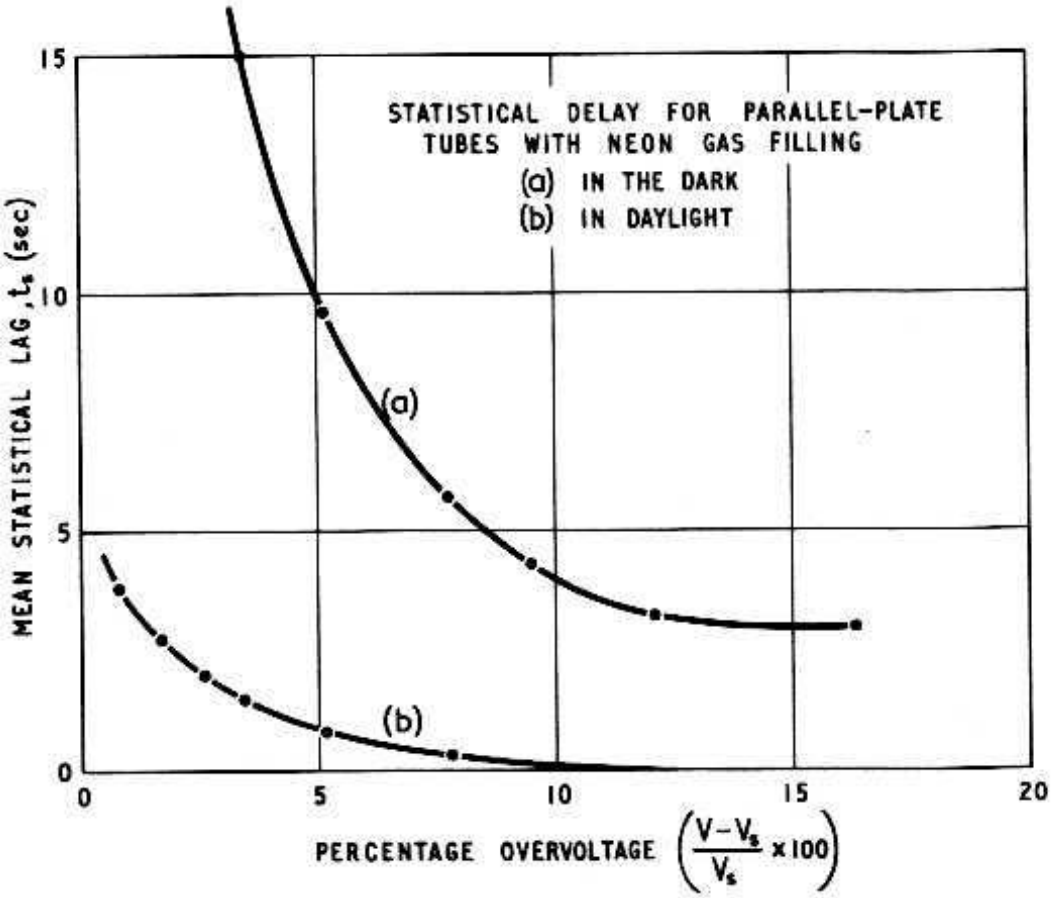


Fig. 5.2. Mean statistical lag for a parallel-plate tube with neon gas filling as a function of the percentage overvoltage. Curve (a) measured in darkness, curve (b) in daylight

mean statistical lag when exposed to daylight had been observed in commercial stabilisers where no priming was present with molybdenum cathodes and neon-argon gas mixtures although the lags involved were rather shorter. This effect was rather surprising since the work function of molybdenum is 4·2 eV corresponding

to a threshold wavelength for photoemission of 2960Å, which should not penetrate the glass envelope. Fox[13] investigated this using a pulse technique where the number of breakdowns out of the total number of pulses applied is counted for pulses of known amplitude and duration. He found that photoemission occurred at much longer wavelengths ($\sim$ 6000 Å) corresponding to a photoemitting surface in the discharge tube with a work function as low as 2·0 eV. In a series of experiments, Fox established that the photoemission arose from the glass surface upon which a thin sputtered layer had been deposited, particularly the bead to the cathode support, and he attributed the effect to sodium released from the glass either by electrolysis or by bombardment during the sputtering process.

Using the same technique Fox[14] has also investigated the effect on the statistical lag of introducing tritium into glow discharge tubes. Tritium is a particularly useful radioactive priming source for inclusion in glow discharge tubes, since it has a half-life of thirteen years and produces only β-particles of less than 18·3 keV energy, which cannot penetrate the glass envelope. For experimental tubes, 0·2 mc of tritium in 20 torr of 99% neon + 1% argon reduced t_s to approximately 0·3 μsecs, and 1 microcurie gave t_s approximately 0·6 msec. Since most of the tritium was taken up by the sputtered film, the actual delay depended on the geometry and degree of sputtering. Many commercial stabiliser tubes with pure metal cathodes now employ tritium as a priming agent to reduce the strike delay on switching on the apparatus in which they are incorporated.

5.3. The Formative Lag

5.3.1. ANALYSIS OF THE FORMATIVE LAG

The analysis of the growth of a discharge from the initiatory electron to a self-maintained discharge (a Townsend discharge with a current density of the order of 10^{-7}A cm^{-2} is usually considered) has been tackled by a number of workers. In general Townsend's basic concepts were employed, and no account was taken of space-charge distortion. The simplest analyses have been those which assume that the secondary electron emission is due solely to positive ion impact at the cathode, γ_i. For example, Druyvesteyn and Penning[15] have given an approximate treatment based on the premise that one ionisation cycle takes a time t_i, the transit time of an ion from anode to cathode. Thus a primary current i_o increases after time t_i to

$i_o(1+M)$ (where M is the multiplication factor) after $2t_i$ to $i_o(1+M+M^2)$, and after time t_F (formative time) $= nt_i$ to

$$i = i_o \frac{M^{n+1}-1}{M-1}$$

$$= i_o \frac{M^{(t_F/t_i+1)}-1}{M-1} \tag{5.7}$$

Since $t_F/t_i \gg 1$ for a current increase of, say from 10^{-14} A to 10^{-7} A, Equation 5.7 may be simplified to

$$t_F = t_i \frac{\log_e[(M-1)i/i_o]}{\log_e M} \tag{5.8}$$

A rather better approach was made by Schade,[16] which compared favourably with experimental results. His analysis consisted of equating the current at the anode to that at the cathode at time t, and integrating over the total current range. If $i(t)$ is the current at time t then he showed that if the ion transit time is small

$$i(t) = i_o + M\left[i(t) - t_i \frac{\mathrm{d}i(t)}{\mathrm{d}t}\right] \tag{5.9}$$

which by integration gives

$$t_F = \frac{t_i}{M-1} \log_e\left[(M-1)\frac{i}{i_o}+1\right] \tag{5.10}$$

providing $M-1 \ll 1$.

At breakdown M approaches 1 in which case Equation 5.10 can be written as

$$t_F = t_i\left(\frac{i}{i_o}\right) \tag{5.11}$$

i.e. the current builds up linearly with time.

For small overvoltages, where γ can be considered as constant and α increases by a small amount $\delta\alpha$, Schade obtained an approximate expression

$$t_F = \frac{\mathrm{a}}{V-V_s} \exp\left(\frac{-\mathrm{b}}{V}\right) \tag{5.12}$$

where a and b are constants at a given pressure p derived from the relation

$$\alpha = \mathrm{a} \exp\left(\frac{-\mathrm{b}}{V}\right)$$

(cf. Equation 1.4, when a = Ap and b = Bpd). Agreement with his experimental results is shown in Fig. 5.3.

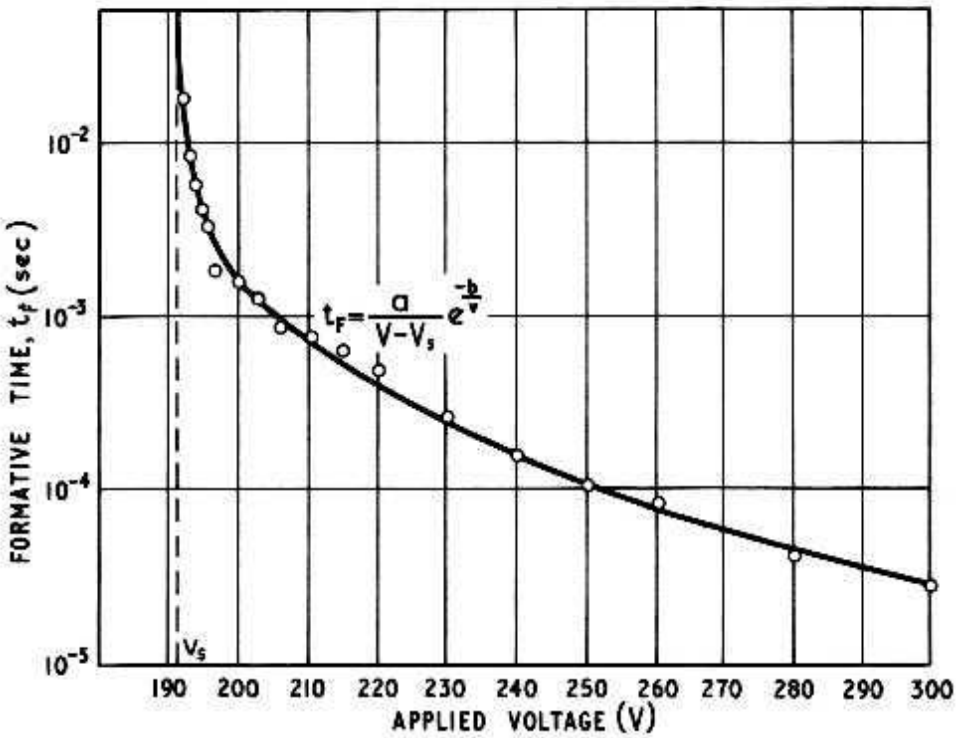


Fig. 5.3. Formative time lag for a neon discharge as a function of the applied voltage. The full curve was obtained from Equation 5.12 by choosing the constants to obtain the best match with the experimental points. Schade[16]

A more accurate calculation has been made by Steenbeck[17] which was later developed by Bartholomyczyk[18] based on the assumption that the current built up exponentially with time. The solution was similar to Equation 5.10, although rather more complex. Bartholomyczyk used the continuity equations for the positive ion and electron current. The general analysis has been expanded by a number of workers to include metastable action at the cathode by introducing the average transit time t_m of the metastable atoms, and to include photoemission, where the electron transit time t_e is important. Thus Raether[19] gives an expression for t_F which takes photoemission into consideration:

$$t_F = \frac{M_i t_i + M_p t_e}{M_i + M_p - 1} \log_e \left[1 + (M_i + M_p - 1) \frac{i}{i_o} \right] \qquad (5.13)$$

where M_i and M_p are the multiplication factors due to the ions and photons respectively.

At higher pressures (particularly of molecular gases) formative times, which appeared too short (10^{-7}–10^{-9} sec) to be consistent with the above general analyses, were observed at modest overvoltages. Further at such pressures the first manifestations of breakdown are narrow brightly luminous filamentary spark channels (streamers) which often take a zig-zag path between the electrodes. These factors suggested another breakdown mechanism occurring at such pressures, and around 1940 the 'streamer'

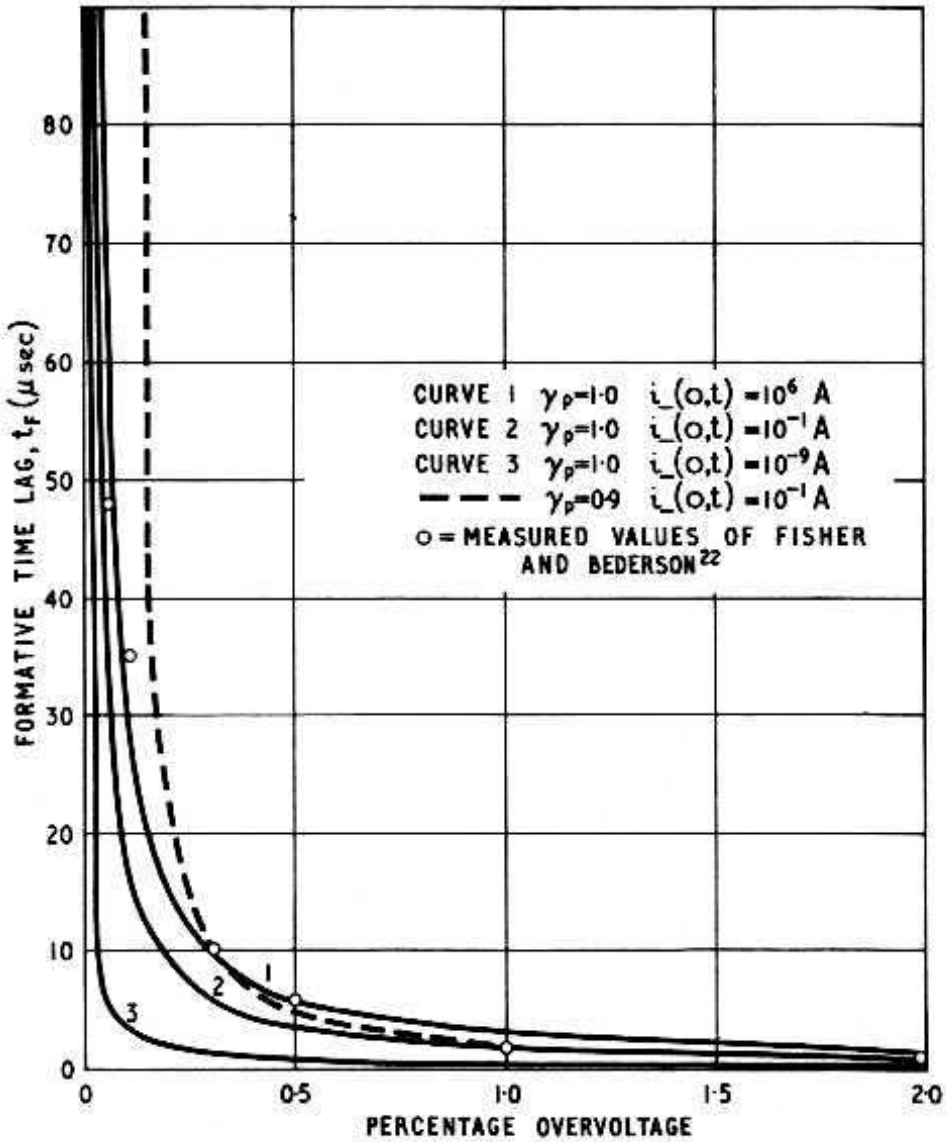


Fig. 5.4. Comparison of the theoretical curves of formative time lag as a function of percentage overvoltage (assuming various values of i_- (o, t) and γ_p) with experimental values. Dutton et al.[23]

mechanism was evolved (e.g. Meek and Craggs[20]). The explanation of the mechanism postulated the formation of a streamer from a single electron avalanche, and the advancement of the streamer across the gap by a photo-ionisation process. The transition from the avalanche to the streamer was considered to occur when the

radial field produced by the positive ions at the head of the avalanche was of the order of the externally applied field. Since the time involved in the growth of the avalanche would be considerably longer than that required for the streamer to advance across the gap, the former was considered as the formative lag. Calculations based on this premise by Fletcher[21] for highly overvolted gaps in air gave good agreement with experimental results. At low overvoltages ($< 0{\cdot}1\,\%$), however, Fisher and Bederson[22] found that, even at atmospheric pressure, formative time lags rose to 100 μsec or more for uniform field air-gaps. They suggested that this could be due to the need to increase the field by space-charge distortion, at the low overvoltages, before a single avalanche could build up sufficiently to produce a streamer. However, Dutton *et al*[23] showed that Fisher and Bederson's results could be equally well explained on the Townsend mechanism by an analysis of Davidson[23] if photo-emission at the cathode (γ_p) was included in the secondary coefficient, i.e. photo-ionisation in the gas does not have to be considered. Comparison of Fisher and Bederson's curves with computed curves obtained from Davidson's analysis by a suitable choice of γ_p to γ_i ratio is shown in Fig. 5.4. Davidson's analysis is based on the solution to the continuity equations in a similar way to that of Bartholomyczyk, but with more accurately defined boundary conditions. Also his solution satisfies the initial conditions completely, giving any prescribed initial distribution of positive and negative charge.

The exact solution is complex, but Davidson has also given an approximate expression for the current at time t_F

$$i_-(x,t)\exp-\alpha x = \frac{i_o}{1-M_i-M_p}+\mathrm{C}\exp\left[\lambda\,(t_F-\frac{x}{v_-})\right] \qquad (5.14)$$

where C is a constant which can be chosen to give approximate agreement with initial conditions and λ is given a real value satisfying $\mathrm{F}(d) = 0$ where

$$\mathrm{F}(x) = 1-\frac{\alpha\gamma_i}{\Phi}(\exp\Phi x-1)-\frac{\alpha\gamma_p}{\Psi}(\exp\Psi x-1)$$

$$\Phi = \alpha-\frac{\lambda}{v}$$

$$\Psi = \alpha-\frac{\lambda}{v_-}-\mu$$

$$\frac{1}{v} = \frac{1}{v_+}+\frac{1}{v_-}$$

v_+ and v_- being the drift velocities of the ions and electrons respectively, and μ being the absorption coefficient of the gas to the photons.

The expression can be further simplified if it is assumed that $\mu = 0$ and that there is no current in the gas at $t = 0$; then

$$t_F = \frac{1}{\lambda}\log_e\left[1+\frac{i}{i_o}(M_i+M_p-1)\right] \tag{5.15}$$

which is similar to Equation 5.13.

Davidson[24] has since extended his analysis to include other secondary mechanisms at the cathode, in particular photoemission from delayed radiation i.e. radiation from excited states which have a long lifetime. Although these analyses give formative time lags in agreement with experimental results, the theoretical current or voltage versus time curves do not agree with empirical data. This is illustrated in Fig. 5.5 where theoretical curves for two ratios

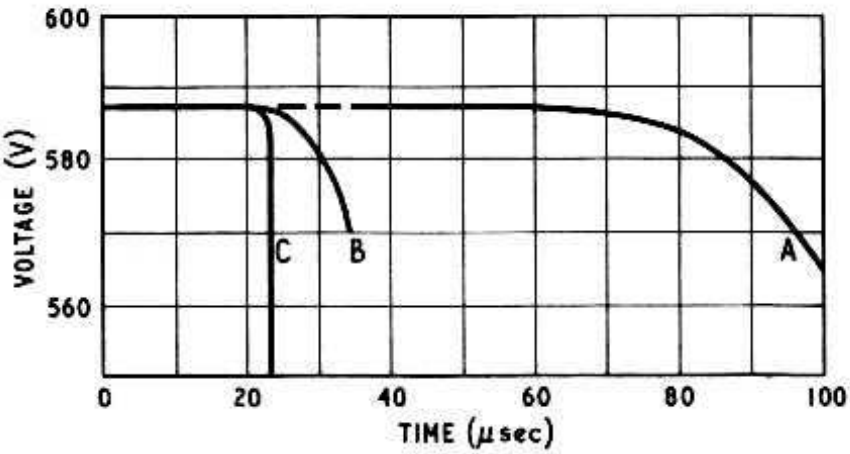


Fig. 5.5. Comparison of an experimental voltage-time curve in hydrogen with curves calculated from Davidson's analysis on the assumption of undistorted fields. (A) *theoretical curve* $\gamma_p = 70\%$, $\gamma_i = 30\%$; (B) *theoretical curve* $\gamma_p = 88{\cdot}3\%$, $\gamma_i = 11{\cdot}7\%$. (C) *experimental curve. Davies, Evans and Llewellyn Jones*[26]

of γ_p to γ_i are compared with an experimental voltage versus time curve. The discrepancy is due to space-charge distortion which tends to accelerate the current growth in the later stages of the build up, thus leading to a more rapid fall in voltage. Analyses of t_F to include space-charge distortion have been made by Ward[25] and Davies, Evans and Llewellyn Jones[26] using computers. Ward combines the continuity equations with one dimensional Poisson's equation. Although this results in current-time characteristics of the same form as those obtained experimentally, Davies *et al.* point out that this procedure is only valid if the radius of the discharge is infinite, or large compared with the electrode separation. Accordingly they have given a method which takes into account

the radius of the discharge, R. Since the effective radius of the discharge, which must be used to calculate the field distortion, is not known, it introduces another variable into the calculations. Consequently it is possible to find several combinations of R, γ_i and γ_p which will produce a collapse of voltage at time t_F. This is shown by Table 5.1, where values of γ_p for a hydrogen discharge are calculated for different R values to give the value of t_F found experimentally.

Table 5.1 indicates that the experimental formative lag can be explained with a secondary process entirely due to positive ions when the discharge is taken to be narrow.

Table 5.1

Summary of γ_p obtained by Jones and Llewellyn Jones and those computed by Davies *et al.*[26] for different radii of the discharge, R

P (*torr*)	E/p (V cm^{-1} *torr*$^{-1}$)	γ (*total*) *J. and L.J.*	γ_p as a fraction of γ		
			J. and L.J.	*computed with* $R = 2{\cdot}9$ cm*	*computed with* $R = 0{\cdot}1$ cm
1·6	140·8	0·0124	0·95	0·75	0
3·59	88·1	0·00313	0·7	0·73	0
5·35	70·36	0·00283	0·7	0·8	0

*2·9 cm was the radius of the electrodes.

Davidson's analysis taking into account space charge is probably the most satisfactory theory of the formative lag, giving complete agreement with experiment if the parameters, such as R, are suitably chosen. However, for low pressures, the simple expressions of Schade are probably adequate and give a reasonable indication of the behaviour to be expected in a cold cathode tube. For example, in Table 5.2, t_F has been calculated for a 10% overvoltage applied to a 1 cm gap between molybdenum electrodes using Equation 5.12. Values of A and B have been taken from Chapter 1 and V_s from Chapter 2.

Table 5.2

Values of t_F calculated from Equation 5.12 for a 10% overvoltage applied to a 1 cm gap between molybdenum electrodes

Gas	V_s *at* 20 *torr*	t_F *at* 20 *torr*	V_s *at* 40 *torr*	t_F *at* 40 *torr*
Helium	340 V	102 msec	480 V	47 msec
Neon	320 V	4 msec	470 V	500 μsec
Argon	510 V	1·3 msec	760 V	90 μsec

These values are of the right order and show the decrease with increase of pressure which is generally observed in this pressure range.

5.3.2. EXPERIMENTAL VALUES OF FORMATIVE LAG

The formative lag can be measured simply by applying a voltage pulse to the electrode system, and observing the point at which the potential collapses (see Fig. 5.5). Although this point may be ill-defined, particularly at low pressures, and depends on the amplification in the oscilloscope, the error introduced is small for most

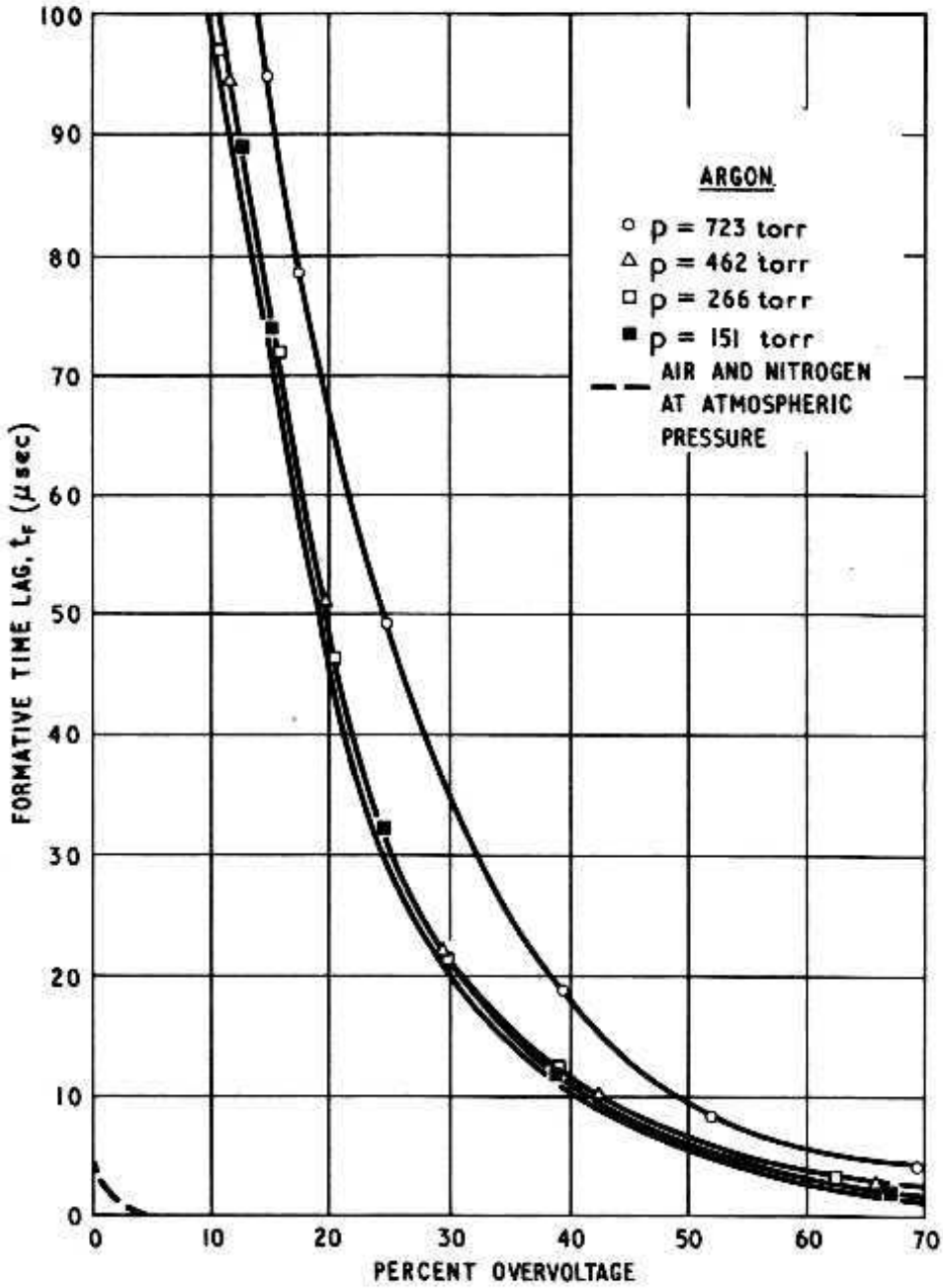


Fig. 5.6. Formative time lag as a function of percentage overvoltage for argon at various pressures. Electrode spacing d = 1·0 cm. Kachickas and Fisher[28]

practical purposes. The gap must be sufficiently primed to reduce the statistical lag to negligible proportions.

The majority of the published measurements of t_F have been for gases at high pressures, made in an attempt to establish whether Townsend or streamer mechanism is involved and whether a

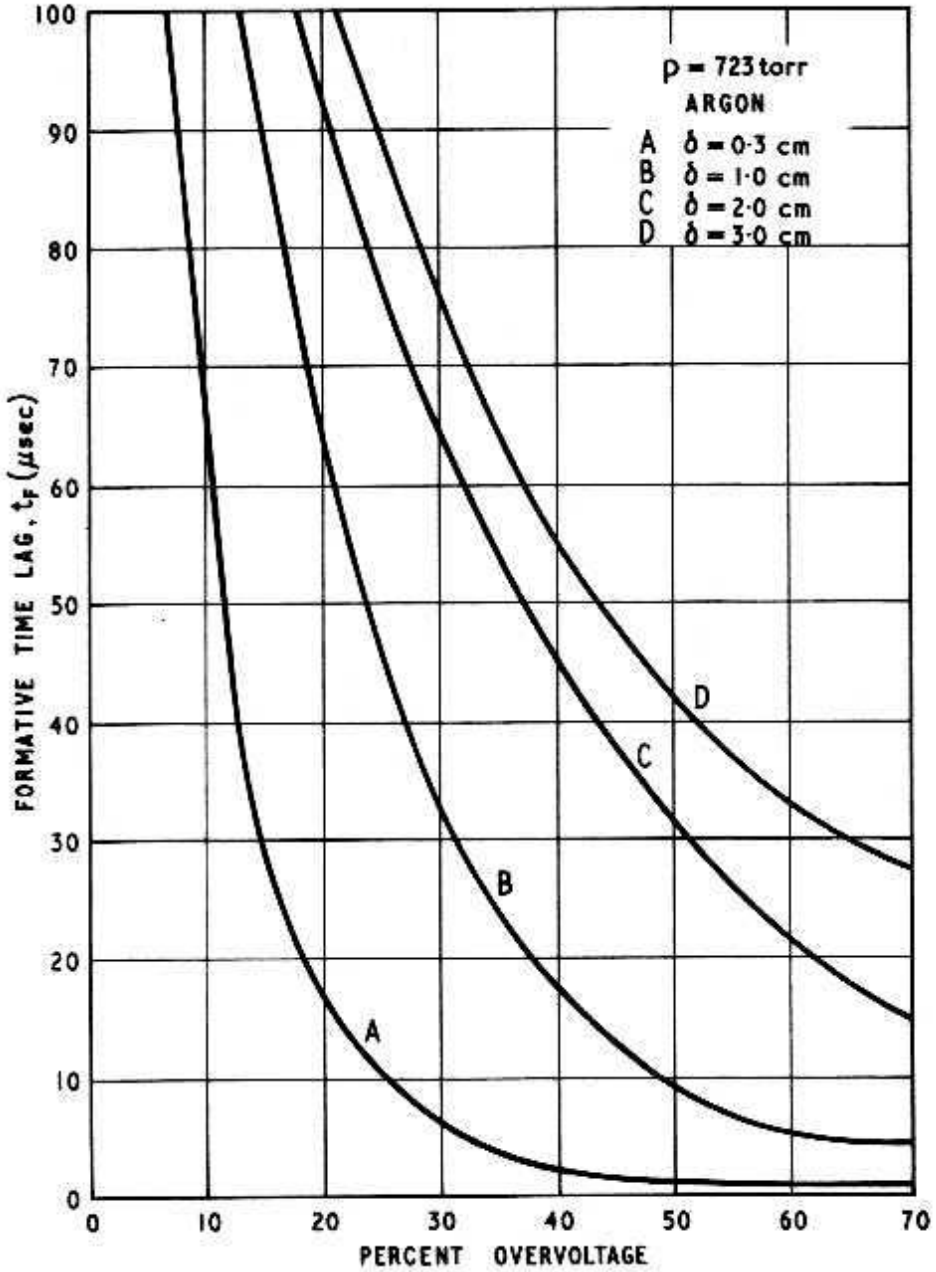


Fig. 5.7. Formative time lag as a function of overvoltage for argon at 723 torr at various electrode spacings. Kachickas and Fisher[28]

transition occurs from one to the other. Thus Fisher and Bederson's measurements in air, mentioned in the last section, were followed up by Kachickas and Fisher in nitrogen[27] and argon[28] in the pressure range 150–700 torr. Their values of t_F against overvoltage for a 1 cm gap in argon at various pressures are shown in Fig. 5.6. t_F was very much larger for the inert gas and showed a marked dependence on pressure not observed in air or

nitrogen. As expected, t_F depends on the interelectrode spacing, increasing with larger gaps, and this is demonstrated in Fig. 5.7.

A fairly extensive study of t_F as a method of examining the secondary processes in a discharge has been carried out at Swansea University under the leadership of Llewellyn Jones. Morgan[29] measured formative time lags at low overvoltages in hydrogen with copper electrodes and showed the dependence on E_s/p where E_s is the field at breakdown. His results are reproduced in Fig. 5.8.

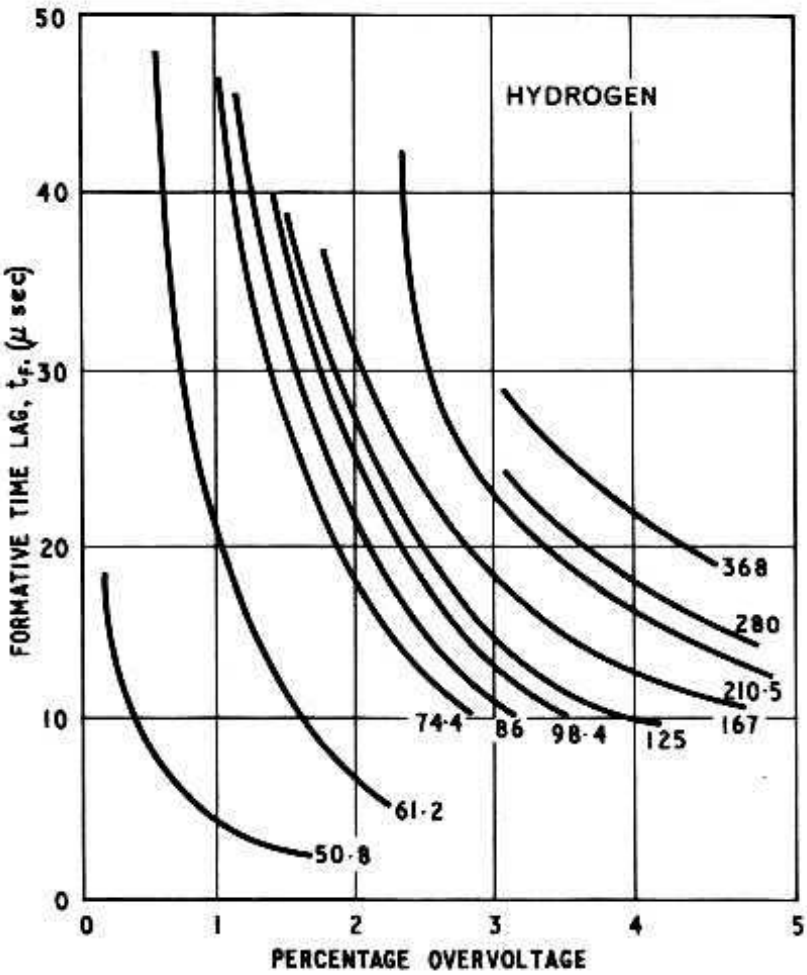


Fig. 5.8. Formative time lag as a function of percentage overvoltage. The numbers on the curves refer to the corresponding values of $E_{s/p}$ in $V cm^{-1} torr^{-1}$. Morgan[29]

By sputtering aluminium, platinum or nickel on to the copper cathode Llewellyn Jones and Jones[30] showed that the values in hydrogen were dependent on the cathode material, since the values and ratios of γ_i and γ_p varied with the work function of the cathode surface. Later, measurements were extended to helium, and Davies, Llewellyn Jones and Morgan[31] gave values of t_F for low overvoltages for pressures between 5 and 38·5 torr, which were in the tens of milliseconds range. Their values are given in Fig. 5.9. They could not explain these results in terms of electron emission by positive ion and/or metastable atom bombardment of the cathode,

and suggested delayed radiation as a possible cause of the long delays. Menes[32] had previously interpreted the lag in argon by such a process.

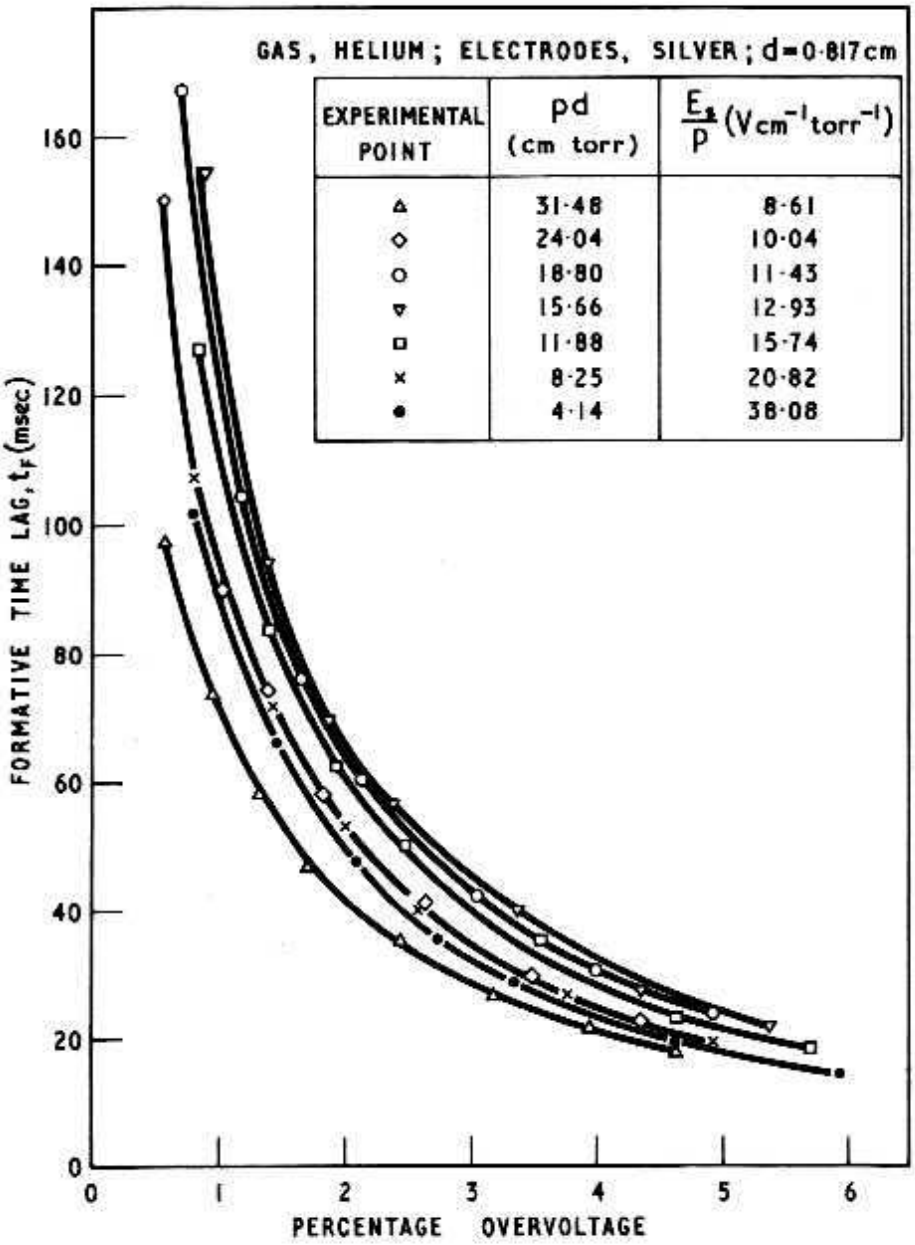


Fig. 5.9. Formative time lag as a function of the overvoltage for helium at various E_s/p values. Davies, Llewellyn Jones and Morgan[31]

The author has measured the formative lag in a number of experimental and commercial tubes, and some typical results for individual tubes are given in Fig. 5.10. In the experimental tubes and the two stabilisers, priming was effected by placing a small radioactive source (about 1μ curie of cobalt 60) close to the tube, whereas in the trigger tube and decade counter a priming discharge was passed to a subsidiary electrode. This made a difference of several orders to i_o and explains the smaller t_F for the latter tubes.

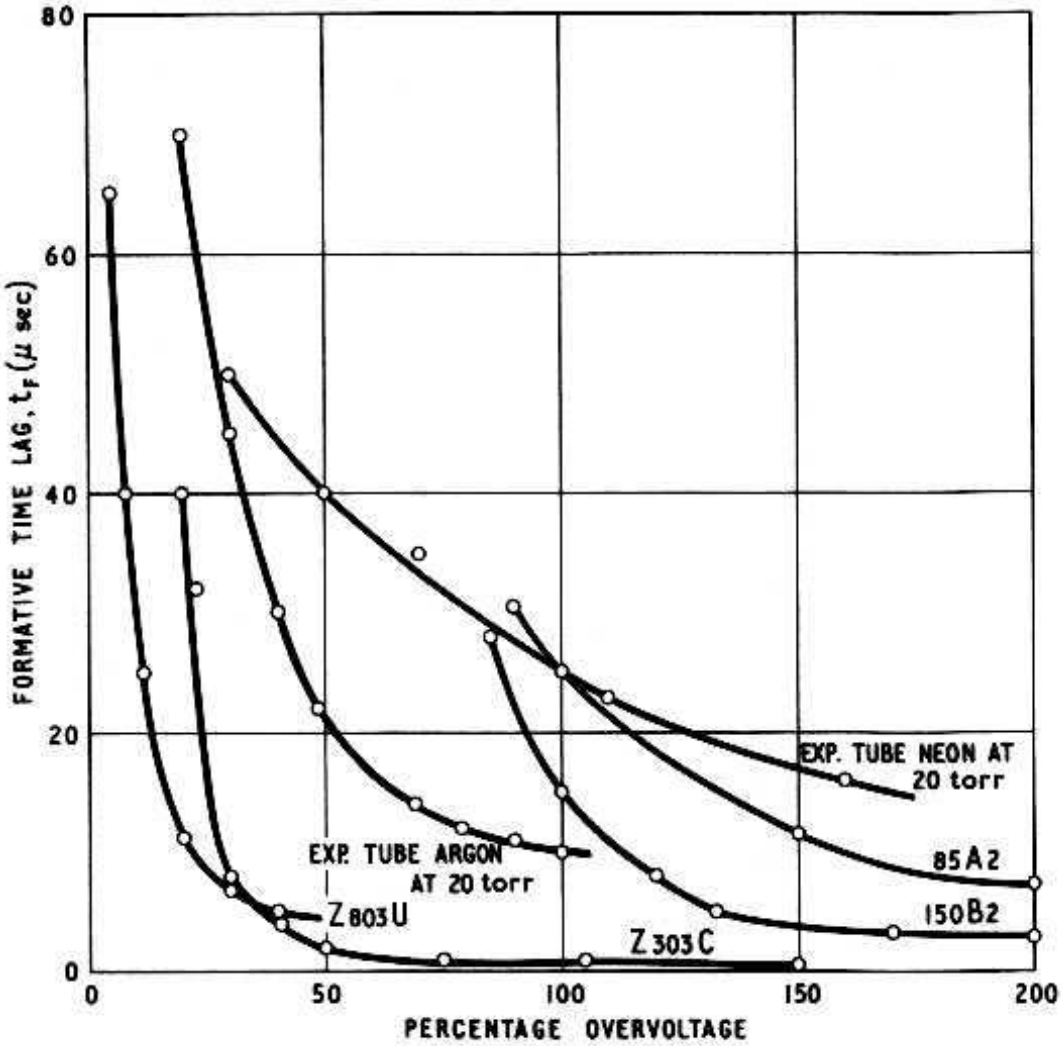


Fig. 5.10. Formative time lag as a function of the overvoltage for a number of sample glow discharge tubes

Most inert gas tubes show similar formative times, although argon usually gives rather lower values.

5.4. Deionisation Time

The electrons and ions created in a gas discharge can be lost, other than to the electrodes, by two processes, (1) the recombination of an ion and an electron to form a neutral particle, and (2) diffusion of the charged particles out of the discharge area where they subsequently recombine on the walls. These loss mechanisms are associated with a delay time, so that when the production of charged particles is stopped by the removal of the electrode potentials a finite time will elapse before the ions and electrons disappear and the gas reverts to the non-ionised state. This is known as the deionisation time. However, the recovery time, defined as the time taken to recover to the condition where the original

breakdown characteristic applies, rather than complete deionisation, is of more practical interest.

Both loss mechanisms can take place simultaneously although, in general, one will predominate depending on the pressure and nature of the gas. The loss mechanism can be determined by examining the variation in charge density during the deionisation period. This can be determined by probe electrodes in the gas, but the most satisfactory method is by microwave techniques. In practice the loss mechanism is of less interest than the variation of breakdown characteristic during deionisation and this can be determined by applying probing pulses to the electrodes.

5.4.1. THEORY OF DEIONISATION

5.4.1.1. *Recombination*

If an ionised gas is considered where the concentration of positive ions is N_+ and electrons (or negative ions) is N_- per c.c., then the loss of ions dN_+ in a time dt will be proportional to N_+ and N_- and also to dt i.e.

$$dN_+ = -\alpha N_+ N_- dt \tag{5.16}$$

where α is the constant of proportionality and is called the coefficient of recombination.

For a plasma such as the positive column $N_- = N_+$ and therefore

$$\frac{dN}{dt} = -\alpha N^2 \tag{5.17}$$

(where N is the number of ions or electrons) which integrated gives the concentration at time t after removing the source of ionisation, gives

$$\frac{1}{N} - \frac{1}{N_o} = \alpha t \tag{5.18}$$

or

$$N = \frac{N_o}{N_o \alpha t + 1}$$

where N_o is the initial concentration at $t = 0$.

From this equation it is seen that N/N_o decreases hyperbolically with time at a rate depending on $N_o\alpha$.

A number of recombination processes are possible in an ionised gas depending on whether more than one gas ion is present. The possible processes are as follows:

1 $A^+ + e \rightarrow A + h\nu$		radiative recombination
2 $A^+ + B^- \rightarrow AB + h\nu$		
3 $AB^+ + e \rightarrow A + B$		dissociative recombination
4 $A^+ + B^- \rightarrow A' + B'$		non-radiative mutual charge exchange
5 $A^+ + e + B \rightarrow A + B$		three body recombination
6 $A^+ + B^- + C \rightarrow A + B + C$ or $AB + C$		

Where A' and B' are excited states.

The probability of recombination occurring depends on the time interval for which particles are close together during a collision. This in turn depends on the velocities of the particles or their kinetic energies, the lower they are the longer the particles will remain near each other. For positive and negative ions where high energy losses occur on collision with neutral molecules, the likelihood of the ions having low kinetic energy when they approach each other is large and the probability of recombination is high. Recombination between an electron and a positive ion is less likely since an electron loses little energy at collision.

From the same considerations, as the pressure increases, and hence the collision frequency with the neutral molecules, the probability of recombination also increases. At pressures above an atmosphere, however, the probability, as reflected in the value of α, decreases again due to the decreasing chance of a recombination collision occurring between the charged particles. In the inert gases negative ions do not occur, and therefore recombination between electrons and positive ions only has to be considered.

5.4.1.2 *Diffusion*

Charged particles are said to diffuse in a gas if they move from points of high concentration to those of low concentration due only to the difference in concentration and not for instance to a field. The motion will be due to the thermal velocity of the particles, which will consequently take zig-zag paths. If diffusion in one direction only is considered (the x-direction) then the number of particles dN, diffusing in time dt, is given by the expression

$$\frac{dN}{dt} = -DA\frac{dN}{dx} \tag{5.19}$$

where D is a constant for a given gas pressure, known as the diffusion coefficient, A is the cross-sectional area, and the negative sign indicates that the motion occurs in the direction of decreasing concentration. It can be shown that D depends on the mobility μ and kinetic energy, as expressed by its temperature T, of the charged particle. Accordingly

$$\left.\begin{aligned} \mathrm{D}_+ &= \frac{\mu_+ k T_+}{e} \quad \text{for ions} \\ \text{and} \qquad \mathrm{D}_e &= \frac{\mu_e k T_e}{e} \quad \text{for electrons} \end{aligned}\right\} \tag{5.20}$$

where k is the Boltzmann constant.

When electrons and positive ions are present in a gas with equal concentration, then, providing the density is reasonably high, as the electrons diffuse out of the plasma (normally electrons diffuse faster than ions) the positive ions will set up a space charge, and the resulting field will tend to retard the electrons whilst accelerating the diffusion of positive ions. Thus an equilibrium will be established when the number of electrons and ions leaving the plasma are equal. The motion of the charged particles can then be considered jointly as having a common velocity governed by a common diffusion constant, this is known as ambipolar diffusion, and the constant is called the ambipolar diffusion coefficient D_a.

Thus the diffusion in one direction is given by

$$\frac{\mathrm{d}N_e}{\mathrm{d}t} = \frac{\mathrm{d}N_+}{\mathrm{d}t} = -\mathrm{D}_a A \frac{\mathrm{d}N_e}{\mathrm{d}x} \tag{5.21}$$

D_a is related to the diffusion coefficient of the ions and electrons respectively by the expression

$$\mathrm{D}_a = \frac{\mu_e \mathrm{D}_+ + \mu_+ \mathrm{D}_e}{\mu_e + \mu_+} \tag{5.22}$$

which from Equation (5.20), and since $\mu_e \gg \mu_+$, can be simplified to

$$\mathrm{D}_a = \frac{k\mu_+}{e}(T_e + T_+) \tag{5.23}$$

In a plasma where the electrons are in thermal equilibrium with the ions, $T_e = T_+ = T_g$ the gas temperature, and then D_a depends on the mobility of the positive ions.

For the practical tube, where diffusion occurs in all directions and the density distribution is time dependent, solution of the diffusion equation is extremely complex and N can only be derived for special cases where simplifications can be introduced. Assuming D_a is independent of time and space co-ordinates, then the equation can be solved for elementary shapes of containers if the boundary condition at the walls of $N = 0$ can be applied.[33, 34] The solution obtained consists of a product of a function of time and a function of space co-ordinates. The variation of N as a function of time proves to be an exponential of the form $\exp(-t/\tau)$ where τ is a time constant involving the reciprocal of the ambipolar diffusion coefficient. A plot of the logarithm of the density N against time, therefore, results in a straight line with a negative slope. Hence

$$\frac{dN}{dt} = \frac{-N}{\tau} \tag{5.24}$$

Equations 5.17 and 5.24 can be combined to give a general equation for deionisation where recombination and diffusion are both present:

$$\frac{dN}{dt} = \alpha N^2 \frac{N}{\tau} \tag{5.25}$$

The predominant process will depend on the gas composition and also on the pressure (see Section 5.4.2.1). As already stated, for a plasma where the electrons are in thermal equilibrium with the gas atoms, D_a is proportional to μ_+, which in turn depends on $1/p$. Thus τ will depend on p and the deionisation rate due to diffusion will increase with decreasing pressure. In general, diffusion will therefore dominate the deionisation process at low pressures, less than say 10 torr, whilst at higher pressures it will become insignificant in comparison with recombination.

5.4.2. EXPERIMENTAL INVESTIGATION OF DEIONISATION

Measurements of deionisation phenomena fall into two categories, those in which the electron density in a decaying plasma or afterglow is examined as a function of time and from which the mechanism and appropriate constants can be determined, and those concerned with the determination of recovery time in actual glow discharge tubes. Unfortunately, correlation between the two types of measurement has not often been established, and because

parameters such as geometry and potentials on subsidiary electrodes can materially affect deionisation, data useful to the valve engineer are limited.

5.4.2.1. *Microwave Studies of the Afterglow*

Examination of electron densities in the afterglow has been greatly facilitated over the last decade with the development of microwave techniques. The detection of microwaves by highly sensitive devices allows low amplitude signals to be used which have virtually no disturbing effect on the discharge under investigation. The technique, developed at the Massachusetts Institute of Technology, was first used by Biondi and Brown[35] in 1949. The discharge tube in the form of a short cylinder was inserted in a microwave cavity, and the resonant frequency, which was found to increase in the presence of a plasma, determined. A block diagram of the apparatus is shown in Fig. 5.11. The discharge is produced by a magnetron

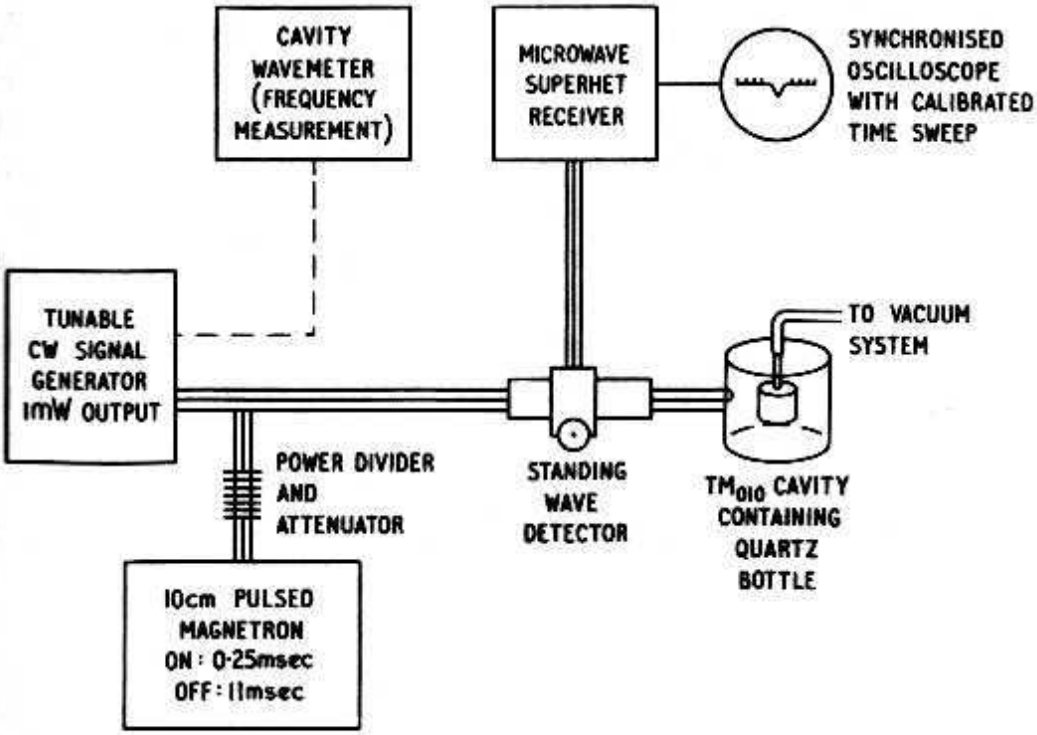


Fig. 5.11. Block diagram of the apparatus used by Biondi and Brown[35] for determining electron density in the afterglow. The oscilloscope sweep is synchronised to start as the magnetron is turned off

pulse during which time the electron density N_e, and also the ion density, builds up to a steady equilibrium state. At the end of the pulse, N_e decreases at a rate dependent on the deionisation process. During this period the gas is probed by a low amplitude c.w. signal

set at a chosen frequency and the reflected wave examined on an oscilloscope. Whilst the signal frequency differs from the plasma frequency ω_p given by

$$\omega_p^2 = \frac{4\pi N_e e^2}{m\varepsilon_o},$$

little absorption occurs, but when the electron density has decayed to a value where the plasma frequency is equal to the frequency of the signal, large absorption occurs and a negative pulse is observed on the reflected signal. By varying the frequency of the probe signal, the decay characteristic can be plotted. In fact, the frequency shift from the resonance frequency of the non-ionised gas is directly proportional to N_e.

Biondi and Brown[35] first used the method to examine helium plasmas. At pressures of 21·2 torr and 28·7 torr $1/N_e$ plotted against t gave a straight line, showing that recombination was predominant (cf. Equation 5.18). At 5·05 torr, however, diffusion was more important; Fig. 5.12 shows a typical experimental curve and the effect of correcting for recombination losses. They found $D_a p$ to be constant over the pressure range of 2–12 torr, although the absolute value of $D_a p$ varied slightly with different gas samples.

In a later paper, Biondi and Brown[36] extended their measurements to other gases at pressures sufficiently high to ensure that recombination was the dominant process. The recombination coefficients obtained are given in Table 5.3.

For the atomic gases α was independent of pressure at 300°K and temperature over the measured range, whereas for the diatomic gases α showed a variation with temperature and, in the case of N_2 and O_2, a marked dependence on pressure. The values of α for the atomic gases were several orders higher than theoretical

Table 5.3

Recombination coefficients at T = 300°K according to Biondi and Brown[36]

Gas	α (cm^3/sec^{-1})
Helium	$1{\cdot}7 \times 10^{-8}$
Neon	$2{\cdot}03 \times 10^{-7}$
Argon	3×10^{-7}
Hydrogen*	$2{\cdot}46 \times 10^{-6}$
Nitrogen	$1{\cdot}4 \times 10^{-6}$ at 2 torr to $1{\cdot}8 \times 10^{-6}$ at 9 torr
Oxygen	$2{\cdot}8 \times 10^{-7}$ at 2 torr to $1{\cdot}3 \times 10^{-6}$ at 18 torr

*Probably due to impurities.

predictions for radiative recombination, and higher than previous reported values. It was suggested[37] that formation of molecular ions and their subsequent neutralisation by dissociative recombination was taking place. Evidence in support of this was gained

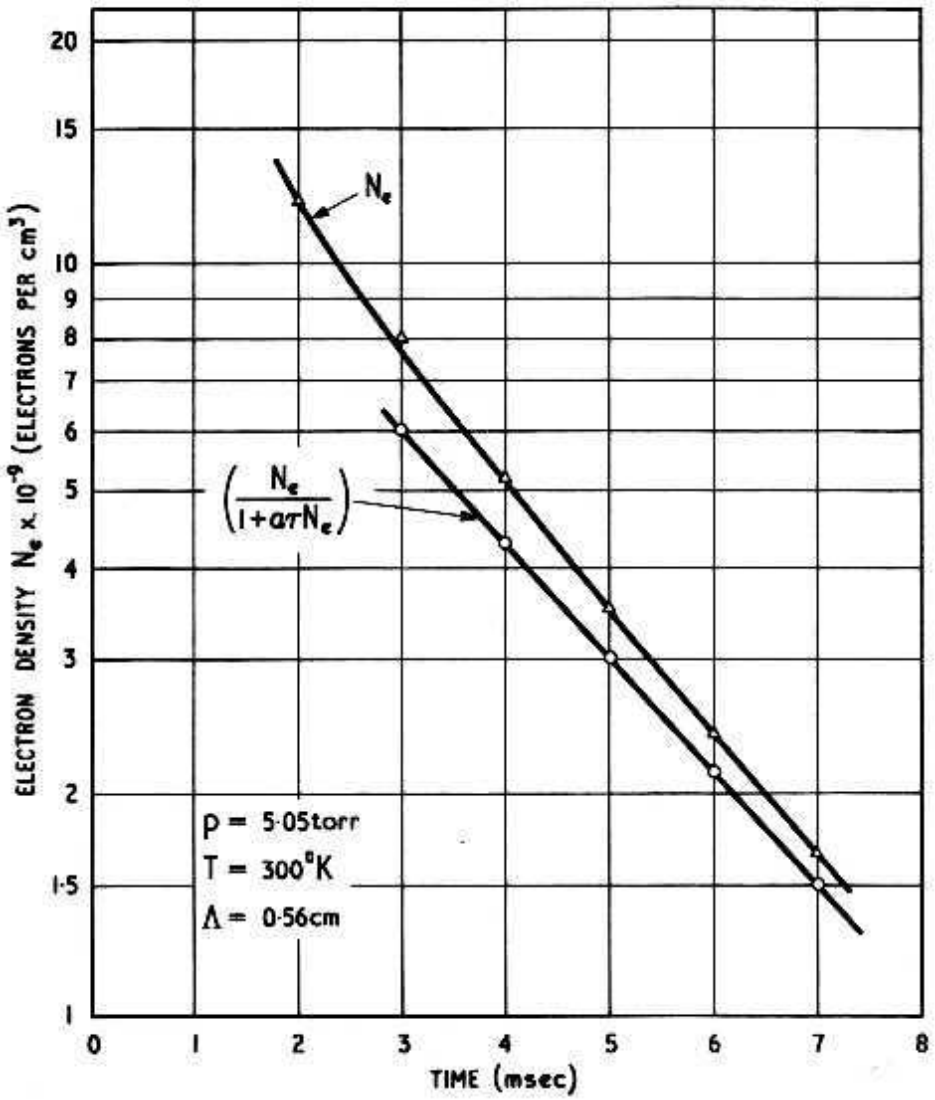


Fig. 5.12. Plot of the electron density against time in the afterglow period of a helium discharge at 5·05 torr. The lower curve is corrected for recombination losses. Biondi and Brown.[35] *α = recombination coefficient, τ = mean decay due to diffusion and Λ = a geometry factor knoen as the characteristic diffusion length*

from the lower values of α observed in helium-argon and neon-argon mixtures[38] where the main ion would be atomic argon and only radiative recombination could result. Dissociative recombination implies that in the pure atomic gases a considerable population of molecular ions must be present. This was confirmed for helium by Phelps and Brown[39] who attached a mass spectrometer to the discharge tube, and measured the density of molecular and atomic ions, as well as the electron density, in the afterglow. They found that as the pressure increased from 1–5 torr the dominant ion changed from He^+ to He_2^+.

The effect of metastable atoms on deionisation has been observed by Biondi[40] for helium and neon. He found that after the maintaining field was removed, electrons continued to be produced due to metastable-metastable collisions. Typical curves for helium are shown in Fig. 5.13. From studies of the ionisation produced by

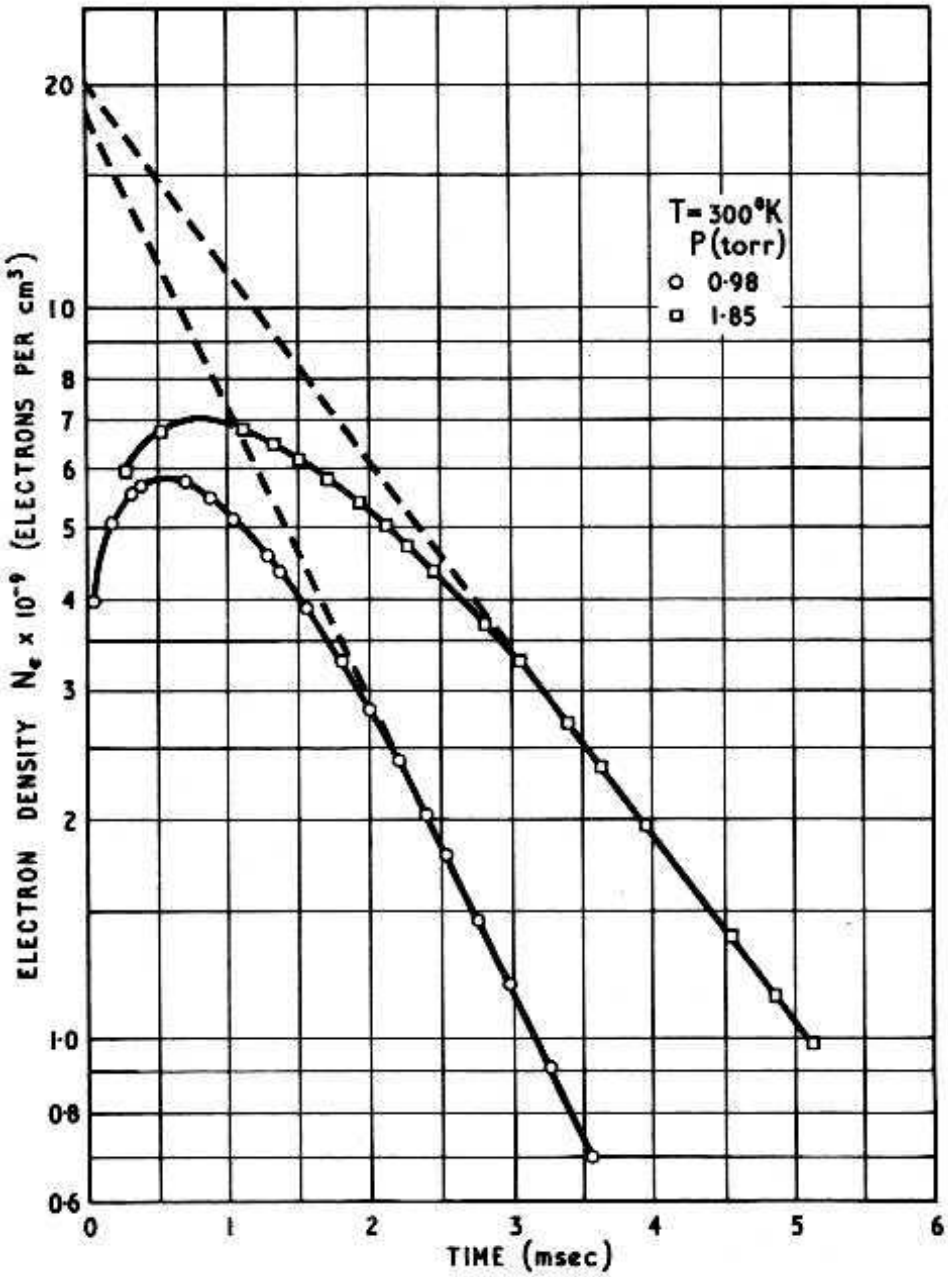


Fig. 5.13. The electron density in the afterglow of a helium discharge as a function of time, showing the production of electrons after switching off the discharge. Biondi[40]

metastables in the afterglow Biondi[41] determined ionisation cross sections for Penning mixtures, see Chapter 1, of helium-argon, helium-mercury vapour and neon-argon.

A more comprehensive study of inert gases has been made by Oskam;[34] in particular he investigated the influence of small admixtures of atoms having a lower ionisation potential than the main gas. For pure helium at pressures of 5 and 22·8 torr he found

that the ions were mainly molecular and that deionisation took place by ambipolar diffusion, in contrast to the observations of Biondi and Brown. Introducing small quantities of neon had a large effect on the behaviour of the afterglow, the rate of electron loss increasing with increasing neon concentration. This is shown in Figs. 5.14 and 5.15 giving results at 5 torr and 22·8 torr respectively. At low pressure the increase in electron loss was due to an

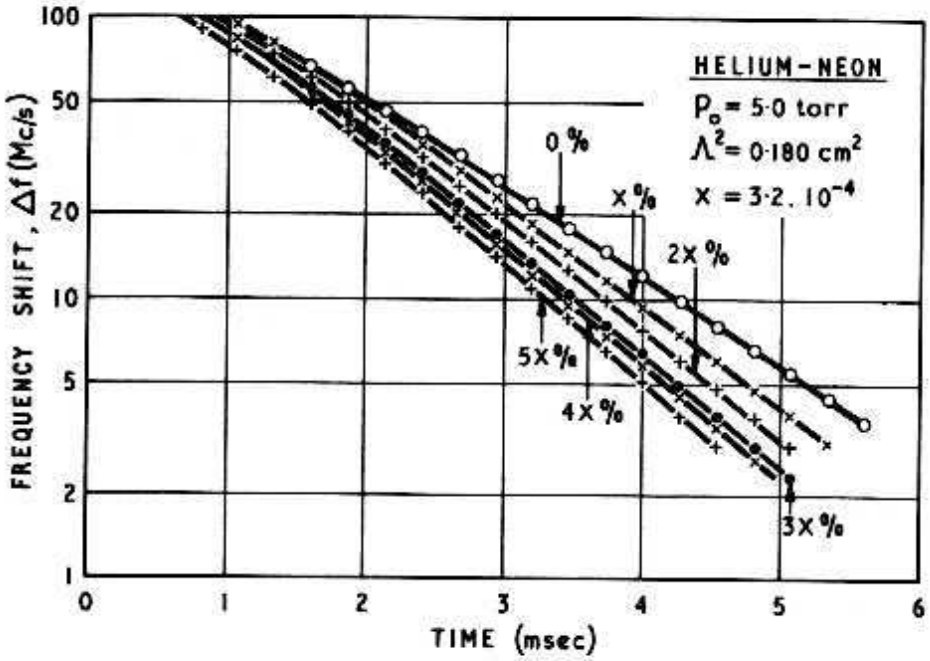


Fig. 5.14. The electron density in the afterglow of discharges in helium and helium-neon mixtures, expressed as the frequency shift, as a function of time. Filling pressure was 5 torr. Oskam[34]

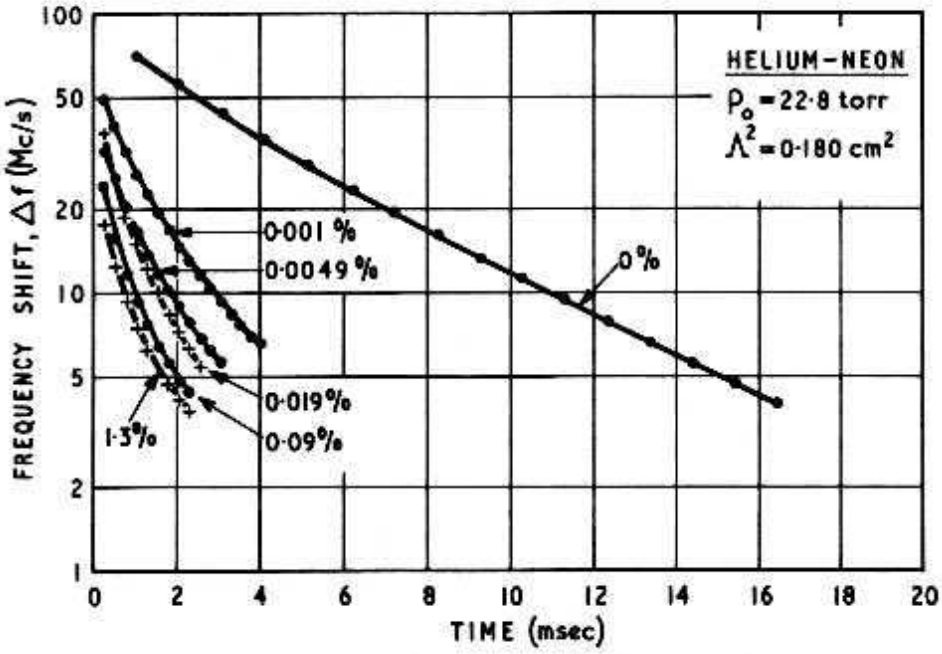


Fig. 5.15. The electron density in the afterglow of discharges in helium and helium-neon mixtures, expressed as the frequency shift, as a function of time. Filling pressure was 22·8 torr. Oskam[34]

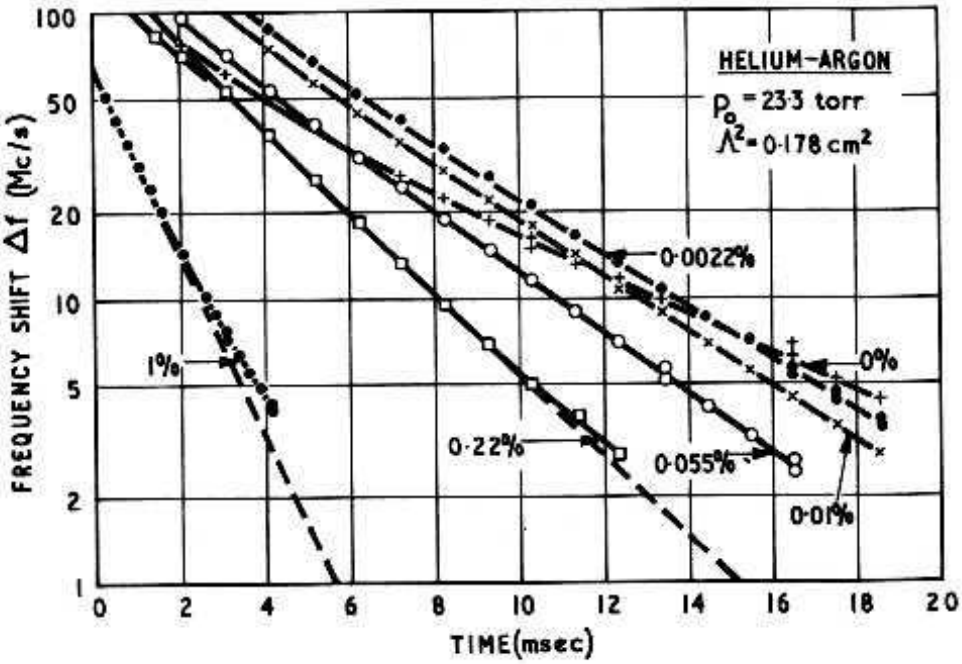


Fig. 5.16. The electron density in the afterglow of discharges in helium and helium-argon mixtures, expressed as the frequency shift, as a function of time. Filling pressure was 23·3 torr. Oskam[34]

increase in ambipolar diffusion, and was explained by a replacement of the helium molecular ions by neon atomic ions. Since neon cannot be ionised by the helium metastable a charge-transfer process was assumed:

$$He_2^+ + Ne \rightarrow Ne^+ + 2He$$

At the higher pressure and larger neon concentrations recombination was evident and involved the production of molecular ions. The recombination coefficient however did not correspond to He_2^+ or Ne_2^+ and Oskam postulated a mixed-gas ion $(He\,Ne)^+$ by the three-body reaction:

$$Ne^+ + Ne + He \rightarrow (Ne\,He)^+ + Ne$$

Support for the above two reactions has been obtained from spectrographs of neon-helium discharges;[42, 43] when small concentrations of neon are added to helium the helium bands due to He_2^+ disappear, but as the concentration increases a new band structure is observed.

The effect of adding argon or krypton to helium was similar to that of adding neon in reducing the deionisation time. The effect however was less marked, Figs. 5.16 and 5.17. For these mixtures the Penning reaction was the main effect.

For pure neon, recombination was an important process even at pressures as low as 3·5 torr. On adding small concentrations of

argon or krypton, the recombination phenomenon disappeared completely and deionisation was due entirely to ambipolar diffusion, giving a slower rate of decay. However, for larger concentrations the recombination loss process reappeared; this is shown in Fig. 5.18 for neon-argon at 19·8 torr. In general, however, whereas admixed gases in helium reduce the deionisation time, similar admixed gases in neon increase it. The significance of Oskam's results on the recovery time in cold-cathode glow discharge tubes

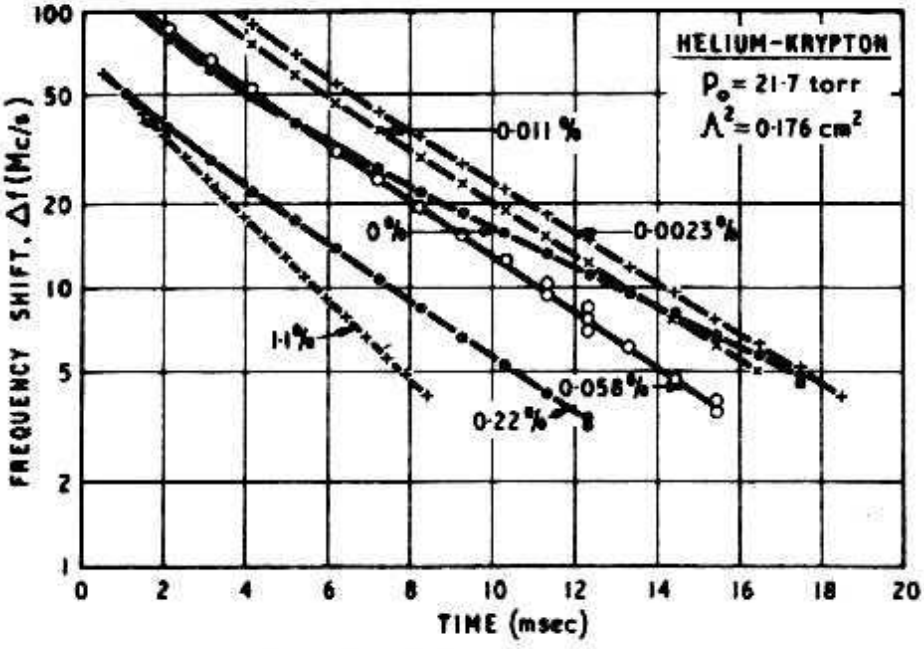


Fig. 5.17. The electron density in the afterglow of discharges in helium and helium-krypton mixtures, expressed as the frequency shift, as a function of time. Filling pressure was 21·7 torr. Oskam[34]

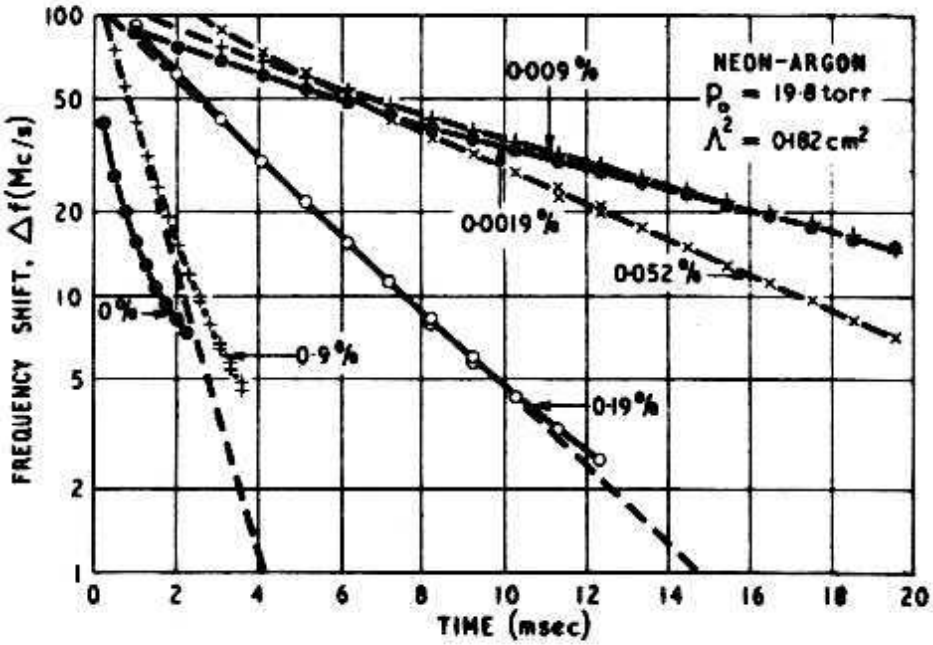


Fig. 5.18. The electron density in the afterglow of discharges in neon and neon-argon mixtures, expressed as the frequency shift, as a function of time. Fillin pressure was 19·8 torr. Oskam[34]

has been investigated by Ahsmann[44] and is reported in the next section. A table of recombination coefficients and mobilities derived by Oskam is given in Table 5.4.

Table 5.4

Mobilities and recombination coefficients from microwave measurements in inert gases according to Oskam[34].

Main Gas	*Admixed Gas*	*Mobility* ($cm^2\ V^{-1}\ sec^{-1}$) (*reduced to* $T = 273°K$)	*Recombination coeff.* $cm^3\ sec^{-1} \times 10^{-7}$
Helium	—	$\mu_0(He_2^+) = 17{\cdot}3 \pm 0{\cdot}7$	—
Helium	Neon	$\mu_0(Ne^+) = 24{\cdot}0,\ 24{\cdot}8$	$a(He\ Ne)^+ = 4{\cdot}1,\ 4{\cdot}4$
Helium	Argon	$\mu_0(Ar^+) = 23{\cdot}2,\ 23{\cdot}4$	—
Helium	Krypton	$\mu_0(Kr^+) = 21{\cdot}0,\ 22{\cdot}3$	—
Neon	—		$a\ (Ne_2^+) = 2{\cdot}3,\ 2{\cdot}6$
Neon	Argon	$\mu_0(Ar^+) = 6{\cdot}9,\ 7{\cdot}9$	
Neon	Krypton	$\mu_0(Kr^+) = 6{\cdot}0,\ 7{\cdot}0$	

The two sets of figures represent separate measurements or calculations, in some cases involving results at different pressures.

Four years later Mulcahy and Lennon[45] reported on their afterglow experiments with helium, neon and argon at pressures low enough to give appreciable ambipolar diffusion. In helium they found no evidence of recombination over the pressure range 0·8–22 torr. However, expressing the diffusion constant as $D_a p_o$ (p_o = pressure reduced to °C) to overcome the variation of D_a with p, they found that $D_a p_o$ increased from 540 cm^2 torr sec^{-1} at 0·85 torr to 780 cm^2 torr sec^{-1} over the range of 3–22 torr (i.e. μ increased from $\sim 14{\cdot}2\ cm^2\ V^{-1}\ sec^{-1}$ to $\sim 20{\cdot}3\ cm^2\ V^{-1}\ sec^{-1}$). A similar increase was observed in neon and argon, 100–260 cm^2 torr sec^{-1} ($\mu \sim 2{\cdot}8\ cm^2\ V^{-1}\ sec^{-1}$– $\sim 7{\cdot}0\ cm^2\ V^{-1}\ sec^{-1}$) and 40–90 cm^2 torr sec^{-1} ($\mu \sim 1{\cdot}1$– $\sim 2{\cdot}3\ cm^2\ V^{-1}\ sec^{-1}$) respectively. However in these gases recombination was important at pressures of the order of 3 or 4 torr. They ascribed the increase as due to the change of ion species from atomic to molecular as the pressure increased.

The deionisation in xenon and krypton with impurity of up to 0·5% of the other gas has been investigated by Lennon and Sexton.[46] Above 3 torr ambipolar diffusion was negligible. The recombination coefficient for xenon which was unaffected by the krypton impurity was $2{\cdot}3 \times 10^{-6}\ cm^3\ sec^{-1}$. The value for krypton, however, was affected by the xenon impurity and an upper limit of $1{\cdot}1 \times 10^{-6}\ cm^3\ sec^{-1}$ was given.

The afterglow measurements in hydrogen had given rather conflicting results,[36, 47] until Persson and Brown[48] in 1955 demonstrated the importance of the purity of the hydrogen; an impurity level of 1 part in 10^8 could cause attachment losses of electrons which would mask other deionisation processes. Taking this into account, they showed that ambipolar diffusion accounted for the main loss process giving $D_a p = 700\,\text{cm}^2\,\text{torr}\,\text{sec}^{-1}$, and suggested that α was less than $3 \times 10^{-8}\,\text{cm}^3\,\text{sec}^{-1}$. Whitmer,[49] using a different technique which allowed higher electron densities, obtained a value of approximately $5{\cdot}9 \times 10^{-11}\,\text{cm}^3\,\text{sec}^{-1}$ for the recombination coefficient. Mixtures of a small percentage of hydrogen (10^{-4} to 1%) in the inert gases, helium, neon and argon, have been examined by Oskam and Mittelstadt.[50] They found that in the early part of the afterglow period the decay was neither exponential nor reciprocal in time, and must, therefore, involve more than one type of ion. The large observed rate of change of electron density during this period suggested that at least one of the ion species had a recombination coefficient of the order of $10^{-6}\,\text{cm}^3\,\text{sec}^{-1}$. They did not believe this was due to contamination. During the later part of the afterglow period the time dependence was found to be exponential showing a diffusion process even up to pressures of 25 torr. The effective mobilities deduced from the measured ambipolar diffusion coefficients were $38 \pm 4\,\text{cm}^2\,\text{V}^{-1}\,\text{sec}^{-1}$ (the highest mobility yet recorded) for helium-hydrogen, $17 \pm 1{\cdot}0\,\text{cm}^2\,\text{V}^{-1}\,\text{sec}^{-1}$ for neon-hydrogen, and $6{\cdot}3 \pm 0{\cdot}5\,\text{cm}^2\,\text{V}^{-1}\,\text{sec}^{-1}$ for argon-hydrogen. Compared with the mobilities for the inert gases alone the values were considerably higher, and suggested that the diffusion was determined either by hydrogen ions or mixed gas ions i.e. $(\text{HeH})^+$, $(\text{NeH})^+$ and $(\text{ArH})^+$ respectively. Admixtures of hydrogen in helium or neon constitute a Penning mixture, and thus hydrogen ions would be expected to be formed by collisions of the second kind, involving the metastables of the relevant parent gas. In argon-hydrogen such a reaction is impossible, and processes similar to those described for helium-neon must be invoked.

5.4.2.2. *Recovery-time Measurements in Glow Discharge Tubes*

Recovery time can be determined by applying a short duration probe pulse to the electrodes of the discharge tube, at a time t which can be varied, after the extinction of the main discharge. By varying the amplitude of the pulse until re-ignition occurs, a plot of pulse amplitude for breakdown against t gives the recovery

characteristics. Alternatively an extinguishing pulse can be applied to a d.c. discharge, and the duration required to prevent re-ignition measured as a function of the applied voltage. Since diffusion plays a prominent role at low pressures, geometry and the residual field applied to the gap during deionisation may well affect the measured recovery time. Thus results obtained from the different methods cannot always be compared, and this must be borne in mind when considering published data.

Measurements using the probe pulse have been carried out by the author and Fig. 5.19 shows examples of a plot of probe strike-

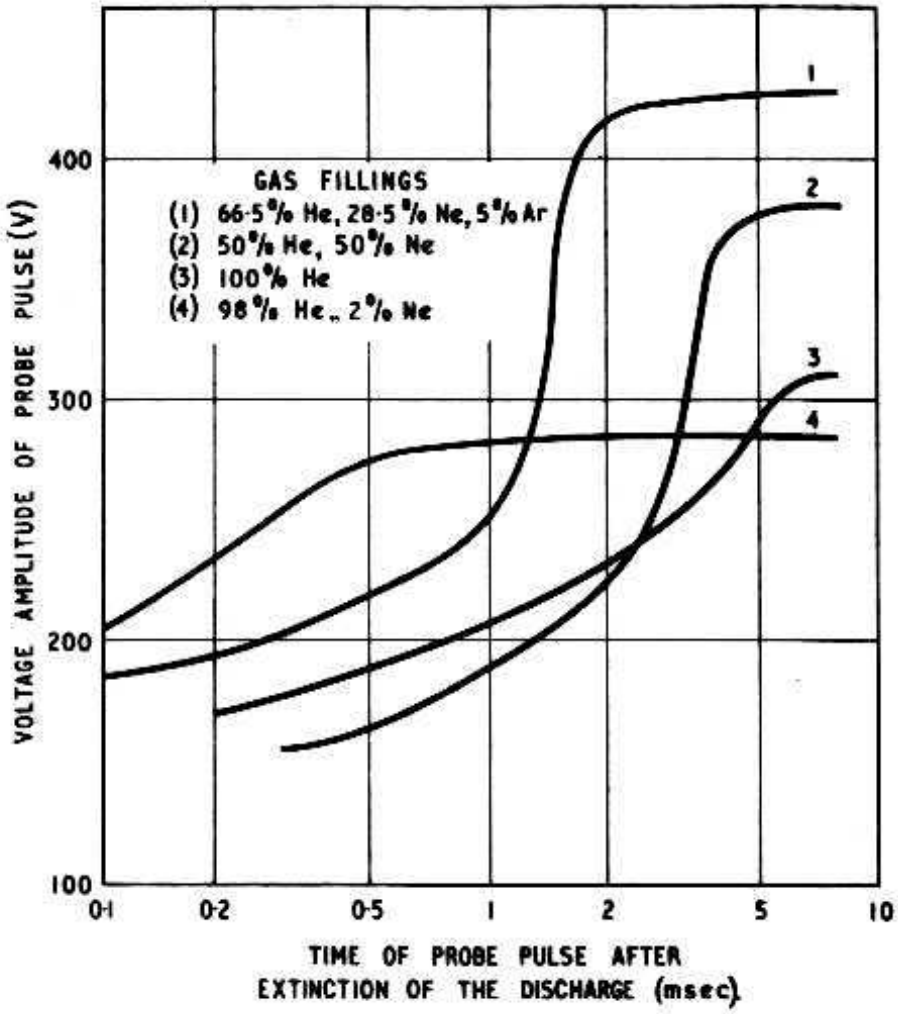


Fig. 5.19. Recovery time curves for experimental tubes filled with various inert gas mixtures at 40 torr pressure. (Breakdown potential of probe pulse plotted against time from extinction of the discharge)

voltage against the logarithm of time which are characteristic of recovery curves in cold cathode tubes. The final saturation value represents the breakdown potential of the tube, whilst the initial value represents the maintaining potential. It is seen that 98% He+2% Ne has the shortest recovery time and pure helium the longest. Since the diffusion and recombination processes are pressure dependent, the recovery time will also depend on pressure;

in Fig. 5.20 a plot of recovery time t_r, defined as the time at which a pulse of amplitude $= V_m + (V_s - V_m)/2$ just restrikes the tube, against pressure is shown for an experimental tube containing neon with molybdenum electrodes. Over the pressure range investigated recombination is probably the dominant process. The results were taken at 1 mA main current, and the effect of varying this current is shown in Fig. 5.21. In the author's experiments no voltage was

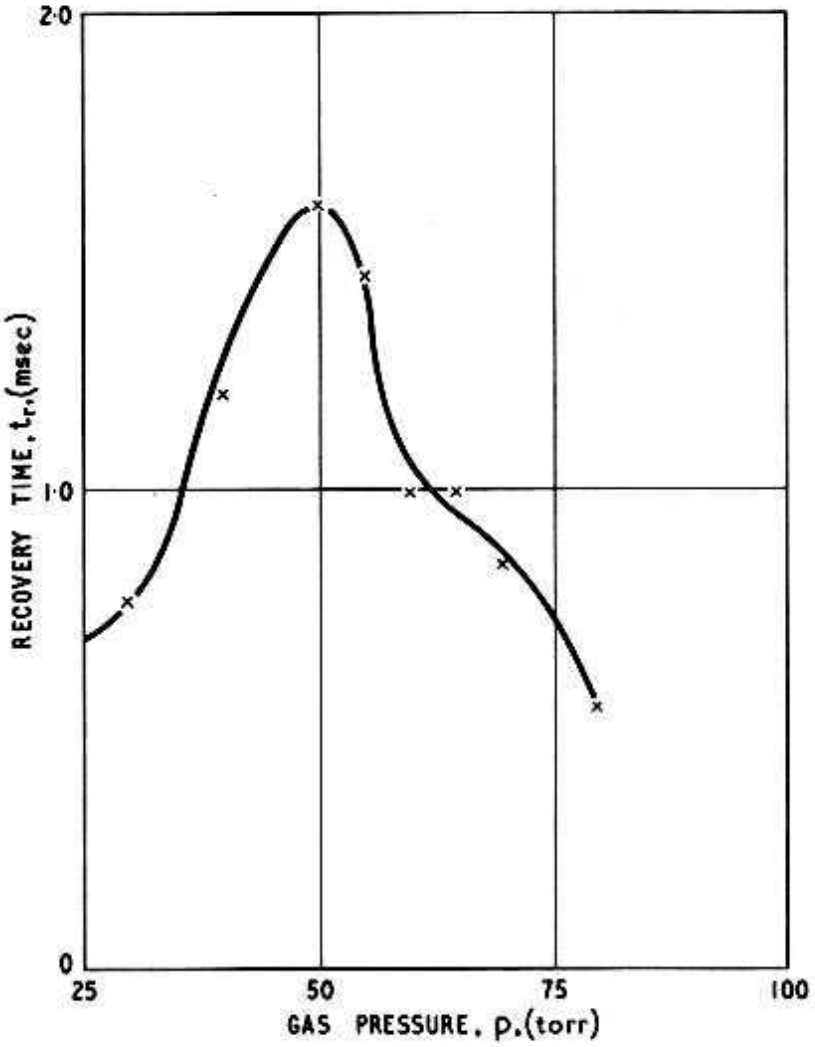


Fig. 5.20. Recovery time t_r, defined as the maximum time after extinction of a discharge at which a pulse of amplitude $V_m + \frac{1}{2}(V_s - V_m)$ restrikes the tube, as a function of pressure for a neon filled tube

applied to the gap during the deionisation period. The importance of the applied potential during this period, however, is seen in Fig. 5.22 where recovery time, defined as the length of an extinction pulse required to just prevent a tube restriking, is plotted against residual voltage, as a function of the applied voltage. The results were obtained by van Lottum[51] using a tube filled to 15 torr of argon and with a BaO-coated cathode; V_s was 200 V and V_m was 60 V. As the residual potential increases from zero, the recovery time decreases due to the field sweeping the ions and electrons from

the gap. At higher voltages, approaching the maintaining potential, the field is sufficiently high to cause ionisation and the recovery time is consequently increased. Fig. 5.22 is a convenient way of presenting results using the extinction pulse method, and was used by Ahsmann[44] to show the effect of admixtures of neon in helium

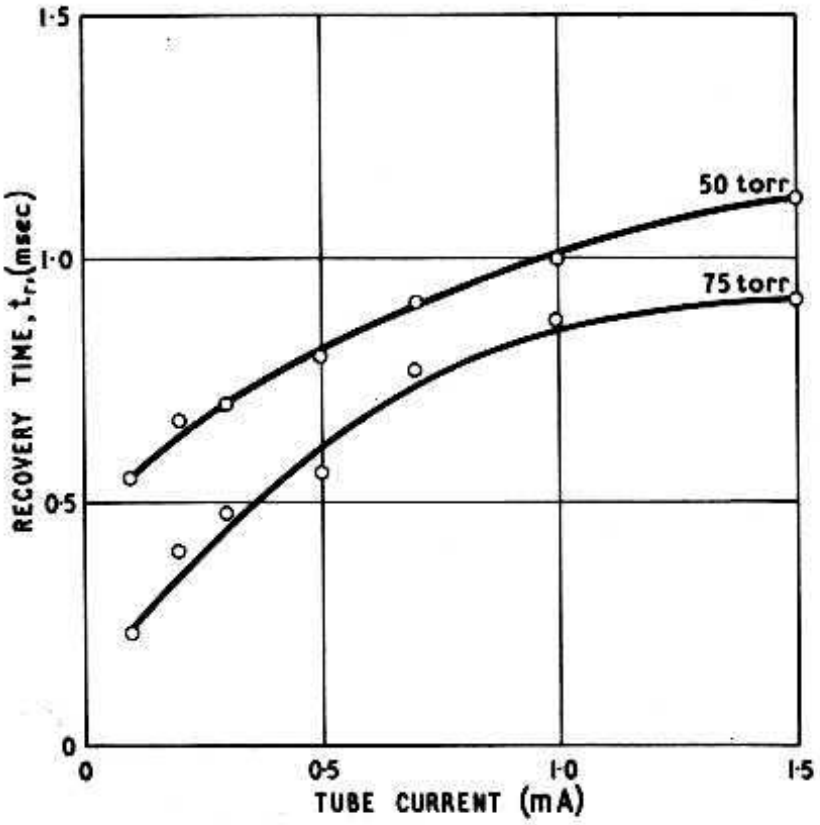


Fig. 5.21. Recovery time t_r, defined as the maximum time after extinction of a discharge at which a pulse of amplitude $V_m+\frac{1}{2}(V_s-V_m)$ restrikes the tube, as a function of the discharge current for a neon filled tube

and of krypton in argon. The energy levels of the gases in each of the two mixtures are such that similar reactions should be expected. In the case of helium, where deionisation at the pressure used is by ambipolar diffusion of He_2^+, the charge-exchange reaction and production of $(He\,Ne)^+$ with their subsequent loss by dissociative recombination on adding a small percentage of neon results in a considerable reduction in the recovery time, Fig. 5.23. On the other hand, in argon, where deionisation is by recombination, the charge exchange on adding a small percentage of krypton increases the recovery time. Presumably at this stage Kr^+ ions are formed and ambipolar diffusion is the predominant process. With a much larger percentage of krypton, molecular ions are formed and the recovery time is reduced again, Figs. 5.24 and 5.25.

Yablonskiy[52] has measured recovery times in all the inert gases using the probe pulse technique, and with diodes having hemi-

spherical molybdenum cathodes and nickel plate-anodes. He found that in Ne, Ar, Kr, and Xe the recovery rate defined as the slope of the voltage-time recovery curve (see Fig. 5.19) increased with increasing pressure over the measured range (20–150 torr for Ne and Ar, and 10–20 torr for Kr and Xe). For helium the recovery rate decreased with increase of pressure from 20–40 torr, but increased for a pressure rise from 70–150 torr. Table 5.5 gives comparative values at 20 torr pressures.

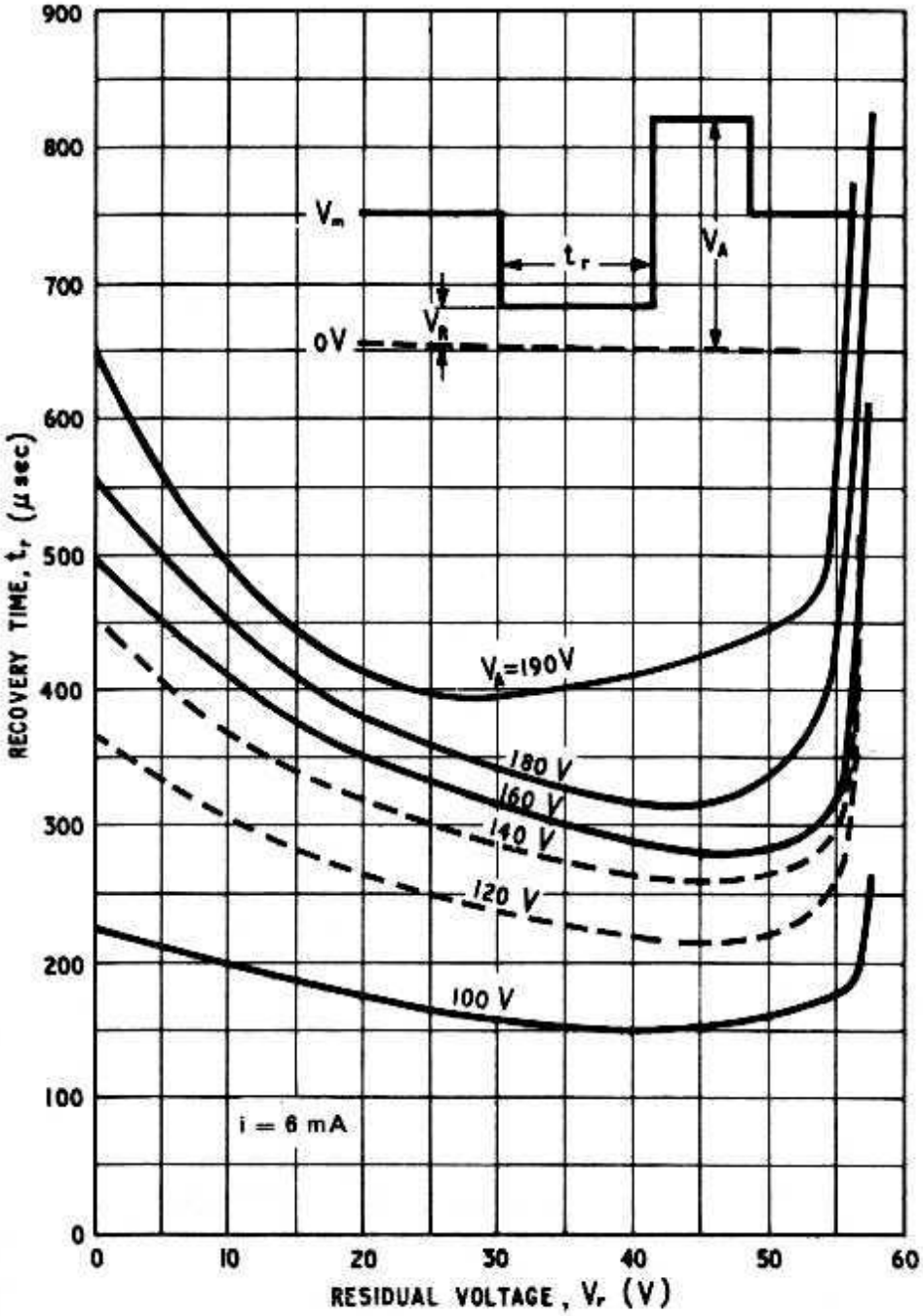


Fig. 5.22. Recovery time t_r, defined as the maximum duration of the extinguishing pulse for the tube to restrike, as a function of the residual voltage for an argon filled tube. Voltage wave form inset. Van Lottum[51]

Table 5.5

Recovery times for the inert gases according to Yablonskiy[52]

Gas	*He*	*Ne*	*Ar*	*Kr*	*Xe*
Initial recovery rate (V μ sec^{-1})	0·26	0·11	0·4	0·85	1·6
Time to recover to 90% of unprimed breakdown value, μ sec	300	400	200	150	75

Table 5.6

$t_r.p/d$ according to Acton and Swift[2]

Gas	$t_r.p/d$(*msec torr cm*$^{-1}$)
Helium	70
Neon	30
Argon	10
Hydrogen	2·0
Neon-argon (99·5% : 0·5%)	120
Helium-argon (99·5% : 0·5%)	25
Helium-neon (95% : 5%)	10
Helium-hydrogen (85% : 15%)	0·6
Neon-argon-hydrogen (95·5% : 0·5% : 4%)	3·0

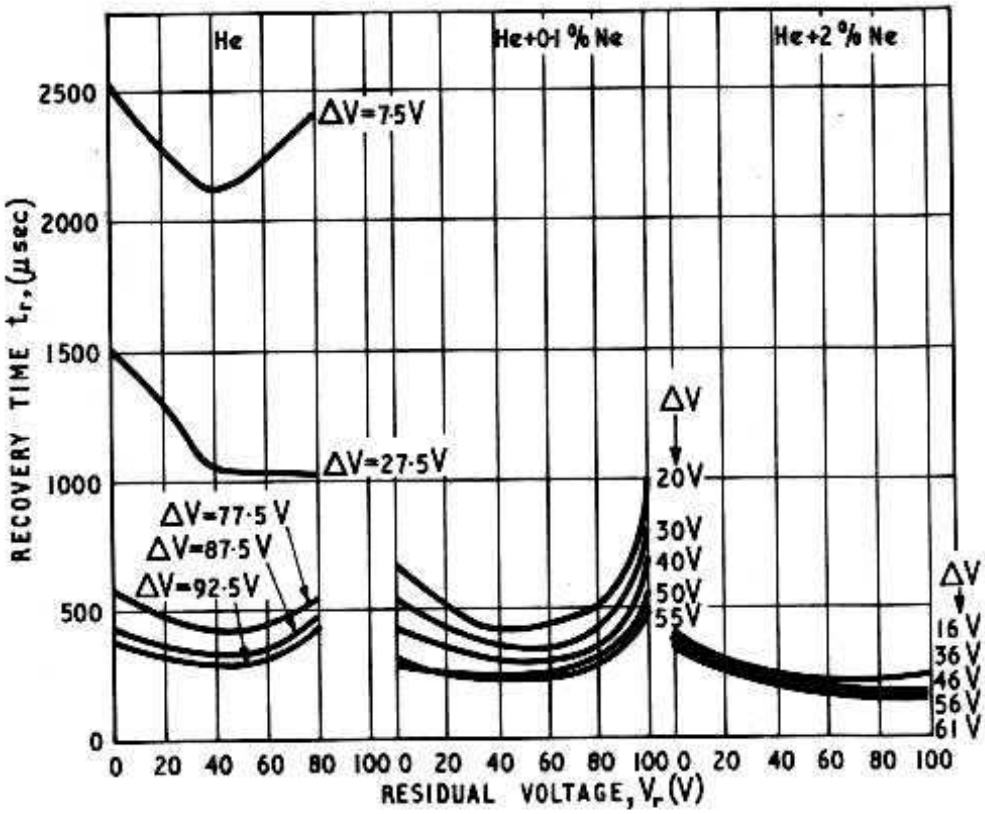


Fig. 5.23. Recovery time t_r, defined as the maximum duration of the extinguishing pulse for the tube to restrike, as a function of the residual voltage, for helium, helium+ 0·1% neon, and helium with 2% neon. Ahsmann[44]. $\Delta V = V_A - V_m$ see inset fig. 5.22

Acton and Swift[2] gave a table of values for $t_r.p/d$ which they found to be roughly constant for various gases; t_r was defined as the time taken after extinction for the re-ignition voltage to have risen to halfway between V_m and V_s. During deionisation the voltage was at $V_m/2$. The table is reproduced as Table 5.6.

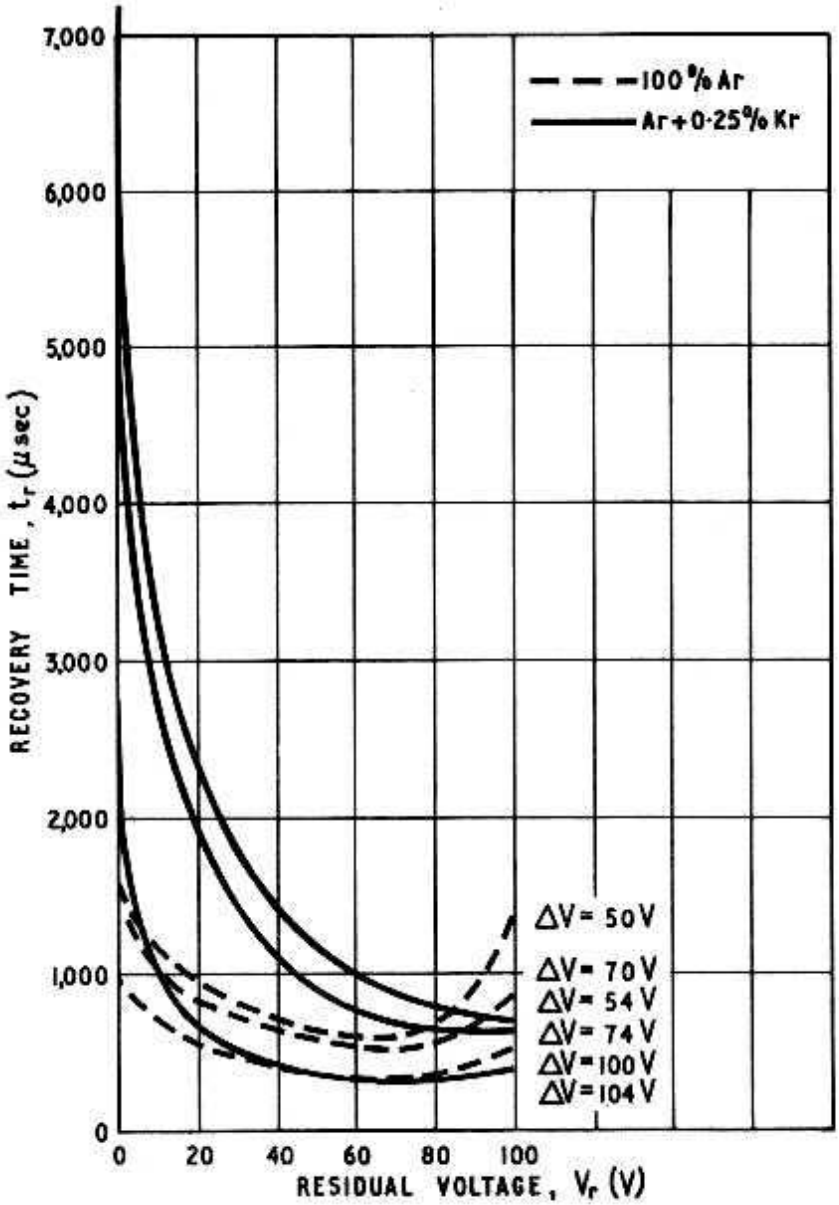


Fig. 5.24. Recovery time t_r, defined as the maximum duration of the extinguishing pulse for the tube to restrike, as a function of the residual voltage, for argon and argon+ 0·25% krypton. Ahsmann[44]. $\Delta V = V_A - V_m$ see inset fig. 5.22

Of interest is the low value for hydrogen and hydrogen-inert gas mixtures. It is well known that adding small concentrations of hydrogen to the inert gases[53, 54] considerably increases the switching speed of cold cathode tubes due to a reduction of the recovery time. What is not clear, however, is the mechanism involved. It has been argued that the Penning reaction between the inert gas and hydrogen reduces the number of metastable atoms, and this

reduces the deionisation time. If this is true, then the metastables must play a much larger part in the recovery time of glow discharges than is evident in microwave studies. It also conflicts with the observation that neon-argon which would be expected to have a lower concentration of metastables than pure neon, gives a longer deionisation time. Nor do the results of the microwave study of

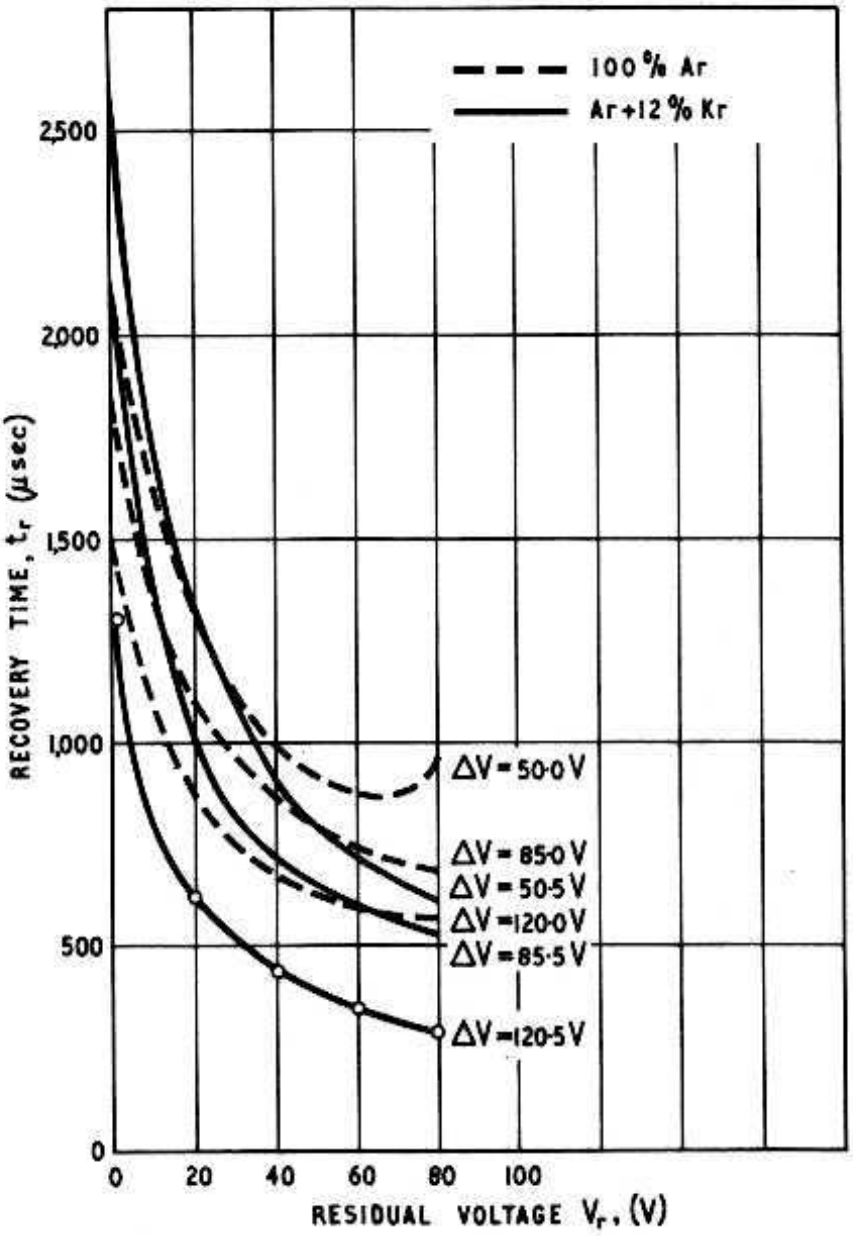


Fig. 5.25. Recovery time t_r, defined as the maximum duration of the extinguishing pulse for the tube to restrike, as a function of the residual voltage, for argon and argon+12% krypton. Ahsmann[44]. $\Delta V = V_A - V_m$ see inset fig. 5.22

inert gas-hydrogen mixtures by Oskam and Mittelstadt give a completely satisfactory answer, since the increase in the ambipolar diffusion coefficient still results in a value for neon plus admixed hydrogen lower than for helium plus admixed neon, whereas the deionisation time is faster. It is possible that the high recombination coefficient observed in the early part of the afterglow period is the

controlling factor. If this is so, then for the observed fast deionisation time of hydrogen one must also assume that recombination is the predominant process, in which case for a high value of α one must assume that in practical devices the gas is always contaminated.

REFERENCES

1. VON LAUE, M., *Ann. Phys.*, **76**, 261, (1925)
2. ACTON, J. R. and SWIFT, J. D., *Cold Cathode Discharge Tubes*, Heywood, London, (1963)
3. BRAUNBECK, W., *Z. Physik*, **107**, 180, (1937)
4. HERTZ, G., *Z. Physik*, **106**, 102, (1937)
5. WIJSMAN, R. A., *Phys. Rev.*, **75**, 833, (1949)
6. FULLER, D. W. E., *Rad. Electron. Engr.*, **28**, 229, (1964)
7. LLEWELLYN JONES, F., *Rad. Electron. Engr.*, **28**, 219, (1964)
8. ZUBER, K., *Ann. Phys.*, **76**, 231, (1925)
9. STRIGEL, R., *Wiss. Veröff. Siemens-Konz*, **11**, 52, (1932)
10. TILLES, A., *Phys. Rev.*, **46**, 1015, (1934)
11. LLEWELLYN JONES, F. and DE LA PERRELLE, E. T., *Proc. Roy. Soc.*, **A216**, 267, (1953)
12. LLEWELLYN JONES, F. and MORGAN, C. G., *Proc. Roy. Soc.*, **A218**, 88, (1953)
13. FOX, A. J., *Proc. Intern. Conf. Ionisation Phenomena Gases, 3rd, Venice*, Societa Italiana di Fisica, 325, (1957)
14. FOX, A. J., *Internal report of the Mullard Research Lab.*, **No. VPL 59**, (1954)
15. DRUYVESTEYN, M. J. and PENNING, F. M., *Rev. Mod. Phys.*, **12**, 87, (1940)
16. SCHADE, R., *Z. Physik*, **104**, 487, (1937); **108**, 353, (1938)
17. STEENBECK, M., *Wiss. Veröff. Siemens-Konz*, **9**, 42, (1930)
18. BARTHOLOMYCZYK, W., *Z. Physik*, **116**, 235, (1940)
19. RAETHER, H., *Erge. Exakt. Natur.*, **22**, 73, (1949)
20. MEEK, J. M. and CRAGGS, J. D., *Electrical Breakdown of Gases*, Clarendon Press, Oxford, p. 251, (1953)
21. FLETCHER, R. C., *Phys. Rev.*, **76**, 1501, (1949)
22. FISHER, L. H. and BEDERSON, B., *Phys. Rev.*, **81**, 109, (1951)
23. DUTTON, J., HAYDON, S. C. and LLEWELLYN JONES, F., with mathematical appendix by DAVIDSON, P. M., *Brit. J. Appl. Phys.*, **4**, 170, (1953)
24. DAVIDSON, P. M., *Proc. Phys. Soc.*, **80**, 143, (1962)
25. WARD, A. L., *Proc. Intern. Conf. Ionisation Phenomena Gases, 5th, Munich*, North Holland Publishing Co. Amsterdam, **Vol. 2**, 1595, (1962)
26. DAVIES, A. J., EVANS, C. J. and LLEWELLYN JONES, F., *Proc. Roy. Soc.*, **A281**, 164, (1964)
27. KACHICKAS, G. A. and FISHER, L. H., *Phys. Rev.*, **88**, 878, (1952)
28. KACHICKAS, G. A. and FISHER, L. H., *Phys. Rev.*, **91**, 775, (1953)
29. MORGAN, C. G., *Phys. Rev.*, **104**, 566, (1956)
30. LLEWELLYN JONES, F. and JONES, E., *Proc. Phys. Soc.*, **75**, 762, (1960)
31. DAVIES, D. K., LLEWELLYN JONES, F. and MORGAN, C. G., *Proc. Phys. Soc.*, **81**, 677, (1963)
32. MENES, M., *Phys. Rev.*, **116**, 481, (1959)
33. MORSE, P. M. and FESHBACK, H., *Methods of theoretical Physics*, McGraw-Hill, New York, (1953)
34. OSKAM, H. J., *Philips Res. Rep.*, **13**, 335, (1958)
35. BIONDI, M. A. and BROWN, S. C., *Phys. Rev.*, **75**, 1700, (1949)

36. BIONDI, M. A. and BROWN, S. C., *Phys. Rev.*, **76,** 1697, (1949).
37. BATES. D. R., *Phys. Rev.*, **77,** 718, (1950); **78,** 492, (1950)
38. BIONDI, M. A., *Phys. Rev.*, **83,** 1078, (1951)
39. PHELPS, A. V. and BROWN, S. C., *Phys. Rev.*, **86,** 102, (1952)
40. BIONDI, M. A., *Phys. Rev.*, **82,** 453, (1951)
41. BIONDI, M. A., *Phys. Rev.*, **88,** 660, (1952)
42. OSKAM, H. J. and JONGERIUS, H. M., *Physica*, **24,** 1092, (1958)
43. WESTON, . G. F., *Brit. J. App. Phys.*, **10,** 523, (1959)
44. AHSMANN, G. J. M., *Proc. Intern. Conf. Ionisation Phenomena Gases, 4th, Uppsala*, North Holland Publishing Co., Amsterdam, **Vol. I,** 309, (1960)
45. MULCAHY, M. J. and LENNON, J. J., *Proc. Phys. Soc.*, **80,** 626, (1962)
46. LENNON, J. J. and SEXTON, M. C., *J. Electron. Control.* **7,** 123, (1959)
47. VARNERIN, L. J., *Phys. Rev.*, **84,** 563, (1951)
48' PERSSON, K. B. and BROWN, S. C., *Phys. Rev.*, **100,** 729, (1955)
49. WHITMER, R. F., *Phys. Rev.*, **104,** 572, (1956)
50. OSKAM, H. J. and MITTELSTADT, V. R., *J. Appl. Phys.*, **36,** 3740, (1965)
51. VAN LOTTUM, J. TH. A., *Philips Natuurkundig Lab. Internal Report*, **No. 2753,** (1952)
52. YABLONSKIY, F. M., *Radio Engng. Electron. Phys.* (USSR), **7,** 125, (1962)
53. HOUGH, G. H. and RIDLER, D. S., *Electron. Engng.*, **24,** 152, (1952)
54. BECK, A. H. and JACKSON, T. M., *Elec. Commun.*, **35,** 251, (1958–9)

CHAPTER 6

Stabilisers and Reference Tubes

6.1. General Criteria

The characteristics of a glow discharge detailed in Chapter 3, in particular the small variation of potential over a wide range of current, commends the use of the cold cathode tube as a constant voltage-level supply. It may either be used to provide a reference voltage for thermionic valve stabilising circuits, or be employed for direct stabilisation by connecting in parallel with the load; hence the terms reference tube and stabiliser. For direct stabilisation essentially simple diode glow-discharge tubes are employed, connected in series with a limiting resistor R_a, across the voltage supply V_a, as shown in Fig. 6.1. Fluctuations in the supply voltage

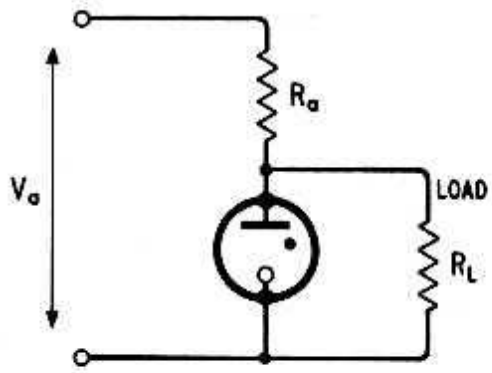


Fig. 6.1. Basic circuit for parallel stabilisation

are taken up by the resistor, whilst the changes in current resulting from variations in the load impedance are absorbed in the tube.

Apart from needing good regulation over a wide range of current, if a discharge tube is to be used for stabilisation it must itself be inherently stable during life and with changes in ambient temperature. It must also be free from discontinuities or hysteresis in the current-maintaining potential characteristic, have a low noise, and present a low impedance to a.c. ripple. Another desirable

characteristic is a small difference between the breakdown and maintaining potential, since the breakdown potential will appear across the load before the tube ignites, when the circuit is first switched on, and could cause excessive dissipation if it were too high. Also if the breakdown voltage is too high it may place a lower limit on the supply voltage.

The usefulness of a glow-discharge tube for stabilisation will, therefore, depend on the degree to which these desirable parameters can be built into a practical device. As has been shown in the previous chapters, some of these parameters are properties of the gas filling, electrode material, geometry, and processing schedule, and it is the exploitation of these properties in the design of stabilisers and reference tubes which is discussed in this chapter.

6.2. The Principle of Operation

6.2.1. PARALLEL VOLTAGE-STABILISATION

The voltage-current characteristics of a glow discharge have been detailed in Chapter 3. Although some curves were shown with substantially flat characteristics in the normal glow region (Fig. 3.19), in general there is a small increase in voltage with current over this region. This is partly a thermal effect resulting from the rise in cathode temperature as the current is increased. The temperature rise reduces the gas density in the cathode fall zone with a consequent increase in potential (see Section 3.4). Inhomogeneity of the cathode surface, and of the electric field can, however, also cause a rise in voltage with current. The latter is particularly important for the non-uniform fields which are present in the majority of stabilisers. Thus a stabiliser, although working in the normal glow region, has an internal impedance R_g associated with it, and V_m is given by

$$V_m = iR_g + V_o \tag{6.1}$$

Referring to Fig. 6.1, and assuming the load is resistive with an impedance R_L, then if the tube is to ignite, the load and series resistances must be proportioned to ensure that the breakdown potential V_s will be attained across the tube even when the supply voltage is at its minimum value

i.e.
$$V_{a(\min)} \frac{R_L}{R_a + R_L} > V_s$$

or in terms of the load current at breakdown $i_{L(V_s)}$

$$V_{a(\min)} > V_s + i_{L(V_s)} R_a \tag{6.2}$$

The further condition which must be satisfied in a stabiliser is that, once the tube has ignited, the current must lie within the specified range of the tube, i.e. between $i_{\min}$ and $i_{\max}$. Clearly the lowest current through the tube will occur when the load current is a maximum, and therefore

$$i_{\min} < \frac{V_{a(\min)} - V_m}{R_a} - i_{L(\max)} \tag{6.3}$$

where V_m is the maintaining potential.

Conversely, the highest current will occur when the load current is at a minimum, and therefore

$$i_{\max} > \frac{V_{a(\max)} - V_m}{R_a} - i_{L(\min)} \tag{6.4}$$

The three expressions, 6.2, 6.3 and 6.4 give the necessary information for designing a parallel stabilising circuit. From 6.3 and 6.4, if it is assumed that V_m does not vary greatly over the current range,

$$i_{\max} - i_{\min} > i_{L(\max)} - i_{L(\min)} \tag{6.5}$$

i.e. the working range of the stabiliser must exceed the load current variations. Whereas, combining 6.2 and 6.4 gives a single expression from which the parameters determining the circuit components can be deduced:

$$\frac{V_{a(\max)} - V_m}{i_{\max} + i_{L(\min)}} < R_a \quad \frac{V_{a(\min)} - V_s}{i_{L(V_s)}} \tag{6.6}$$

To cover the variations from tube to tube, V_m must be taken as the minimum value and V_s as the maximum value given by the manufacturer for the particular tube type. The way in which expression 6.6 can be used is illustrated in the following example. It is desired to provide a stabilised supply at a nominal 150 V for a load current range of 1–6 mA. From expression 6.5 it can be seen that the 150B2 with a current range of 5–15 mA should be suitable. For this tube $V_{s(\max)} = 180$ V and $V_{m(\min)} = 146$ V.

Substituting in expression 6.6 and taking $i_{L(V_s)}$ as $\frac{180}{146} \times 6$ mA i.e. taking R_L as minimum at breakdown, then

$$\frac{V_{a(\max)} - 146}{16 \times 10^{-3}} < R_a < \frac{V_{a(\min)} - 180}{7{\cdot}4 \times 10^{-3}}$$

allowing a $\pm 10\%$ variation on V_a, then for $V_a = 300$ V

$$11{\cdot}5\ k\Omega < R_a < 12{\cdot}2\ k\Omega$$

which would place an impractical tolerance on R_a unless the stability of V_a can be improved. At $V_a = 350$ V a wider tolerance is possible, $15\ k\Omega < R_a < 18{\cdot}3\ k\Omega$. Thus, increasing the supply voltage eases the tolerance on the series resistor, although of course its value will be higher. Using a tube with a larger current range (larger value of i_{max}) would also allow a wider tolerance resistor, but this would involve a higher power dissipation.

The degree of stabilisation that can be achieved can be deduced from considerations of the circuit of Fig. 6.1 and Equation 6.1. If the V-i characteristic is assumed to be linear, i.e. R_g constant in Equation 6.1, then

$$V_a = V_m\left[\frac{R_a}{R_g}+\frac{R_a}{R_L}+1\right]-\frac{R_aV_o}{R_g} \tag{6.7}$$

from which

$$\frac{dV_a}{dV_m} = \frac{R_a}{R_g}+\frac{R_a}{R_L}+1 \tag{6.8}$$

dV_a/dV_m is known as the stabilisation ratio and gives the factor by which variations in the supply voltage are reduced. Clearly for the most efficient stabilisation dV_a/dV_m should be as high as possible.

Expression 6.8 can be written in terms of V_a and V_m to eliminate R_L whence

$$\frac{dV_a}{dV_m} = \frac{R_gV_a+V_oR_a}{R_gV_m} \tag{6.9}$$

Thus not only do a high supply voltage and series resistor give greater tolerance on the operating criteria but they also result in better stabilisation. On the other hand, higher supply voltages mean a greater fractional loss in voltage, $(V_a - V_m)/V_a$, which again results in wasted power. Thus in designing a circuit, a compromise between circuit tolerances and degree of stabilisation on one hand must be made against wastage of power on the other.

In the majority of circuits R_a is chosen to be less than R_L so that $R_L > R_a \gg R_g$ and expression 6.8 can be simplified to

$$\frac{dV_a}{dV_m} \simeq \frac{R_a}{R_g} \tag{6.10}$$

For most stabilisers R_g is in the range of 100–1,000Ω whereas R_a

is a few tens of kilohms, so that the stabilisation ratio is of the order of 100.

The stabilisation ratio can be improved by connecting two tubes in cascade. The output voltage from the first tube is stabilised by a second tube of lower voltage with the necessary series resistor R_{a2}, Fig. 6.2. Benson[1] in his comprehensive book on voltage stabilisation, describes a cascade circuit giving approximately 83 V at

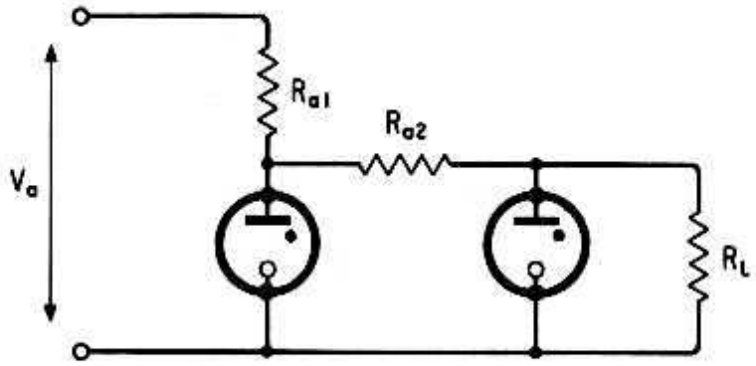


Fig. 6.2. Circuit using two stabiliser tubes in cascade

0–30 mA from a 250 V supply where variations of ~ 40 V on the supply voltage are reduced to ~ 0·01 V at the output i.e. a stabilisation ratio of 4,000.

Tubes can be connected in series and in cascade to obtain output potentials outside the normal range of available stabilisers. However, care must be taken to ensure that all the tubes are operating within their ratings, and that breakdown of one tube does not prevent breakdown of others.

6.2.2. REFERENCE TUBES

An alternative use for the glow-discharge diode is to provide a reference voltage for thermionic valve stabilisation. The simplest circuit arrangement, where the thermionic valve is placed in parallel with the load is shown in Fig. 6.3. The grid voltage V_g is controlled by the potentiometer R_1R_2 and the reference tube. An increase in V_a drives the grid positive so that the anode current increases. This results in an increased drop across R_a which offsets the increase in V_a. As Hunt and Hickman[2] have pointed out, a balance can be obtained if R_1, R_2 and R_a are suitably chosen such that the increase in current through the valve completely compensates for the

increase in V_a. This condition is satisfied when the mutual conductance of the valve, g_m, is given by

$$g_m = \frac{(r_a+R_g)}{V_a} \cdot \frac{(R_1+R_2)}{R_aR_2-R_g(R_1+R_2)} \tag{6.11}$$

where R_g is the resistance of the glow discharge tube and r_a that of

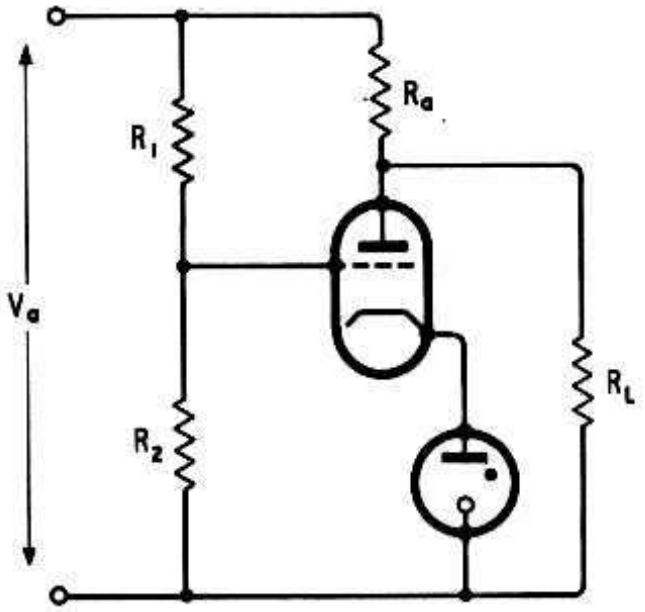


Fig. 6.3. Stabilising circuit with the thermionic valve connected in parallel with the load

the valve. At balance the stabilisation ratio is infinite, although this will only apply over a limited current range.

From Fig. 6.3 it is seen that the reference tube is taking the whole valve stabilising current, and, as with the parallel stabilising circuit (Fig. 6.1), the current range is limited to that of the discharge tube. An alternative system is to connect the valve in series with the load such as in Fig. 6.4. This is a cathode follower circuit, the cathode

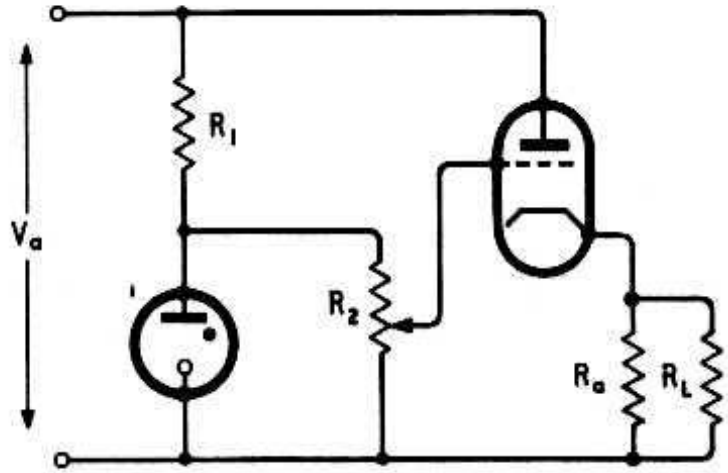


Fig. 6.4. Stabilising circuit with the thermionic valve connected in series with the load

potential 'following' the grid potential as it is varied, independent of variations in the load. R_1 and R_2 can be large resistors since the grid draws very little current. The reference tube can then be run at a modest current, and only has to cope with variations in V_a. Under these circumstances, a tighter tolerance can be placed on

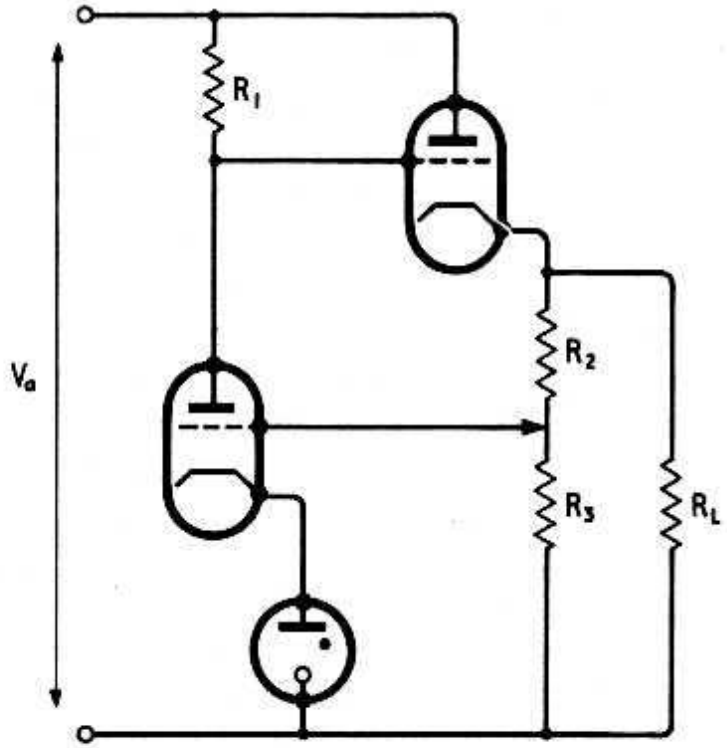


Fig. 6.5. Series-parallel thermionic valve stabilising circuit

the tube. In practice most thermionic valve stabilising circuits use a combination of series and parallel connected valves such as shown in Fig. 6.5.

6.3. Stabiliser Design

In the early glow discharge stabilisers the cathodes were normally coated with a low work function material, either by applying a stable compound to the cathode before assembly and 'activating' it during processing, or by evaporation of an active coating, such as barium or potassium, on to the cathodes whilst on the pumps. These early stabilisers showed considerable drifting on life and were not particularly reproducible from tube to tube; typically, tolerances of 10–15 V about a nominal maintaining potential of 100 V were quoted by the manufacturers.

In 1946, however, as a result of the work by Penning and his colleagues[3, 4] in the Philips Laboratory at Eindhoven, the first

pure metal cathode tube prepared by a sputtering process, the 85A1, became available.[5] Although not a good tube by modern standards, the 85A1 represented a considerable improvement on most of the stabilisers available at that time, and paved the way for a new generation of stabilisers with superior performance, based on the sputtering technique for cleaning the cathode. Comparison of the characteristics of the 85A1 with other tubes at that time were reported by Titterton[6] and Benson, Cain and Clucas.[7] Of note were the improved life characteristics; for a coated-cathode tube the maintaining potential increased by 5–10% to a value outside the specified range after 1,500 hr, change, whereas the 85A1 showed, on average, only a 0·6% change over the same period.

Since the introduction of the 85A1 most manufacturers have produced stabilisers having pure metal cathodes which have been sputtered, and modern tubes are available as reference sources with stabilities at a given current as good as, if not better than, that obtained for a Weston standard cell. An example is the 83A1 with a maximum variation in maintaining potential after the first 300 hr of 0·4% during the next 10,000 hr. In general, the closest tolerance tubes have molybdenum cathodes, but other metal cathodes processed by the same sputtering technique are used to give a variation of the stabilising voltage level. To cover as wide a range of voltage as possible, cathodes with low work function coatings are still manufactured. For example, the OB2, a 108 V tube, has a mischmetall layer (an alloy of cerium and other rare earth metals) sputtered on to a nickel cathode during processing. There is in fact a wide range of modern tubes available (Neale[8] in his book on cold-cathode tube circuit design lists over 150 covering the voltage range of 50–150 V), and many of the tubes having the same characteristics differ markedly in design between manufacturers. However, the general features and the premises on which they are based, detailed in the following sections, are fairly common to the design of all stabilisers.

6.3.1. THE GAS FILLING

The gas filling of stabiliser tubes has been chosen mainly from considerations of the breakdown potential V_s relative to maintaining potential V_m. $V_s - V_m$ is required to be as low as possible, and therefore the gas pressure and electrode spacing are normally chosen to give the *pd* for minimum breakdown potential $V_{s(\min)}$. At the same time, since any anode fall or positive column is a

possible source of oscillations, the anode is generally positioned in the Faraday dark space (see Section 3.1) so that V_m is also a minimum close to the normal cathode fall V_n. To reconcile these two requirements the anode or cathode often has a projecting portion to obtain minimum striking voltage.

Figure 3.13 shows that a rough correlation exists between $V_{s(\min)}$ and V_n for inert gases and their mixtures. It is seen from this figure that $V_{s(\min)} - V_n$ will be smallest for the lowest $V_{s(\min)}$ values. These occur in Penning mixtures having some 0·5–1% of an admixed gas whose ionisation potential is lower than the metastable energy level of the parent gas. Thus 99% neon + 1% argon and 99% helium + 1% argon are common gases used in present day stabilisers, giving $V_{s(\min)} - V_n$ of ~ 30 V or less, depending on the electrode material.

Although chosen from this premise, Penning mixtures have several other attributes which make them particularly suitable for stabilisers and reference tubes. Because of the relatively 'flat' characteristic of the Paschen curve at the minimum (cf. Fig. 2.10) the electrode spacing is not very critical. Also, the variation in V_n with pressure is less than for the respective pure gases alone, which means that the filling pressure has a wider tolerance, and the temperature coefficient is smaller. More important, however, is the ease with which tubes with pure metal cathodes and filled with Penning mixtures can be processed to give reproducible and stable values of maintaining potential. This is partly due to the fact that although the sputtering rate in, for example, neon-argon, is similar to that in pure neon for the same current, the power dissipation is less and the envelope temperature is therefore lower. Since the sputtering technique relies on cleaning the surface, and also trapping gaseous impurities in the sputtered film at a faster rate than such impurities can be released from the envelope, a lower bulb temperature is obviously advantageous.* In the case of helium, little or no sputtering can be obtained without the addition of another inert gas. A second reason for the more stable results is the fact that small traces of impurities which could affect the breakdown and maintaining potentials in a pure gas tube by creating a Penning reaction, have little effect in a gas mixture where the Penning reaction already plays a major role in determining the tube characteristics. Of the possible Penning mixtures, 99% neon + 1% argon has been found to be the most satisfactory from the

*By pulsing the tube it is possible to obtain a high sputtering rate and yet maintain a low envelope temperature, even in pure gases such as neon and argon (see Chapter 4). Tubes with pure gases, particularly trigger tubes, are now manufactured using this technique.

above consideration, and is used in the high stability reference tubes.

The pressure of the gas filling is partly determined by similar considerations of processing. Since the sputtering rate is proportional to $1/p^n$, where n is greater than 2, at high pressures difficulty is experienced in cleaning the cathode, whereas at low pressures sputtering during normal use would be too high and result in erosion of the cathode in a relatively short time. This, however, is not generally the controlling factor, and the pressure range is limited mainly by its effect on the voltage-current characteristic. For good regulation, the discharge should run in the normal glow region over the whole current range, the glow expanding smoothly as the current is increased until at maximum current the whole cathode is covered. Theoretically, the current density in this region is proportional to the square of the pressure, although due to temperature effects strict adherence to the law is rarely found in practice. At low pressures, where the current density is low, the limitation is the prohibitive cathode size which would be required for the discharge to remain in the normal glow region over a reasonable current range. At the high pressures, the limitation is associated with 'jumps' in the characteristic. These are evident as sudden changes in voltage which occur at certain points when the current range is scanned. They can be due to local variations in γ of the cathode surface, to edge effects or geometry effects. Although primarily these are cathode effects, the magnitude is related to the current density and the pressure.

At low pressures the current density is low, and the glow is diffuse, with ill-defined edges. At high pressures the current density is high, and furthermore the glow edges are clearly defined, giving an abrupt change from maximum to minimum current density when going from glowing to non-glowing areas of the cathode. As a result, a barrier can be established, when running the tube, which will cause a discontinuity on varying the current. This barrier may be caused by impurities, by sputtered material diffusing on to the surface, or just by variation in the surface texture. At lower pressures a defined barrier is less likely to occur, and although the regulation may well be affected by contamination or surface variations, discontinuities are less evident. A jump will also occur when the glow expands over the edge of the cathode i.e. from the front to the back of the cathode, and again this is more pronounced for high pressure discharges. In practice, an upper limit of the order of 50 torr is found acceptable for jumps etc., although most reference tubes and stabilisers are filled to a slightly lower pressure. The lower limit of pressure is about 15 torr.

The manufacturer rarely divulges details of the pressure of gas filling, or, if he does, the reasons for the final selection. Consequently, there is a dearth of published information on the actual effect of the gas composition and pressure on the performance of the glow discharge tube as a stabiliser. However, over the last few years Benson and his co-workers at Sheffield University have made a special study of the effect of gas composition. They have examined characteristics such as maintaining potential, impedance, temperature coefficient of maintaining potential and noise of diodes with cylindrical molybdenum cathodes and concentric rod anodes, filled to 40 torr with various inert gas mixtures. Initially, Benson and Chalmers[9] examined the effect of the argon content, over the range 0·001 %–10 % in neon-argon mixtures. They concluded that the most desirable characteristics for stabilisers, with their geometry and gas pressure, were obtained when the argon content was in the region of 1 %, although minimum impedance and temperature coefficient could be realised for argon contents lying between 1 % and 3 %. Later Benson and Burdett[10] examined krypton and xenon as admixtures in neon, and also some tubes containing triple mixtures of neon-argon-krypton and neon-argon-xenon. The values of maintaining potential were given in Chapter 3; small additions of krypton and xenon to neon lowered the breakdown and maintaining potential in a similar manner to argon, as a result of the Penning reaction. The effect, however, was less marked. Impedance, temperature coefficient, and variations on life, showed minimum values for mixtures which in general gave the lowest values of maintaining potential. Although some tubes were measured with a lower gas pressure, 20 torr instead of 40 torr, no conclusions could be drawn on the effect of pressure with these mixtures. The effect of pressure on the V-i characteristic, however, is well known, see Fig. 3.19.

6.3.2. THE ELECTRODE MATERIAL

Providing the anode is not a source of contamination, for example, by running hot and outgassing, the material of which it is made has no effect on the discharge characteristics. The cathode on the other hand plays an important part in providing secondary electrons by the γ processes (see Chapter 1), and the material and the way it is processed contributes to the values of the breakdown and maintaining potentials and to their stability with time.

As already mentioned, two types of cathode are employed in stabilisers and reference tubes; those which are of pure metal and

are processed by sputtering to attain a high surface purity, and those which have a coating of low work function on a metal substrate, the coating usually being another more volatile metal. The former is the more important group, and several metals are employed by the various manufacturers. The commonest metals are molybdenum, used with 99% neon + 1% argon to give a maintaining potential of approximately 85 V, and nickel, used with 99% helium + 1% argon to give approximately 150 V in the presence of an anode fall. Other metals employed are zirconium, giving approximately 75 V with 99% neon + 1% argon, and tantalum and tungsten, giving values within a few volts of molybdenum cathode tubes. The success of molybdenum as a really stable surface is due to the efficient gettering of the gases evolved from the glass by the sputtered layer deposited on the envelope (see Section 4.4.2). In this respect the molybdenum film is superior to most other metal deposits including zirconium, which is recognised as a better gettering material in bulk. It is essential, however, to seal as much of the glass surface with the deposit as possible, particularly the area of the bulb in the proximity of the discharge, and the tubes are therefore designed to ensure as little obstruction to the free passage of the cathode material as possible. The finally processed tube has a heavy deposit on the walls so that the discharge is scarcely visible.

The actual processing schedule of the cathode is most important, and considerable effort is expended by the manufacturer to assess the optimum procedure for each tube type in terms of performance and manufacturing cost. Some of the early processing schedules involved sputtering the cathode whilst the tube was connected to the pumping system, with several changes of the gas filling at different pressures. With better pumping techniques, where the envelope and electrodes are adequately outgassed, and attention is paid to the cleanliness of the cathode before it is mounted in the tube, the schedules have been simplified, and sputtering can be carried out after the tube has been sealed from the pump; this sputtering schedule is often referred to as 'screening'. Typically, the tube is connected to the pumping system and evacuated to 10^{-4} torr. The tube is baked at 350°C for 10–15 minutes, and when cool, the cathode is eddy-current heated to a bright red heat. The tube is then filled to the required pressure of inert gas and sealed off. With a current some ten times the maximum recommended running current, the cathode is processed by sputtering for 24 hr. (A pulse technique may be used for this process.) The importance of choosing the right current for this screening schedule is illustrated in Table 6.1, where the initial characteristics and their 'drift'

during storage ('on shelf') at room and elevated temperatures are given for experimental reference tubes of the 85A2 type, screened at various currents in the author's laboratory. In some cases the sputtering was followed by an ageing process at a lower current.

At 50 mA the deposited layer on the bulb was inadequate to cope with gases evolved from the glass, particularly at the elevated temperature, and drift on life was relatively large. At 200 mA, presumably the bulb ran too hot during sputtering so that the deposit failed to getter adequately the further outgassing of the glass, and similarly larger drifts resulted.

Since the drift over the first hours of taking current subsequent to the sputtering process tends to be larger than drifts experienced later in the tube's life, an ageing process at the normal operating current is often included in the manufacture of high stability tubes. Ageing the experimental 85A2 tubes at 10 mA for 24 hr showed some evidence of improvement in shelf drift (Table 6.1) but a larger sample would be required for a definite conclusion.

Table 6.1

Initial characteristics and 'drift' of experimental tubes processed with modified schedules and left on shelf

Tube No.	*Screening current* (mA) *(for 24 hr)*		*Initial* V_s (V)	*Initial* V_m (V)	*Drifts in* V_s *and* V_m (V)			
					2,000 hr at room temperature		*400 hr at 100°C*	
	High	*Low*			ΔV_s	ΔV_m	ΔV_s	ΔV_m
1	50	—	105	86·00	3	0·60	13	14
2	50	10	107	86·29	6	2·37	10	11
3	50	10	111	85·50	1	0·11	8	5·79
4	100	—	107	87·65	1	0·12	7	3·03
5	100	10	105	86·05	3	0·01	6	1·71
6	100	10	108	86·89	1	0·03	5	2·03
7	150	—	107	87·49	1	0·07	3	0·47
8	150	10	103	87·69	0	0·03	3	0·34
9	150	10	105	86·13	1	0·02	5	3·49
10	200	—	93	86·79	0	0·35	2	2·76
11	200	10	108	86·56	0	0·20	3	2·14
12	200	10	107	86·83	1	0·01	6	2·11

V_m = maintaining potential at 6 mA

For the coated cathodes, a common method of obtaining a suitable surface is to evaporate material, such as barium, on to a metal substrate at the end of the pumping process, before filling with gas and sealing off the tube. This can be achieved by eddy-current heating a suitable chemical compound of the metal in

pellet form mounted in a metal container. Alternatively, the cathode is coated with the metal azide which is decomposed by heating on the pump. Magnesium, potassium, and barium are suitable metals for coating by these methods, giving maintaining potentials of 80 V, 60 V and 50 V respectively with 99% neon + 1% argon. A coating which is used for voltages of the order of 100 V is cerium or mischmetall, which is an alloy of ~ 50% cerium, ~ 25% lanthanum, ~ 10% neodymium, and other rare metals obtained as a by-product in the extraction of thorium from Monazite. Mischmetall is usually sputtered on to the substrate from a small pellet of the material mounted on the anode. The cerium-coated cathode, although not as stable as sputtered metal cathodes, gives a better life performance than the low work-function surfaces formed by evaporation; typically a 1% drift in V_m can be expected over 1,000 hr at the recommended current rating. An important cathode of the coated type is that formed by coating the metal cathode substrate with barium carbonate, and breaking down the carbonate during processing to barium. The carbonate is first decomposed to the oxide by heating the cathode whilst the tube is under vacuum. After filling with gas and sealing off, the tube is then activated by passing an alternating or pulse current to the cathode. This causes sparking at the cathode, and the cathode surface blackens giving the characteristic maintaining potential of barium. The resulting surface is superior to the evaporated barium surface, giving a much longer life, although stability and reproducibility are no better. In the coated cathode tube, and also in tubes made by some manufacturers where the metal cathode is not sputtered, it is common to include a barium getter in the envelope to reduce the effect of contamination.

6.3.3. THE ELECTRODE GEOMETRY

For early stabilisers, the electrodes were mounted in large bulbs cemented into plastic bases in accordance with vacuum tube techniques of the period. With the introduction of the pressed glass foot, where the pins were sealed into the glass and plugged directly into valve holders, stabilisers with loctal bases and later, in miniature executions were developed. More recently, a number of subminiature tubes with 'flying leads' suitable for wiring into printed circuit boards have become available. Decreasing the envelope size has meant decreasing the cathode surface area, with a consequent reduction in the maximum current rating. However, since there is no particular merit in using the tubes as parallel stabilisers, as

opposed to reference tubes for thermionic valve stabilisation, at currents above 10 mA, this has not proved a retrograde step.

The electrode construction is mainly dictated by considerations of processing, current rating, and mechanical strength rather than from tube performance. Since the cathode needs to be thoroughly outgassed, it is conveniently made in sheet metal to facilitate eddy-current heating. It is normally cylindrical, or part cylindrical to present the maximum surface area in a given bulb, although this is not a necessity and other shaped cathodes are used. In fact in the 83A1 a rod cathode is used which is outgassed and cleaned entirely by ion bombardment. It has the advantage that it has no edges or corners so that there is freedom from discontinuities in the spread of the glow. It is not so important to heat the anode, and in the case of the sputtered tube the anode must not present an obstruction to the flow of sputtered particles from the cathode to the glass envelope. Thus, in most tubes a rod anode is employed. In the unsputtered tubes the electrodes can be supported by micas, but for the sputtered tubes the electrodes must rely on the lead-in wires for holding them in position otherwise leakage paths may develop. Some typical structures from three different manufacturers are shown in Fig. 6.6.

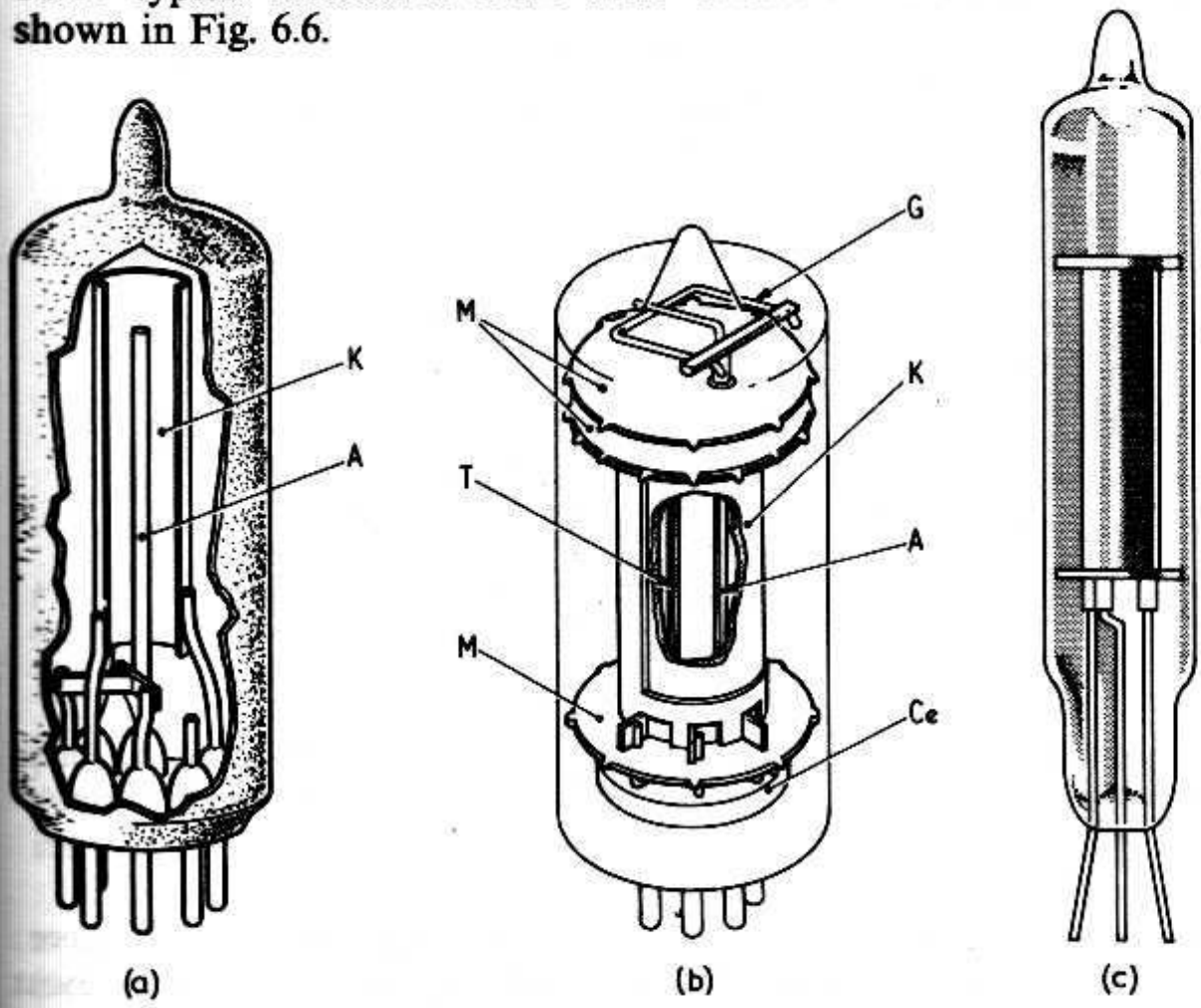


Fig. 6.6. Some typical examples of modern stabiliser or reference tubes
A = anode, K = cathode, G = getter, M = Micas, Ce = ceramic, T = primer

The tube in Fig. 6.6 (a) is a sputtered-molybdenum 85 V reference tube. The cathode and anode are each mounted on three support wires to give rigidity, since micas could not be used in such a heavily sputtered tube. Figure 6.6 (b) shows an unsputtered-nickel cathode tube for 150 V stabilisation. The electrode structure is supported by micas at each end, and a getter is included. The latter has to be well shielded from the nickel cathode otherwise the maintaining potential would be affected. The tube has a priming electrode which, like the anode, consists of a wire running parallel to the axis of the cathode cylinder. Figure 6.6 (c) is a sub-miniature tube giving a V_m of about 100 V, and is constructed of two coaxial nickel cylinders. The outer cylinder, which is the cathode, has a molybdenum lining and the spacing between the cylinders is about 0·5 mm. Although the cathode is sputtered, the electrodes are positioned by micas and a getter is inserted at the top.

The effect of the cathode shape on the performance of stabilisers has been investigated by Benson and Chalmers.[11, 12] They examined molybdenum cathodes mounted with rod anodes. The

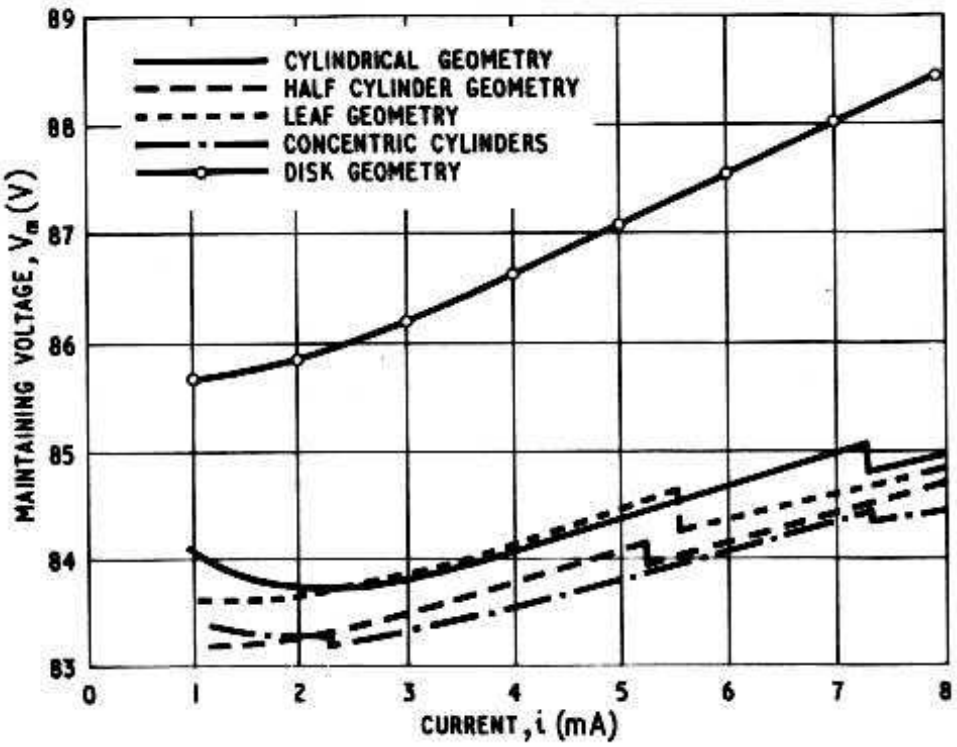


Fig. 6.7. Comparison of typical maintaining voltage—current characteristics of tubes filled to 40 torr of 99·7% Ne+0·3% Ar with various shaped molybdenum electrodes. Benson and Chalmers[11, 12]

various cathode shapes, which had the same surface area, were: cylindrical, half cylindrical, leaf (a plane rectangular cathode bent along its length at 140°), and disk with a central hole. A cylindrical cathode with an inner cylindrical anode was also examined. The

gas filling was 99·7% neon + 0·3% argon at three different pressures, 30, 40, and 50 torr. They found that, due to the variation in the point at which breakdown occurred, the spread in breakdown voltage for tubes of the same type masked any effects of geometry. Similarly, the spread in maintaining potentials, although much smaller, was sufficient to conceal any dependence on geometry, except in the case of the disk cathode which generally appeared to have a value of a few volts higher than the other configurations. The disk was mounted perpendicular to the anode, which passed through the centre hole. The current-voltage characteristics at a given pressure, however, were affected by geometry, particularly the current values at which minimum voltage occurred and at which the glow spread round the edge of the cathode manifest as a step in voltage. Some typical characteristics demonstrating this are shown in Fig. 6.7. Varying the cathode area by using cathodes of different lengths had the effect of changing the slope of the *V-i* characteristic as well as the position of the voltage step. This is shown in Fig. 6.8. The temperature coefficient of maintaining

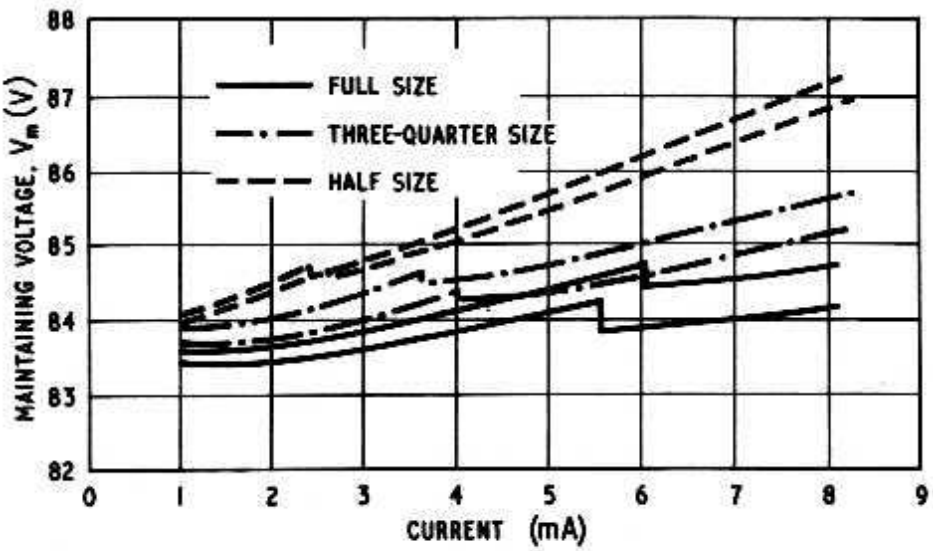


Fig. 6.8. Maintaining voltage-current characteristics for tubes with leaf geometry showing effect of electrode dimensions. Benson and Chalmers[11, 12]

potential, and initial drifts in maintaining potential on drawing current were not significantly affected by geometry, nor were the impedance-frequency characteristics.

If the cathode geometry is such that the glow from one portion overlaps that from another, so that there is a hollow cathode effect (see Section 3.7), then the characteristic will be materially affected. Both the minimum maintaining potential and the d.c. impedance are decreased allowing higher current densities to be taken. This

is illustrated in Fig. 6.9, and has been made use of in the development of the subminiature voltage reference tube type ZZ 1000.[13]

Although the available stabilisers have glass envelopes, alternative constructions have been described in the United States, using ceramic-metal seals, to meet the demand for tubes to withstand the severe environmental conditions experienced in rockets.

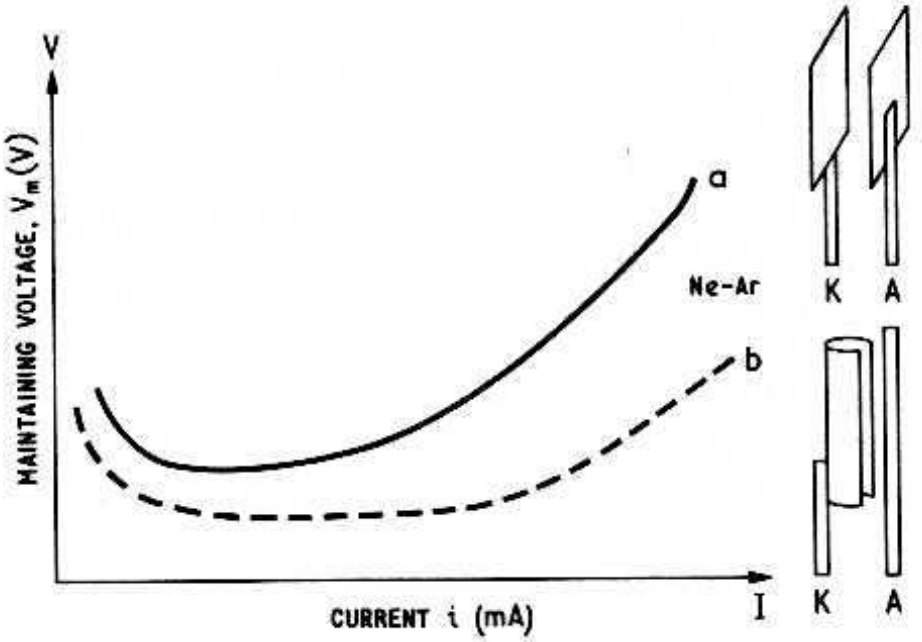


Fig. 6.9. Effect of a hollow cathode, used in the ZZ1000 (geometry b) on the maintaining voltage/current characteristic

Culp and Koskos[14] described an 85 V reference tube in which the molybdenum cathode in the form of a cylindrical cup acted as the tube envelope. A coaxial rod anode was held in position by a ceramic washer sealed to the end of the cathode cup. The sealing process was carried out in a bell jar filled with the inert gas, by heating the assembly to a temperature sufficiently high to braze the metal to the ceramic with a high temperature solder. A similar construction was described by Lafferty[15] using a titanium cup as the cathode which was coated with barium during processing. Because of the gettering action of the titanium to active gases, and the high temperature used during processing, gas contamination on life should be negligible. Lafferty claimed that such a tube could be operated at a temperature of 500°C.

6.3.4. PRIMING

When the stabilising circuit of Fig. 6.1 is first switched on, the potential across the load will be somewhat greater than V_s, until the glow discharge tube breaks down. It is important that the delay

in the breakdown of the tube, which has been discussed in Chapter 5, should not be excessive if damage to the circuit is to be avoided. Most composite cathodes are photoemissive, and the ambient light can produce sufficient electrons to reduce the strike delay to the order of milliseconds; some sputtered tubes also exhibit photemission under ambient light. If, however, the tubes are mounted in the dark or the cathode is not photoemissive, the delay can be considerable, of the order of minutes. A delay of this magnitude would not be acceptable for many applications, and therefore a source of sufficient initiatory electrons (priming) must be incorporated in the tube.

Some stabilisers are provided with an auxiliary priming electrode which is connected to the supply voltage via a high resistor. Since

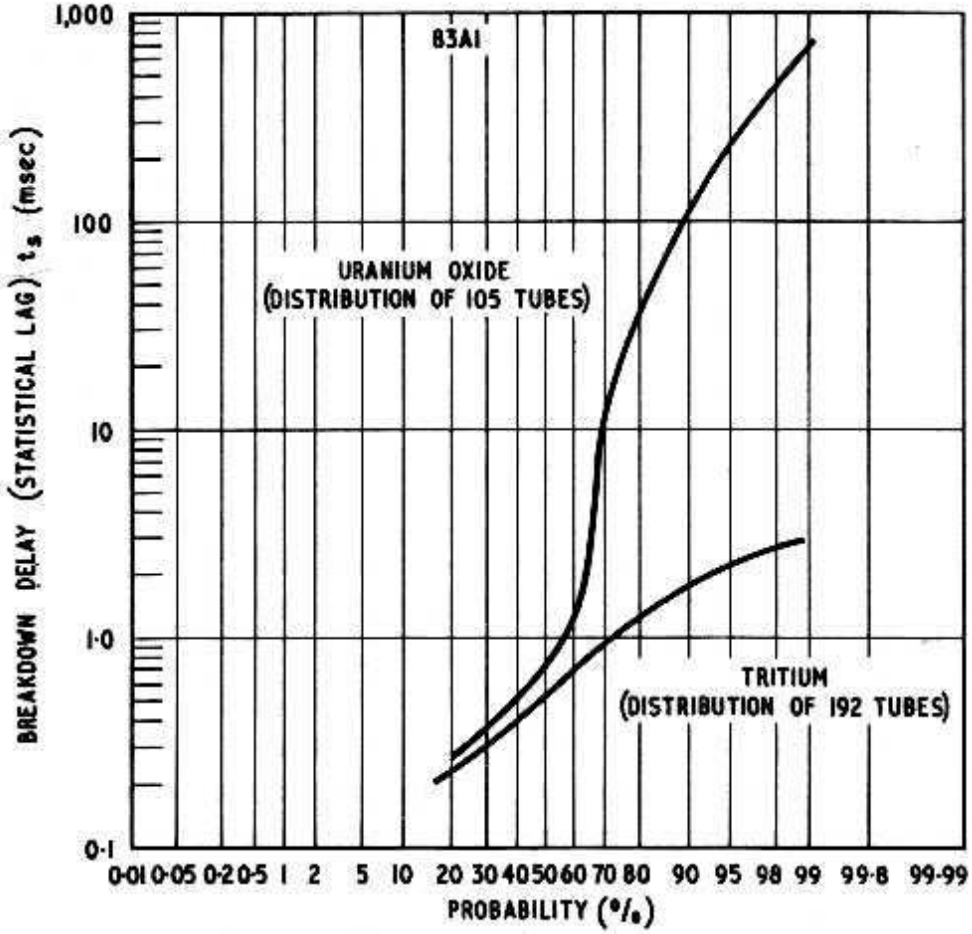


Fig. 6.10. Breakdown delay time (statistical lag) distribution curves for two batches of 83A1 reference tubes, showing the effect of the radioactive priming

the priming gap does not have a load across it, the full supply voltage is applied between priming electrode and cathode. The supply voltage is substantially in excess of the breakdown voltage, and therefore both the statistical and formative lags in the priming gap are reduced (cf. Figs. 5.2 and 5.10 respectively). More important, if the priming current is of sufficient magnitude the breakdown

potential of the main gap is practically equal to the maintaining potential, so that an excess voltage does not have to be applied across the load before breakdown. The priming electrode, however, involves an extra circuit component with appropriate connections, and for most purposes sufficient priming can be obtained more conveniently by introducing a small radio-active source into the tube. Since delays of milliseconds are generally acceptable, the source can be very weak, and will not in any way present a health hazard. One method of providing such a radio-active source, is to coat a small area of the inside of the bulb with uranium oxide in the form of a paint. This reduces the delay to the order of 10–100 msec. However, a more satisfactory source is obtained by introducing a small amount of tritium gas (approximately 1 μc) into the gas mixture. Tritium emits only low energy β-particles which cannot penetrate the glass envelope. The effect of tritium on the delays was discussed in Chapter 5. A comparison of breakdown delays in the 83A1 with uranium oxide and tritium priming is shown in Fig. 6.10, where the probability of a tube firing can be seen as a function of the breakdown time. Thus 70% of the tubes containing uranium oxide could be expected to break down within 10 msec, whilst for tritium priming 70% could be expected to break down within 1 msec.

6.4. The Performance of Stabilisers

In the previous section the design of stabilisers to make the best use of the properties of a glow discharge (see Chapter 3) has been described, and in so doing the performance of some tubes has been outlined. The essential parameters of stabilisers, based on the discharge properties, such as breakdown and maintaining potentials, d.c. impedance, current limitations, and sometimes the spread in these parameters, are given in the data sheets by the manufacturers. Therefore, in discussing the performance of stabilisers such parameters will not be included, however some of the aspects of the performance important to stabilisation, but not always given in detail by the manufacturer, will be described.

6.4.1. DRIFTS IN MAINTAINING POTENTIAL ON LIFE

Unlike many electronic components, the stabiliser or reference tube rarely ends its operational life with a catastrophic failure. Rather, it exhibits a drift of characteristics, in particular the maintaining

potential, during the whole of its operational life, and the tube is considered to have failed when the deviation from initial values becomes unacceptable. In general, the drift is in a positive direction, the voltage increasing with time, but some tubes exhibit a negative drift during the early part of their lives, i.e. over the first few hundred hours. The drifts are caused by changes in the cathode surface as a result of positive ion bombardment, and/or by clean-up of the gas. Ion bombardment is more detrimental in the case of coated cathodes, where the discharge properties are dependent on a relatively thin cathode coating, than in the case of pure metal cathodes. In consequence drifts are considerably larger for tubes with the former type of cathode; this is illustrated in Fig. 6.11

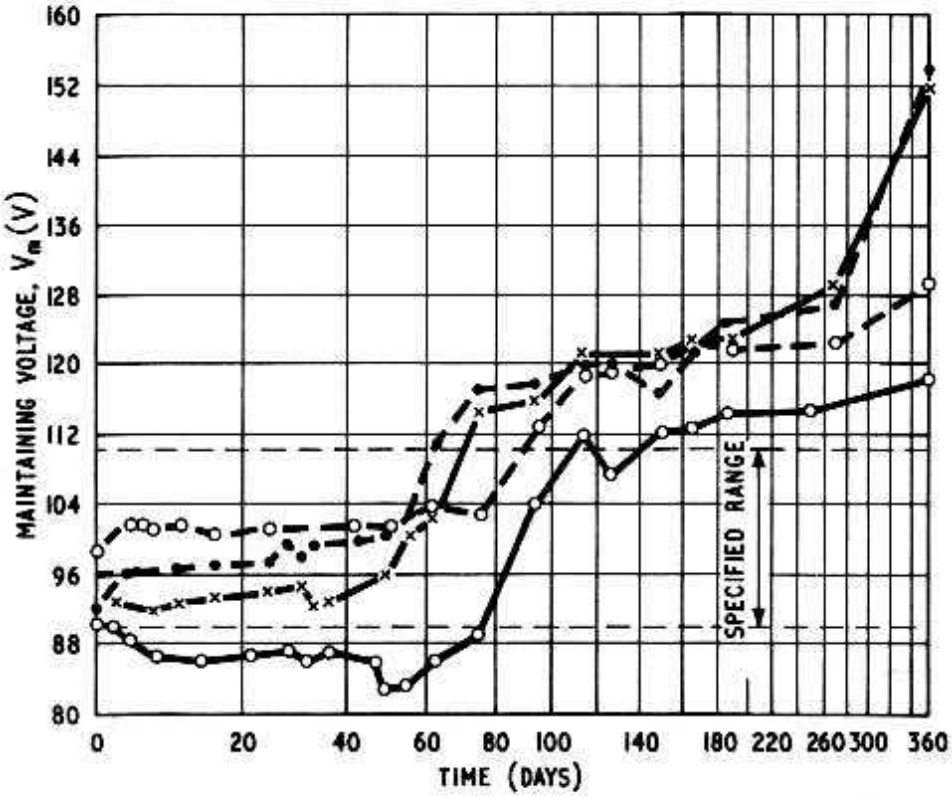


Fig. 6.11. The maintaining potential of tubes type 7475 as a function of time, run continuously at 5 mA, Benson[16]

where the voltage of some typical 7475 tubes (having a barium-coated cathode) is plotted against time. The curves are taken from a paper by Benson[16] in which a number of stabilisers from various manufacturers were compared. A comprehensive study of sputtered cathode tubes has been made by Bond[17] covering results over a period of four years. He found that, for the 83A1, there was a pattern of life behaviour which appeared to consist of a stabilising period of approximately 300 hr followed by very stable operation. The results of the life tests were treated statistically, and the 80% points from the distributions of maintaining potential (i.e. 80%

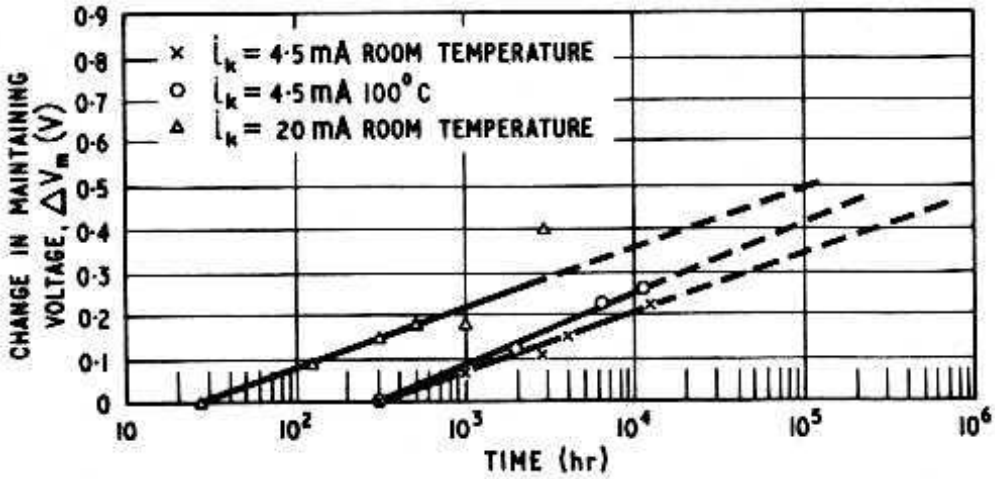


Fig. 6.12. Maximum change in V_m for 80% of batch of 83A1 tubes as a function of time, under various life conditions, Bond[17]

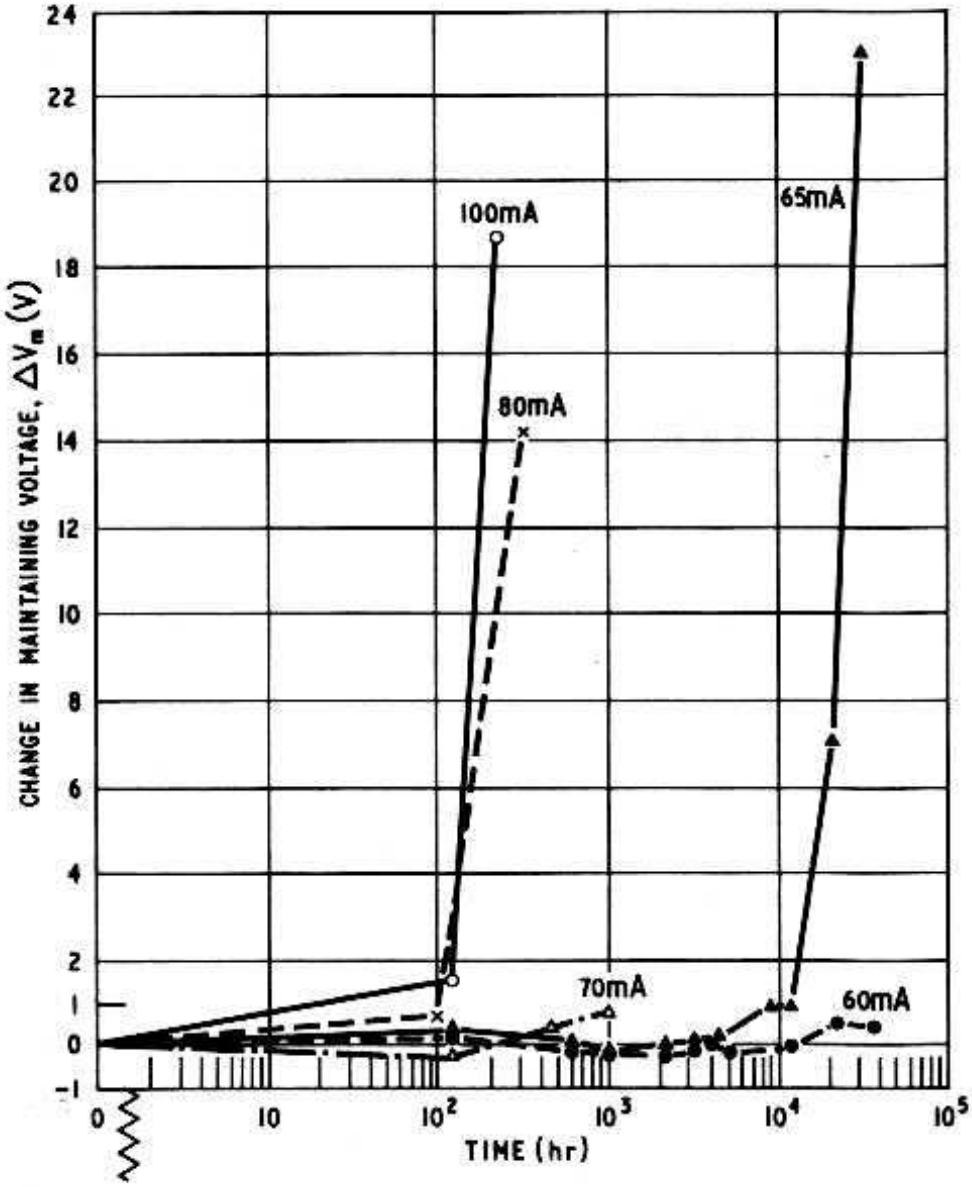


Fig. 6.13. Average change in V_m for 80% of a batch of 75C1 tubes as a function of time, run at various currents at room temperature. Bond[17]

could be expected to have ΔV_m below the plotted value), after the initial stabilising period, plotted against a logarithmic time scale (see Fig. 6.12). He found that the points only deviated slightly from a straight line, and that the curves for two different currents were roughly parallel. Bond also presented life data on the 75C1, a stabiliser with a sputtered zirconium cathode, at several different currents; the curves are shown in Fig. 6.13. The tube showed a relatively small drift until, towards the end of life, the voltage rose markedly. By taking the period for a 3 V change in V_m to occur, and plotting this against current on logarithmic scales, a straight line was obtained (Fig. 6.14), suggesting that there is an inverse

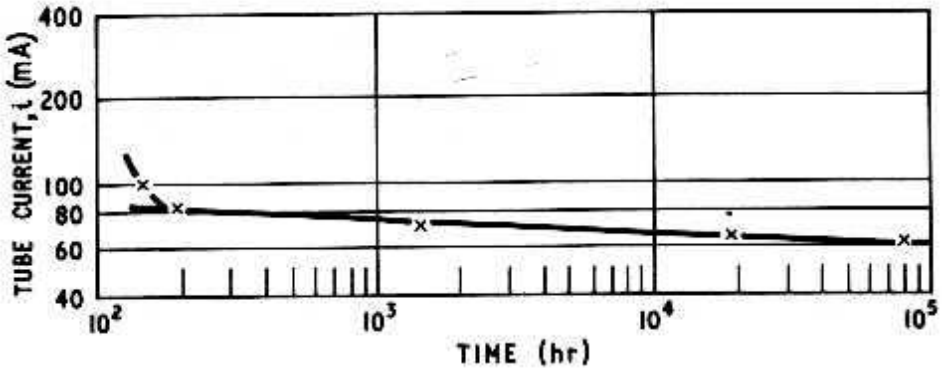


Fig. 6.14. Length of life for a 3 V change as function of the current for tube type 75C1. Bond[17]

power relationship between the length of life and the current. From these results Bond suggests that predictions of life at low current can be deduced from results at higher currents. This being so, it can be assumed that if the working conditions are correctly chosen, and these are usually the manufacturer's recommended

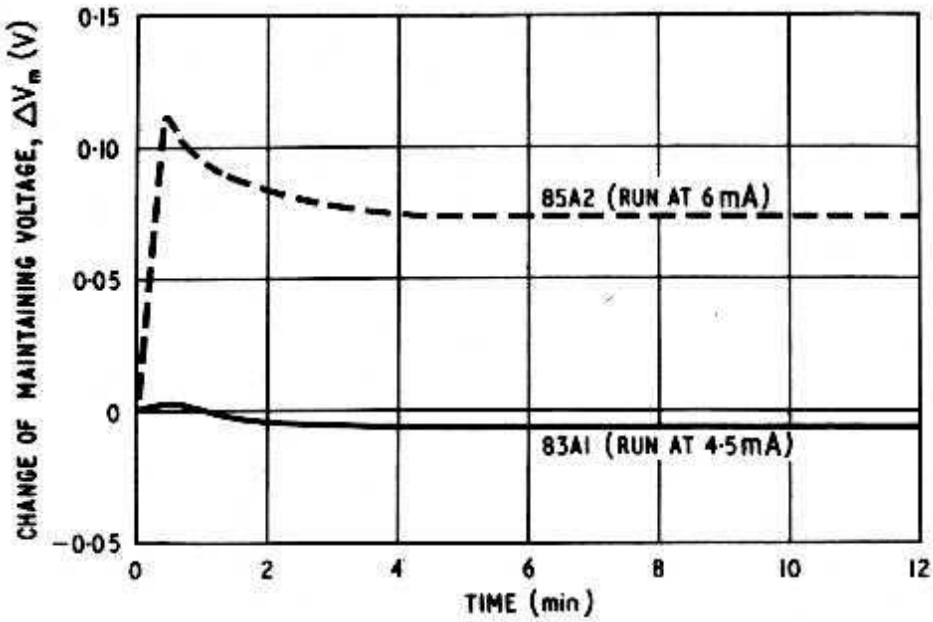


Fig. 6.15. Typical inital drift curves for tube types 85A2 and 83A1

operational conditions, then lives of tens of years can be expected for sputtered cathode tubes.

Apart from drifts on life there is an initial drift which occurs in all glow discharge tubes in the first few minutes after switching on. This is partly due to the temperature of the tube reaching equilibrium, but can be partly due to cleaning of the cathode surface. The magnitude and duration of these initial drifts have been studied by Benson[16] for a number of tubes available in the U.K. This drifting is a very variable quantity from tube to tube of the same type, and also from operation to operation of the same tube. However, in general the drifts are smaller for sputtered cathode tubes than for composite cathode tubes; typical curves for the former type are shown in Fig. 6.15. The low value for the 83A1 is mainly due to the small thermal capacity of the cathode. The initial drifts increase markedly with current as might be expected, and also increase with the age of the tube.

6.4.2. TEMPERATURE COEFFICIENT OF MAINTAINING POTENTIAL

The effect of the envelope temperature on the maintaining potential of a glow discharge has been explained in Section 3.4. Most stabilisers show a negative coefficient, the voltage decreasing with increasing temperature, but a few tubes, particularly those giving V_m of the order of 150 V such as the 150B2 and QS150/15, exhibit positive coefficients. This is due to the presence of an anode fall. Although the variation of voltage with envelope temperature is not strictly linear (an example is shown in Fig. 6.16), for many tubes the deviation from a straight line is only very slight over a temperature range of, say, room temperature to 100°C, and a constant temperature coefficient can be assumed. The coefficient depends on geometry and gas filling and on the tube current. In general the sputtered cathode tubes, particularly those used as reference tubes, have lower coefficients than others. Values are usually quoted by the manufacturers.

The temperature coefficient refers to equilibrium conditions. Since V_m is a function of the density distribution set by the cathode and envelope temperature, sudden changes in envelope temperature will give a non-equilibrium situation which can result in initial values of V_m which will differ from the final equilibrium value. This was demonstrated by Smith[18] who showed that V_m 'overshot' the final value when the tube was suddenly transferred from one oil bath and immersed in another at a different temperature. A typical curve for an 85A2 transferred into baths at different temperatures

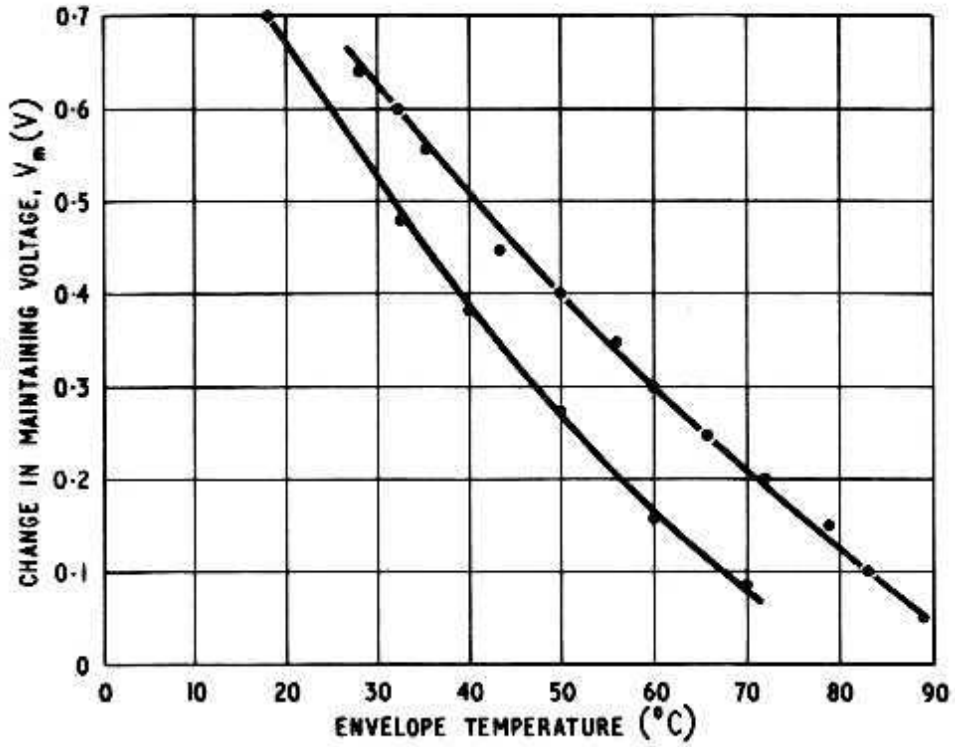


Fig. 6.16. Typical curves for the change in maintaining voltage as a function of envelope temperature for the 90C1 at 20 mA

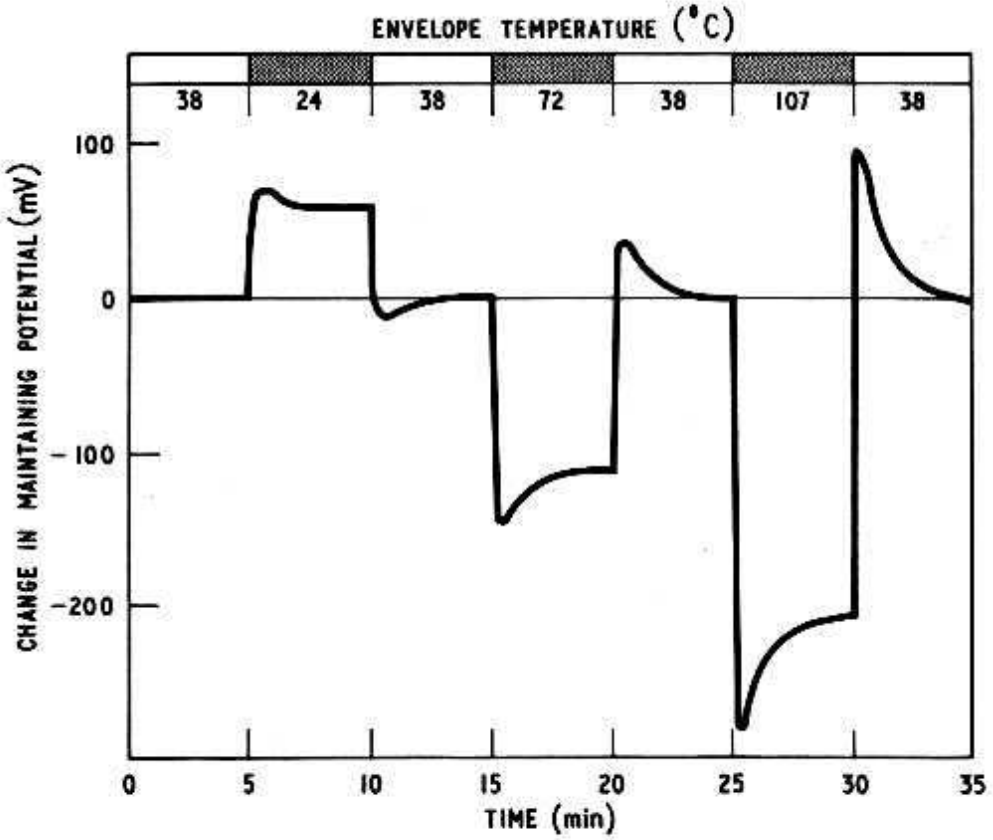


Fig. 6.17. The effect of changes in envelope temperature of an 85A2, by transferring the tube into oil baths at different temperatures, as a function of time

is given in Fig. 6.17. The explanation is as follows: the immediate effect of raising the envelope temperature is to heat the gas in its vicinity which expands and compresses the gas in the cathode region. The increase in density in the cathode region lowers V_m. The cathode temperature, however, also rises although lagging behind the envelope temperature, and this tends to reduce the density in the cathode region, raising V_m again to the equilibrium value. The effect could be important in applications where rapid changes of temperature are likely to take place, for example in rocket-borne equipment.

6.4.3. IMPEDANCE AND NOISE

The impedance of a glow discharge to a.c. signals is higher than suggested by the static characteristics; this has been discussed in Chapter 3. The impedance at relatively low frequencies is important in stabilising circuits since the unstabilised supply is likely to contain a small a.c. ripple, for example a 50 or 100 c/s ripple if

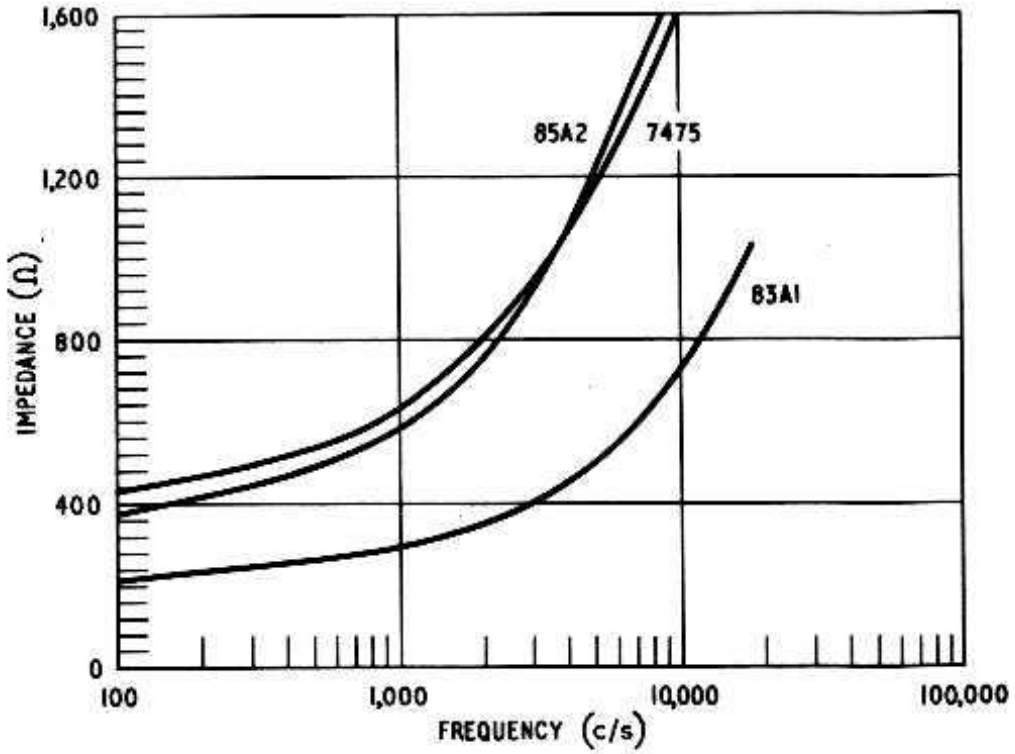


Fig. 6.18. Typical impedance/frequency curves for three types of stabiliser tubes

operated from the mains. At higher frequencies the impedance will determine the tube's response to transients.

The impedance-frequency curves for most stabilisers are very similar and data for several tubes have been given by Benson and his co-workers.[19] Typical curves for three tube types are given in Fig. 6.18; the 85A2 and 7475 were run at 5 mA and the 83A1 at

4·5 mA. At low frequencies up to 100 c/s the impedance is mainly resistive with a value close to the d.c. incremental resistance. At higher frequencies the reactive component becomes comparable to the resistive component, see Figs. 3.29 and 3.30. The spread of values for any tube type follows a similar pattern to other characteristics, i.e. the narrowest spread occurs for the high stability reference tubes with sputtered cathodes, whilst the largest variations occur for coated cathode tubes. Both the reactive and resistive components depend on the gas filling and pressure, and the resistive component also depends on the cathode material. This

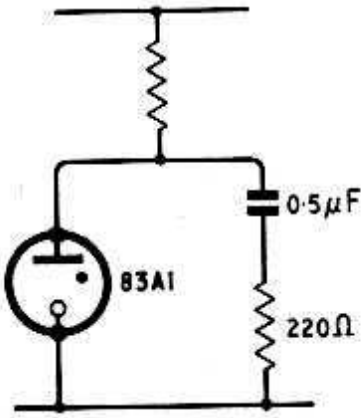


Fig. 6.19. Circuit for reducing the a.c. impedance of an 83A1 reference tube

accounts for the differences between tube types. The impedance also varies with the tube current, being almost linearly related to the reciprocal of the current. These parameters were discussed in more detail in Chapter 3.

The increase in impedance with frequency can be compensated over a fairly wide frequency range by connecting a capacitor across the tube. If, however, the capacitor is too large, > 1 μF, relaxation

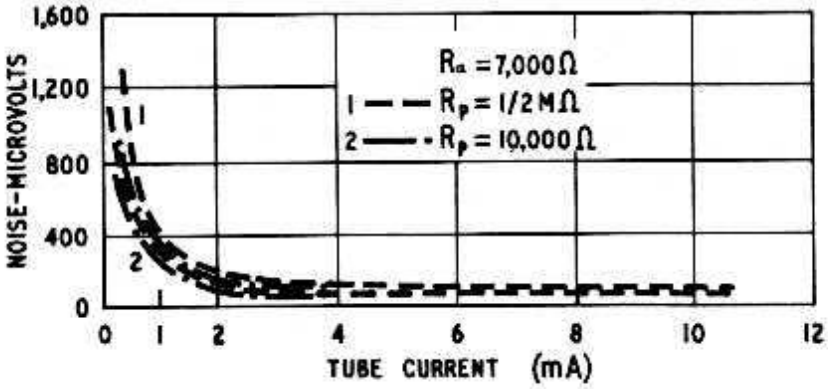


Fig. 6.20. The r.m.s. noise voltage as a function of current for the 85A2. Bache and Benson.[20] *R_p = power supply impedance*

oscillations can occur. A suitable corrector network for the 83A1 is given in Fig. 6.19.

Modern stabilisers are free from spurious noise due, for instance, to the glow wandering or an oscillating anode glow. They do,

however, exhibit shot noise and/or thermal noise (see Chapter 3), which is uniformly spread over all frequencies. The noise level is relatively low, of the order of 100 μV r.m.s. and is no embarrassment for stabilising applications.

As the current is decreased the equivalent noise level rises, and can be a factor of two or three higher near the minimum current. This is shown by typical curves for the 85A2 obtained by Bache and Benson[20] and reproduced in Fig. 6.20. The noise level at low current may therefore be troublesome in the most stringent application.

6.5. The Corona Stabiliser

The voltage of a glow-discharge tube has an upper limit of a few hundred volts even with an anode fall and positive column present. If, however, the electrode geometry is arranged to produce a high electric field in the proximity of one of the electrodes the tube will operate stably in the corona discharge mode (see Section 2.4) at a considerably higher voltage, up to several kilovolts. The field conditions can be realised by making one of the electrodes either a point or a wire. Since the current which can be drawn in the case of a point is too small for useful stabilisation, a coaxial cylindrical geometry is normally employed with the inner cylinder, in the form of a rod or wire, connected as anode and the outer cylinder as cathode. For a practical-sized tube the current is much smaller than for a glow discharge stabiliser, typically less than 1 mA.

Unlike the glow discharge, the breakdown potential for a corona discharge only exceeds the potential at which it operates (corona voltage) by a few per cent. A typical characteristic is shown in Fig. 2.18. Providing this is taken into account, the formulae for designing the required circuit and for assessing the stabilisation ratio, given in Section 6.2, apply equally to the corona stabiliser.

6.5.1. THE DESIGN OF A CORONA STABILISER

From Chapter 2, it was seen that for a coaxial cylindrical electrode geometry, with the central wire connected as cathode, the breakdown potential V_s was a function of the pressure × the radius of the wire according to the expression:

$$\frac{V_s}{\log_e b/a} = f(pa) \tag{6.13}$$

where a is the radius of the wire and b that of the cylinder. Llewellyn Jones[21] and Meek and Craggs[22] have shown experimentally that Equation 6.13 holds also for a positive-wire corona; some empirical curves are shown in Fig. 6.21. As Cohen and Jenkins[23] have

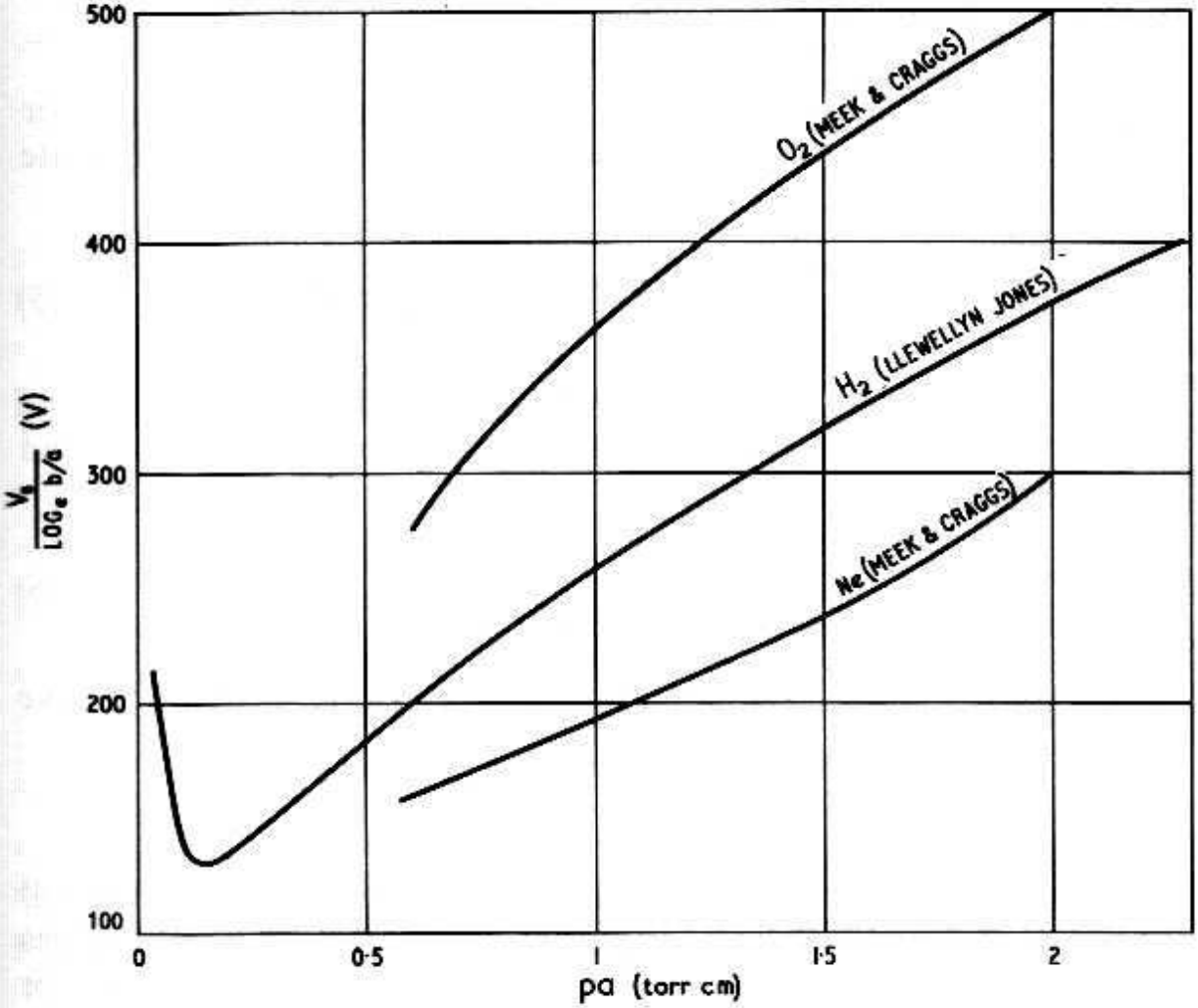


Fig. 6.21. Positive-wire corona breakdown potential for coaxial cylindrical-electrode geometry as a function of pressure × the anode wire radius

pointed out, at the higher values of pa the curves approximate to straight lines so that Equation 6.13 can be written

$$\frac{V_s}{\log_e b/a} = \mathrm{K}pa \tag{6.14}$$

where K is a constant depending mainly on the gas but also, to a lesser extent, on the cathode material. Values of K deduced by Cohen and Jenkins from published curves are given in Table 6.2. Thus, since V_s is close to the corona voltage V_c, by suitable choice of gas and electrode dimensions corona stabilisers can be designed to cover a wide range of stabilising voltages.

There are two other factors to be considered in the design of corona stabilisers, and they are the incremental resistance and the

Table 6.2

Evaluation of K in Equation 6.14 according to Cohen and Jenkins [23]

Gas	*Argon*	*Helium*	*Hydrogen*	*Nitrogen*	*Neon*
Breakdown parameter K	16×10^2	$8{\cdot}3 \times 10^2$	30×10^2	42×10^2	$7{\cdot}5 \times 10^2$

current range. According to Loeb[24] the incremental resistance per unit length R_g, for a coaxial system, is given from approximate calculations as:

$$R_g = \frac{p \log_e b/a}{2V_c k}\left[(b^2 - a^2) - (r_1^2 - a^2)\frac{\log_e b/a}{\log_e r_1/a}\right] \qquad (6.15)$$

where k is the mobility of the ions at 1 torr and r_1 the radius of the ionising sheath round the wire. Since normally $r_1 \simeq a$; and $b \gg a$, Equation 6.15 reduces to

$$R_g \simeq \frac{b^2 p \log_e b/a}{2kV_c} \qquad (6.16)$$

which since $V_s \simeq V_c$ can be combined with Equation 6.14 to give

$$R_g \simeq \frac{b^2}{2akK}$$

Thus for given dimensions the gas with the highest mobilities will give the lowest impedance. Hydrogen and helium, with mobilities several times those of the other inert gases, are most suitable for corona stabilisers, and are most widely used. A high ratio of a/b would also be advantageous, but increasing a/b decreases the current range of stabilisation, until at 0·37, according to Lichtman,[25] no corona discharge region occurs. The current range, which is limited at the low end by intermittent breakdown and at the upper end by transition to a glow discharge, also depends on the gas composition and its pressure; in general the maximum current increases with increasing pressure. Both the incremental resistance and maximum current are directly proportional to the length of the coaxial system, so that it is advantageous to make the electrodes as long as possible within the given envelope.

Most of the early work of the design of practical corona stabilisers was carried out at the United States Naval Research Laboratory.[26] Tube dimensions and several gas fillings were investigated. Hydrogen was found most suitable for the high voltages, but a mixture of helium and hydrogen was more satisfactory for lower voltages $\sim$ 500 V. Culminating from this work

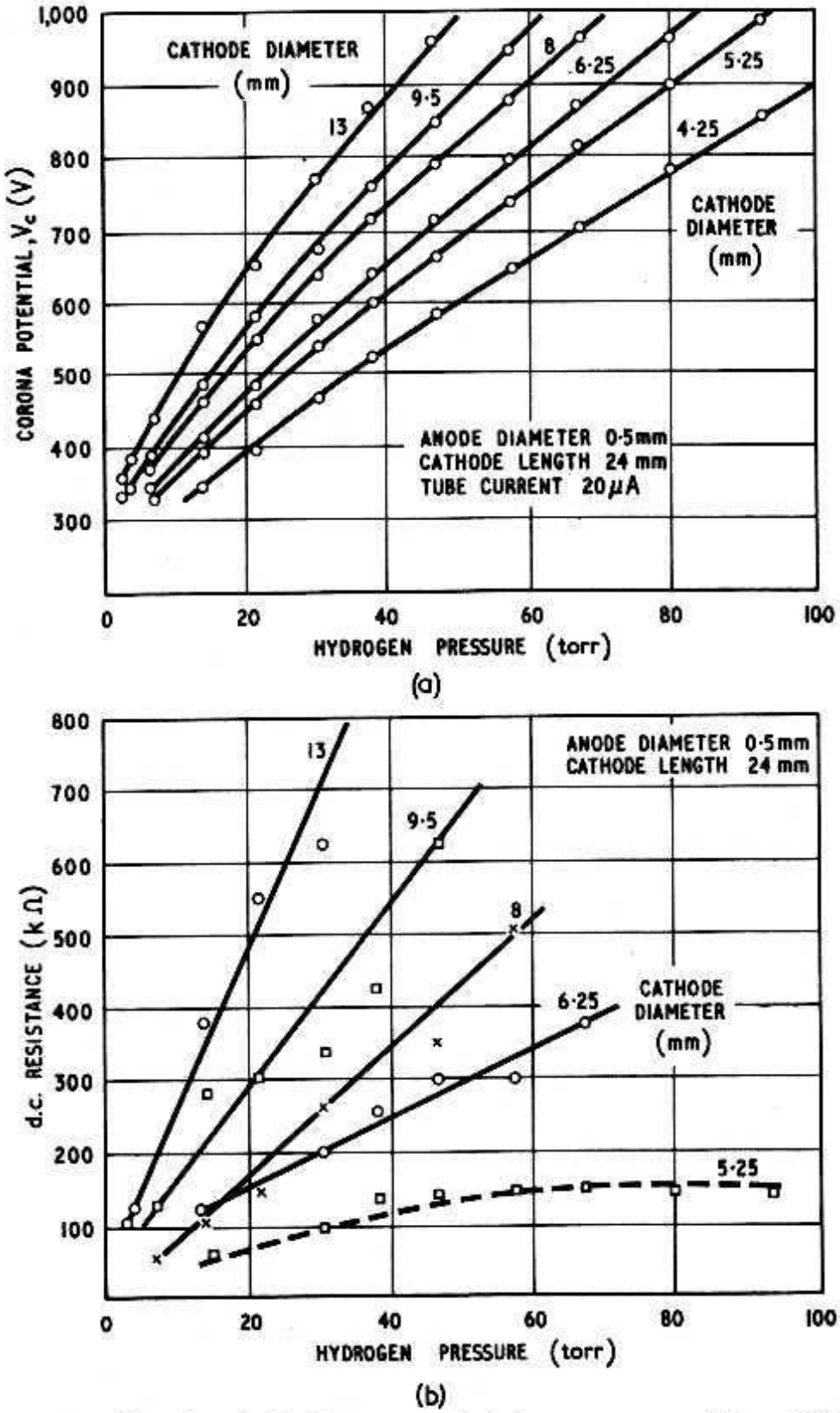


Fig. 6.22. The effect of cathode diameter on (a) the corona potential and (b) the d.c resistance of positive wire coaxial-cylinder corona tubes. Shelton and Wade[27]

was the design of a 700-V stabiliser in a miniature execution having a current range of 50 μA and an incremental resistance of about 1 MΩ. Lichtman[25] reviewed these early studies and commented on the structural design, and performance in circuits. He pointed out that eccentricity of the anode within the cathode cylinder could be expected to cause variations in the incremental impedance.

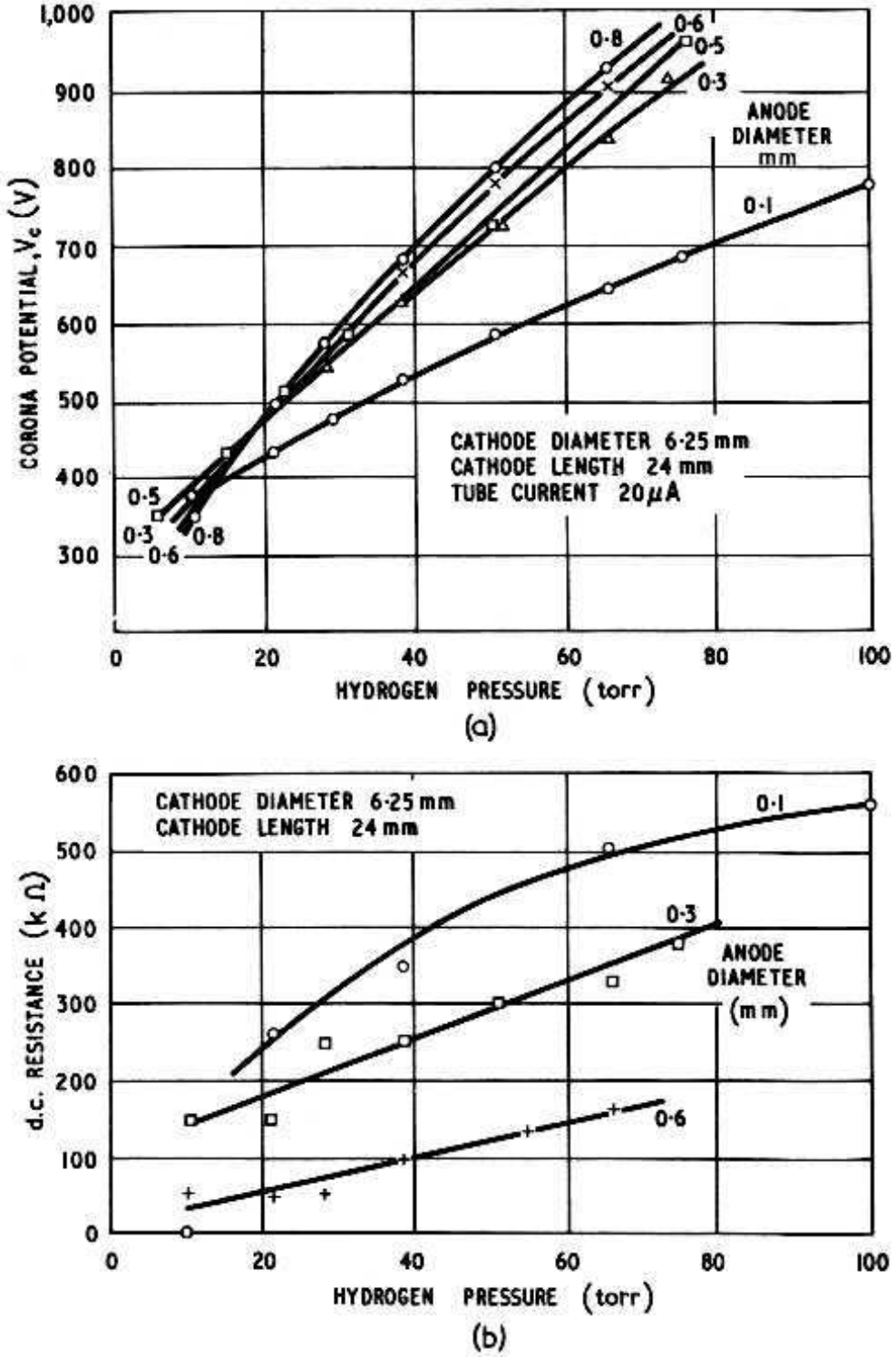


Fig. 6.23. The effect of anode diameter on (a) the corona potential and (b) the d.c. resistance of positive-wire coaxial-cylinders corona tubes. Shelton and Wade[27]

Shelton and Wade[27] have also investigated electrode dimensions and gas filling, comparing hydrogen, helium and argon. The effect on the breakdown potential and impedance of varying the cathode diameter for a hydrogen-filled tube is shown in Fig. 6.22, whilst the effect of changing anode diameter is indicated in Fig. 6.23; in both graphs the breakdown voltage and impedance are plotted as a function of the gas pressure. They found that asymmetry in the electrode structure had little effect on V_s, but a marked effect on the current range by lowering the maximum current. With a nickel

cathode of 6 mm diameter and 30 mm length and an anode of 0·75 mm diameter, currents of several hundred microamps could be passed before the maximum current was reached. In He and Ar the maximum current was limited by onset of a steady glow discharge, but in hydrogen commencement of relaxation oscillations was the determining factor.

A study of the hydrogen-filled corona stabiliser, particularly its processing, has been made by Collinson and Hill.[28] They found that the cleanliness of the cathode not only affected the breakdown voltage but also the current range; the maximum current was more than doubled by sputtering the cathode in a glow discharge at 20 mA for 3 hr with a gas pressure of 5 torr. They also measured the effect of ambient temperature on the corona voltage, and found a small positive temperature coefficient over the whole temperature range, amounting to a few volts for a 100°C variation, i.e. of the order of 0·005% per °C.

More recently Cohen and Jenkins[23] have described the design criteria and the development of a number of hydrogen-filled corona stabilisers covering the range 350–7000 V. Although the required breakdown potentials could be obtained from one design by varying the gas pressure, the incremental resistance would have been unacceptably high for some of the range. It was found that a reasonable compromise to cover the voltage range was to use three different electrode systems. For the high voltages, 4·5–7 kV, the cathode and anode diameters were 13·2 mm and 3 mm respectively, for 2–4·5 kV they were 8 mm and 1·2 mm, and for the lower range 350–2000 V 7 mm and 1 mm diameters were used. The two larger electrode assemblies were mounted in B9G based envelopes whilst the lower voltage range tubes were in B7G based envelopes. The d.c. resistance of some of the tubes, particularly those operating at high voltage, could be quite small or even negative at some currents due to thermal effects; increasing the current raised the gas temperature, thus reducing the density to such an extent that a lower potential resulted. The resistance of such tubes, however, to a.c. signals was very much higher even at low frequencies; for example at 100 c/s it would be several hundred kilohms. The life of the tubes depended on clean-up of the hydrogen manifest in a drop in corona voltage with time. If the useful life of a tube was defined as the time taken for the voltage to alter by the tolerance originally allowed, then lives in excess of 10,000 hr could be expected for most tubes operating at 250 μA.

Tubes of similar dimensions to those described by Cohen and Jenkins are now available from a number of manufacturers. Structurally, they are almost identical; the anode is positioned in

the cathode cylinder by ceramic spacers at each end, which also restricts the discharge to the inside of the cathode cylinder. The connector to the anode is taken out through a top cap to allow adequate insulation. Corona tubes can be used in a similar way to glow-discharge tubes for parallel stabilisation in high voltage power packs supplying currents of the order of a few hundred microamps. The series resistor will normally be in the megohm range, and gives stabilisation ratios of the order of 10–100. Corona tubes find applications in power supplies to Geiger-Müller counters, cathode ray tubes, photomultipliers, image tubes, etc., and can be used for voltage reference purposes in thermionic valve-stabilised circuits.

REFERENCES

1. BENSON, F. A., *Voltage Stabilisation*. Macdonald, London, (1965)
2. HUNT, F. V. and HICKMAN, R. W., *Rev. Sci. Instru.*, **10**, 6, (1939)
3. PENNING, F. M. and MOUBIS, J. H. A., *Philips Res. Rep.*, **1**, 119, (1946)
4. JURRIAANSE, T., PENNING, F. M. and MOUBIS, J. H. A., *Philips Res. Rep.*, **1**, 225, (1946)
5. JURRIAANSE, T., *Philips Tech. Rev.*, **8**, 272, (1946)
6. TITTERTON, E. W., *J. Sci. Instr.*, **26**, 33, (1949)
7. BENSON, F. A., CAIN, W. E. and CLUCAS, B. D., *J. Sci. Instr.*, **26**, 399, (1949)
8. NEALE, D. M., *Cold Cathode Tube Circuit Design*, Chapman & Hall, London, (1964)
9. BENSON, F. A. and CHALMERS, P. M., *Proc. Inst. Elec. Engrs.*, **C106**, 82, (1959)
10. BENSON, F. A. and BURDETT, G. P., *Proc. Inst. Elec. Engrs.*, **B108**, 501, (1961)
11. BENSON, F. A. and CHALMERS, P. M., *Proc. Inst. Elec. Engrs.*, **B109**, 290, (1962)
12. BENSON, F. A. and CHALMERS, P. M., *Radio Electron. Compon.*, **4**, 189, (1963)
13. *Electronic Applications (Holland)*, **22**, 35, (1961)
14. CULP, J. W. and KOSKOS, P., *I.R.E. Trans. Electron Devices*, **ED–4**, 144, (1957)
15. LAFFERTY, J. M., *I.R.E. Trans. on Electron Devices*, **ED–5**, 143, (1958)
16. BENSON, F. A., *Electron. Eng.*, **24**, 396, 456, (1952)
17. BOND, M. E., *Electron. Eng.*, **34**, 798, (1962)
18. SMITH, J., *Brit. J. Appl. Phys.*, **9**, 122, (1958)
19. BENSON, F. A. and MAYO, G., *Electron Eng.*, **26**, 206, (1954); **28**, 124 (1956)
20. BACHE, H. and BENSON, F. A., *Electron Eng.*, **24**, 328, (1952)
21. LLEWELLYN JONES, F., *Ionisation and Breakdown in Gases*, Methuen, London, p. 118, (1957)
22. MEEK, J. M. and CRAGGS, J. D., *Electrical Breakdown of Gases*, Clarendon Press, Oxford, p. 109, (1953)
23. COHEN, E. and JENKINS, R. O., *Proc. Inst. Elec. Engrs.*, **B107**, 285, (1960)
24. LOEB, L. B., *Fundamental Processes of Electrical Discharges in Gases*, Wiley, New York, p. 494, (1939)
25. LICHTMAN, S. W., *Proc. I.R.E.*, **39**, 419, (1951)
26. BLIFFORD, I. H., ARNOLD, R. G. and FRIEDMAN, H., *United States Naval Research Laboratory*, **Rep. No. N–3140**, (1947)
27. SHELTON, E. E. and WADE, F., *Electron. Eng.*, **25**, 18, (1953)
28. COLLINSON, A. J. L. and HILL, D. W., *J. Sci. Instr.*, **32**, 13, (1955)

CHAPTER 7

Switching Diodes and Trigger Tubes

7.1. Switching Diodes

If the d.c. potential V_B applied across a glow-discharge diode lies between the maintaining potential V_m and the breakdown potential V_s, then the tube may be switched from the non-conducting to the glow discharge state by applying a positive pulse to the anode of sufficient magnitude to raise the potential above V_s. The amplitude and duration of the required pulse will depend on the 'over-voltage'/strike delay characteristic of the tube (see Chapter 5). Similarly the tube can be switched back again by lowering the potential below V_m, i.e. by applying a negative pulse to the anode, the duration being dependent on the deionisation time. Thus the glow-discharge diode is equivalent to a switch with a memory, and can be used as a logic circuit element, particularly in applications where very high speeds are not essential. It also has the advantage of giving a visual indication of its output.

In spite of competition from semi-conductor devices, the glow-discharge diode is finding increasing use in such applications. Initially small glow-discharge lamps were employed. By selection and ageing processes a reasonable reproducibility could be obtained, which allowed them to be used in simple non-critical circuits. More recently, diodes have been developed specifically for logic circuit applications. Based on the experience gained in the manufacture of high quality glow-discharge tubes, such as stabilisers and trigger tubes, the switching diodes are designed to give good stability of breakdown and maintaining potential yet are relatively cheap.

The required characteristics of a switching diode depends to some extent on the application. If the breakdown voltage is to be used as a reference (for example in timing circuits) then the

stability and reproducibility of V_s is of prime importance. However, in differential circuits (for example in counters) it is mainly the difference between the breakdown and maintaining potential which is of interest, and the spread in V_s need only be small compared with this difference; the higher the difference between V_s and V_m can be made, the greater can be the tolerance on the respective values. In some applications the tubes are used as visual indicators or to activate light sensitive elements, in which case provision must be made for adequate light output. There are, however, certain desirable parameters which are common to all applications. The tubes should be small, a subminiature execution which can be directly wired into the circuit being particularly suitable; they should have a long operational life and they should be cheap. Also, to ensure that the breakdown potential should depend as little as possible on the rate of change of the supply voltage firing the tube, the statistical delay should be small. This is particularly important for timing circuits, although it applies to a lesser extent to other applications.

7.2. Practical Design of Switching Diodes

Indicator tubes or sub-miniature low-current stabilisers such as the NE2 or the NT2 were initially employed for logic circuit applications. Such tubes have barium-coated cathodes and a neon-argon gas mixture, giving a maintaining potential ~ 60 V and a breakdown potential of ~ 80 V. The barium-coated cathode has two advantages; firstly the bulb is not covered with metal during the processing so that the glow is clearly visible, and secondly in normal ambient lighting the breakdown delay is negligible due to photo-electric priming. To ensure reliability on the latter point, it is wise to install a suitable light source near the tubes. Dimensionally the tubes are 6–7 mm diameter and 25–30 mm in length. The main disadvantage of these tubes is the small difference between the breakdown and maintaining potential, particularly as the spread of the respective values is relatively large. By using a different gas filling and/or changing the electrode spacing a higher breakdown potential can be obtained, and difference diodes based on this principle are manufactured. Examples are the XC25 and XC26 giving breakdown potentials of 130–160 V and 150–175 V respectively, with similar maintaining potentials to the NT2.

In the last few years the need for tubes with closer tolerances

has prompted the development of switching diodes with metal cathodes prepared by the sputtering process discussed in the previous chapter. The afore-mentioned advantages associated with the barium-coated cathode no longer apply and normally the glow is only visible through the ends of the tube where less material is deposited. Also some form of priming is required. Recently, however, a range of tubes has been described[1] (types ZA1000 to ZA1004), which overcome the main objections to a sputtered cathode tube. The tubes are all basically of the same design, but are given characteristics suitable for specific applications by modification of the gas filling and processing schedules. Thus ZA1000 and ZA1002 are switching diodes, ZA1001 is a 'relaxation' diode (see Section 7.3.1) suitable for timing circuits and ZA1003 and ZA1004 are indicators suitable as read-out diodes for transistor circuits. The tubes are approximately 6 mm diameter and 25 mm in length, and are shown in Plate 7.1. The electrode structure of the tubes consists of a central cathode rod of molybdenum surrounded by a concentric anode cylinder in the form of a mesh. When the cathode is sputtered, material deposits on to the tube wall, to act as a getter, in a pattern corresponding to the holes in the mesh. The screening effect of the mesh on the sputtered material is greater than that due to geometrical shadowing, because of the zig-zag path taken by the diffusing cathode particles, with the result that the glow is still clearly visible through the side of the bulb.* The geometry also helps in obtaining reproducibility of the breakdown potential, since slight deviation from concentricity does not materially affect the field at the cathode.

The statistical breakdown delay of the tubes is reduced to less than 1 msec for a 10 to 20 V overvoltage by the inclusion of a small quantity of tritium. The delay, however, is not reduced sufficiently for timing circuits where the tube acts as a relaxation oscillator at relatively low frequencies (see Section 7.3.1), and where it is essential for the tube to strike at the same voltage on each pulse. There is no problem at higher frequencies, since the tube does not have time to deionise completely between pulses, and there is sufficient residual charge to prime the gap. To overcome the low-frequency problem the tube designed for timing circuits, the ZA1001, has a small quantity of a heavy gas (Kr or Xe), which has a long deionisation time, added to the main gas. The resultant deionisation curve, expressed as residual charge against time, is shown in Fig. 7.1. On extinction of the glow the residual charge falls rapidly, the fall being governed by deionisation of the main

*The screening effect of a wire mesh on sputtered material is discussed in more detail in Section 4.4.4.

gas. This allows the tube to function satisfactorily at high frequencies. The presence of the admixed gas prolongs the final decay, however, providing sufficient priming for the tube to function at low frequencies. Thus, by combining a fast and slow deionisation process in one discharge, the tube is able to oscillate over a comparatively wide frequency range.

For application as read-out diodes for transistor circuits the gas filling is chosen to give less spread in the breakdown voltage than

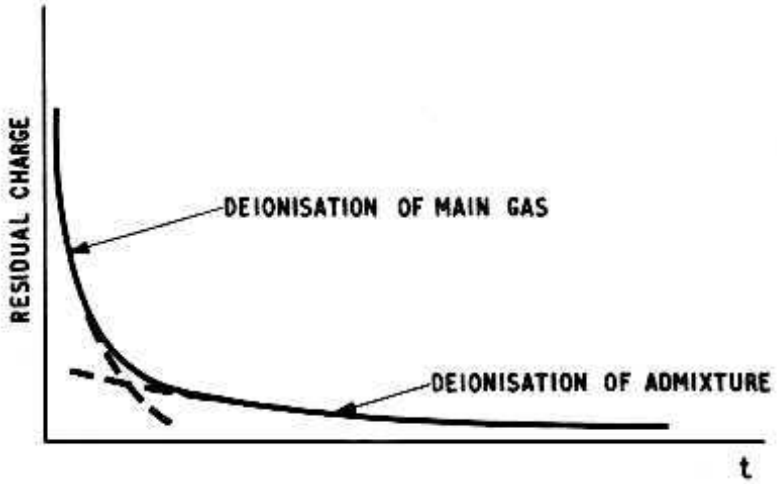


Fig. 7.1. The residual charge as a function of time in the ZA1001

the value of the signal supplied by the transistor. Examination of the breakdown characteristics of gas mixtures, Section 2.3, shows that neon plus a small percentage of argon (for example 0·1%) has a 'flat' characteristic at the Paschen minimum. This implies that the unavoidable small differences in pressure and electrode spacing of individual tubes have hardly any influence on the breakdown potential. Using such a mixture, the spread of V_s has been kept low enough for tube type ZA1003 to be controlled reliably by a 12 V transistor, and type ZA1004 by a 6 V transistor. The difference between breakdown and maintaining potential, however, for these tubes is small.

For tube types ZA1000 and ZA1002 used in differential circuits where a larger difference between V_s and V_m is required, pure neon is used with a filling pressure higher than that corresponding to the minimum of the Paschen curve. With this gas filling $V_s - V_m$ is approximately 60 V. For these tubes the light output is rather higher than for the neon-argon filled tubes, being 50 lux at a distance of 2 mm with a dissipation in the tube of 0·25 W, and they are very suitable for use in conjunction with cadmium sulphide cells.

A diode with a sputtered metal cathode suitable for timing circuits has been described by Wright[2] in which a separate priming or keep alive electrode is incorporated. The stability of the diode

as an oscillator was 0·5% at frequencies above 250 c/s, and at low frequencies of ~ 25 c/s was about 5%.

7.3. Applications of Switching Diodes

7.3.1. SELF-QUENCHING AND RELAXATION OSCILLATIONS

Before discussing circuit applications of the switching diode it is necessary to understand the phenomena of self-quenching and relaxation oscillations. If the circuit of Fig. 7.2 is considered, then

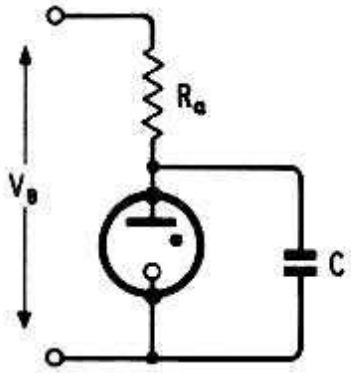


Fig. 7.2. Basic circuit for relaxation oscillations

on applying V_B the capacitor C is charged via resistor R_a until the breakdown voltage of the tube is reached. The capacitor will then discharge through the diode. Since the tube has an inductive component, a resonant circuit is formed by the inductive component and the capacitor. If the damping by R_a and R_g, the resistive component of the tube impedance, is less than critical, the voltage drop from V_s to V_m will produce oscillation about V_m. It is thus possible for the voltage to swing below V_m and the discharge to extinguish or self-quench. In terms of the phenomena in the tube, breakdown of the discharge is followed by a high current which will have passed its peak by the time that the capacitor has discharged sufficiently for the anode voltage to fall to V_m. Consequently, an abnormally large positive-ion cloud will have formed round the anode, and as this cloud moves away the anode can descend well below V_m in an over-shoot. At the end of the overshoot the capacitor will recharge through R_a raising the anode potential up to V_m, but not before the process of deionisation has gone far enough to cause extinction. Thus, on the V-i characteristic, Fig. 7.3, the dotted line represents the excursion of the current and voltage during a capacitor discharge. Whether or not the discharge self-quenches depends on the values of C and R_a. If either is small the circuit is damped, the anode voltage rises to V_m before appreciable deionisation, and the discharge does not extinguish. However,

due to the complex equivalent impedance of the discharge it is not possible to calculate the minimum values required to prevent self-quenching, and critical values of R_a and C can only be determined experimentally. Fig. 7.4 shows the critical value of R_a as a function of C for the ZA1000 at three different values of supply voltage V_B.

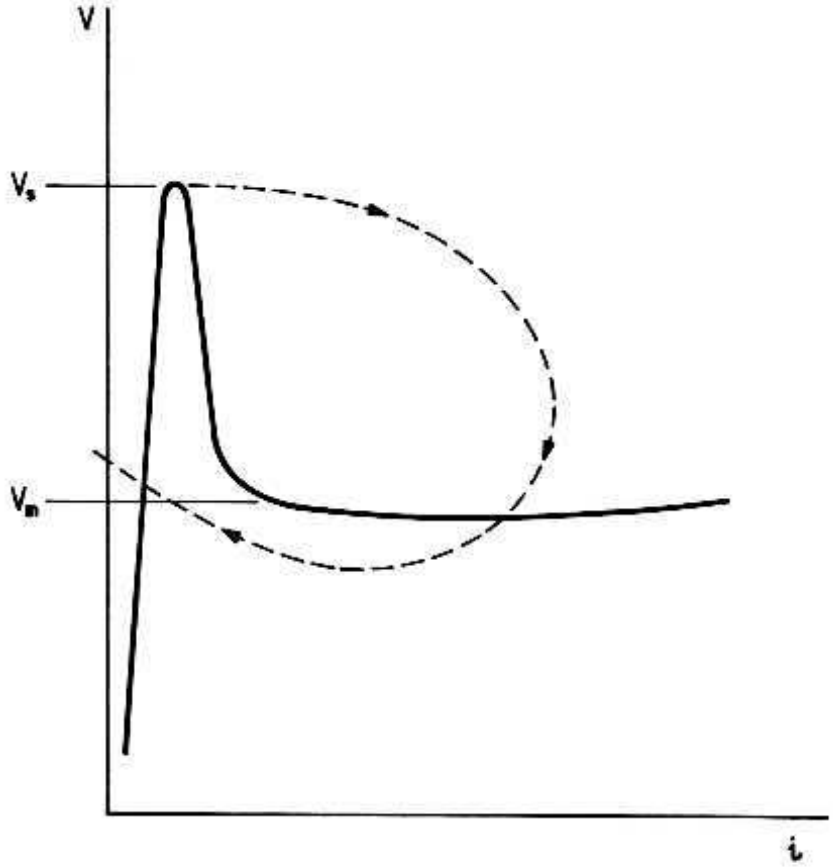


Fig. 7.3. V-i characteristic of a switching diode showing the voltage/current excursion (dotted curve) during a capacitor discharge in the circuit of Fig. 7.3

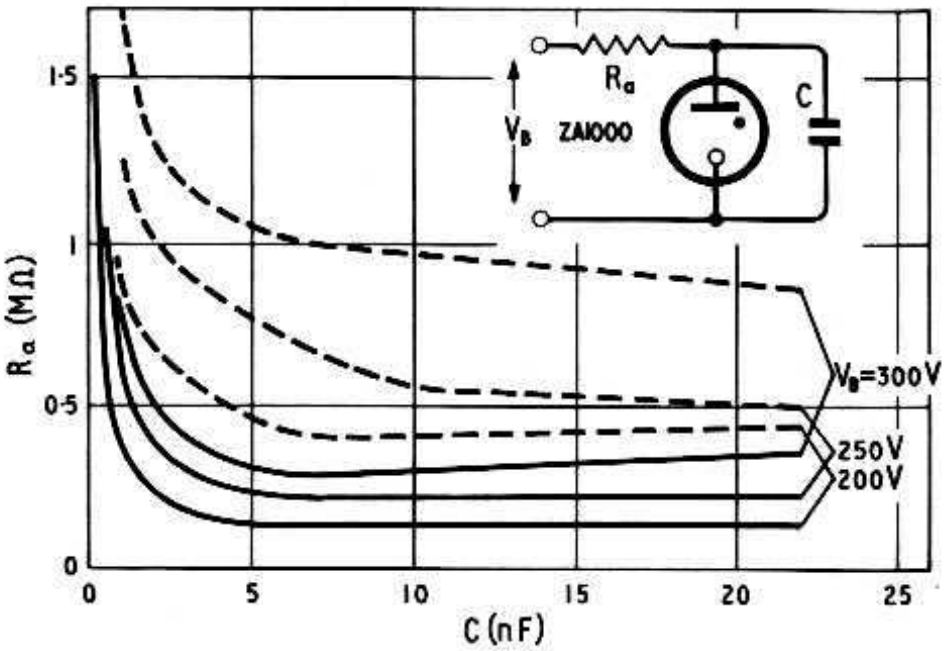


Fig. 7.4. Critical value of anode resistor in the circuit of Fig. 7.2, below which the tube remains ignited, as a function of the capacitor value, and with the supply voltage as parameter. If the tube remains ignited the anode resistor must be increased to a value exceeding the corresponding dotted curve to render the tube self-quenching

Clearly if V_B remains above the breakdown potential the capacitor will recharge to V_s and the process will be repeated giving rise to relaxation oscillations. The characteristic sawtooth waveform of the oscillations is shown in Fig. 7.5. Assuming that the voltage does not drop appreciably below V_m and that the discharge

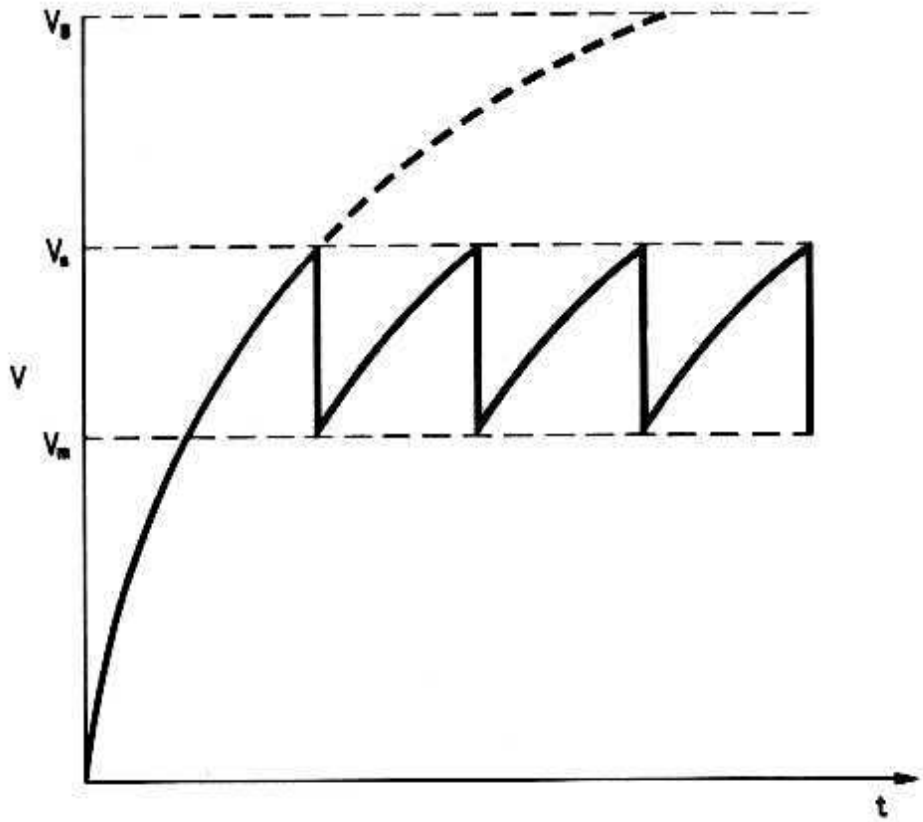


Fig. 7.5. Characteristic sawtooth waveform of the relaxation oscillation obtained with the circuit of Fig. 7.2

time is short in comparison with the charging time, then the periodic time of the oscillations is given by

$$t = CR_a \log_e \frac{(V_B - V_m)}{(V_B - V_s)} \tag{7.1}$$

In practice the frequency will be slightly lower than $1/t$ due to the finite discharge time and the excursion below V_m, but Equation 7.1 is useful for circuit design purposes.

7.3.2. TIMERS AND TONE GENERATORS

The relaxation oscillations of a switching diode can be employed in timers. A simple timing circuit giving a positive pulse output on each cycle is illustrated in Fig. 7.6. The accuracy falls off at lower frequencies, nevertheless a timer of several minutes can be obtained

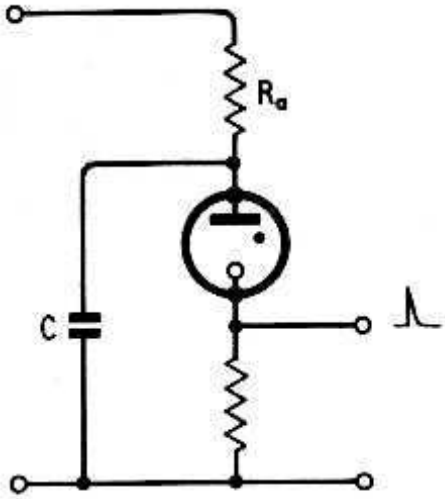
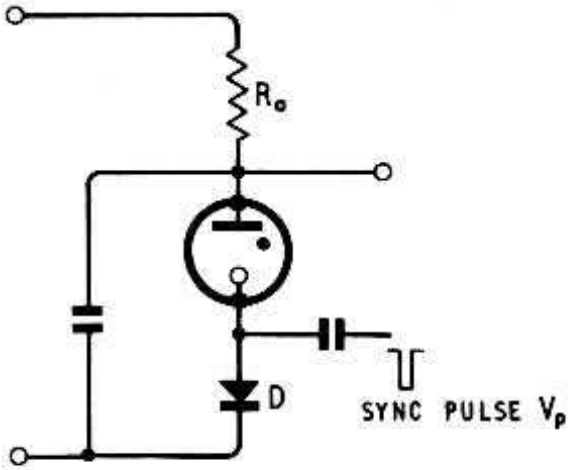
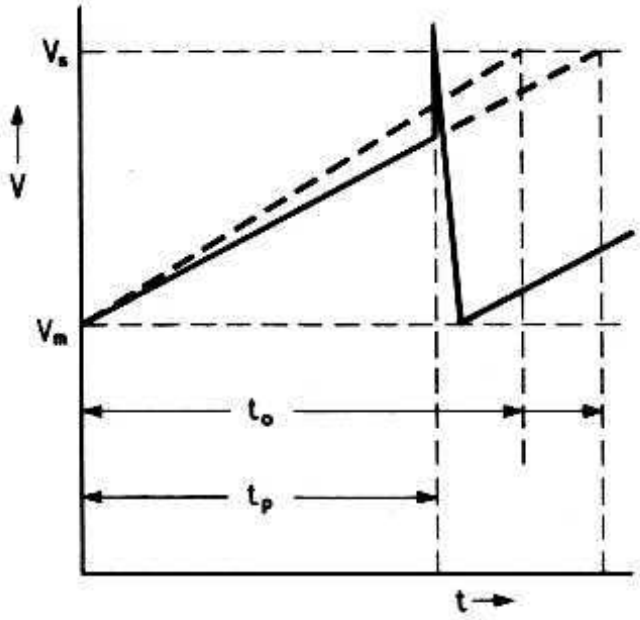


Fig. 7.6. Simple timing circuit giving a positive pulse output



(a)



(b)

Fig. 7.7. A synchronised relaxation oscillator for use as a tone generator. (a) basic circuit, (b) voltage waveform at the anode

from the circuit providing R_a, of the required high value, is suitably screened.

By taking the output from the anode the circuit can be considered as a sawtooth waveform generator, and has found use as a tone generator in electronic organs.[3, 4] The sawtooth signal has the advantage of a regular distribution of even and odd harmonics, and with the aid of electrical filters the tone 'colour' can be varied to match that of a mechanical instrument. In the early designs[3] NE2 tubes were used and rather complicated circuits were required to overcome their limitations. In the more recent work described by v. de. Kerckhoff[4] the ZA1001 was employed. Even so it was necessary to synchronise the oscillations, since variation in the supply voltage, RC components and the tube parameters affected the stability. A suitable circuit is shown in Fig. 7.7 (a) together with the voltage waveform Fig. 7.7 (b). Without synchronisation the glow discharge would occur after a time t_o, which would depend on the slope of the charging curve and the values of V_s and V_m. By applying a synchronising pulse of the order of $\frac{2}{5}(V_s - V_m)$ at a time $t_p < t_o$, variations of over $\pm 10\%$ can be tolerated on the various parameters. An electronic organ utilising glow-discharge diodes as tone generators is commercially available.

7.3.3. COUNTING CIRCUITS

Several circuits based on the connection of two or more switching diodes in parallel, with a common anode or cathode resistor, have been described in the literature. The simple flip-flop binary counter[5] (circuit diagram of Fig. 7.8) illustrates the principle of such circuits.

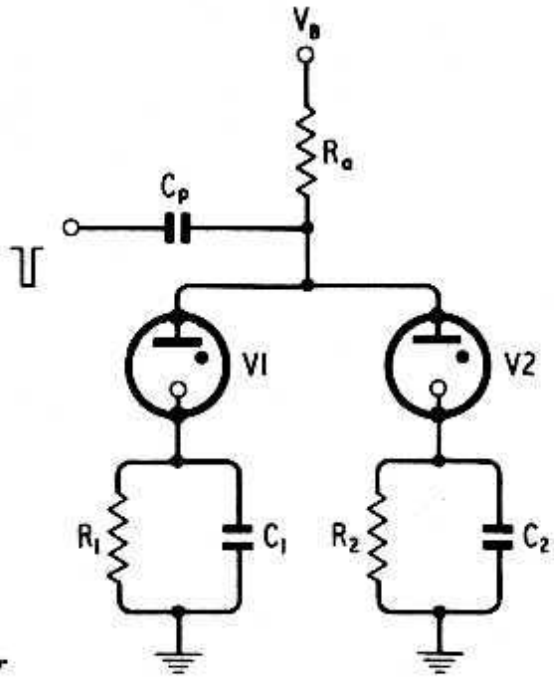


Fig. 7.8. Common-anode binary counter

The supply voltage V_B is above the breakdown potential of the tubes so that, upon switching on, one of the tubes, say V_1, will ignite and conduct with cathode current i. The current charges up C_1 to iR_1, the voltage at the anodes rising from V_m to a final value of V_m+iR_1. Providing this is less than V_s, tube V_2 cannot ignite. Applying a negative pulse to the anodes extinguishes V_1, and if it is of sufficient duration allows the tube to deionise. Because of R_1C_1, however, the cathode potential does not fall to zero immediately, so that at the end of the pulse when the anode voltage rises towards V_s a larger voltage will appear across V_2 than across V_1. When the anode potential reaches V_s, tube V_2 will breakdown and, in the same way as before, prevent V_1 igniting. By a second pulse the circuit can be switched back to its original state. The output can be taken from the cathode of either tube, and several flip-flops can be cascaded to provide a binary scaler.

Of more interest is the possibility of connecting several tubes in parallel, and, by suitably arranging that on extinction of one tube a particular second tube is ignited, produce a ring counter. A simple circuit for a ring counter has been described by Flood and Warman[6] and is shown in Fig. 7.9. If initially V_1 is conducting, the capacitor

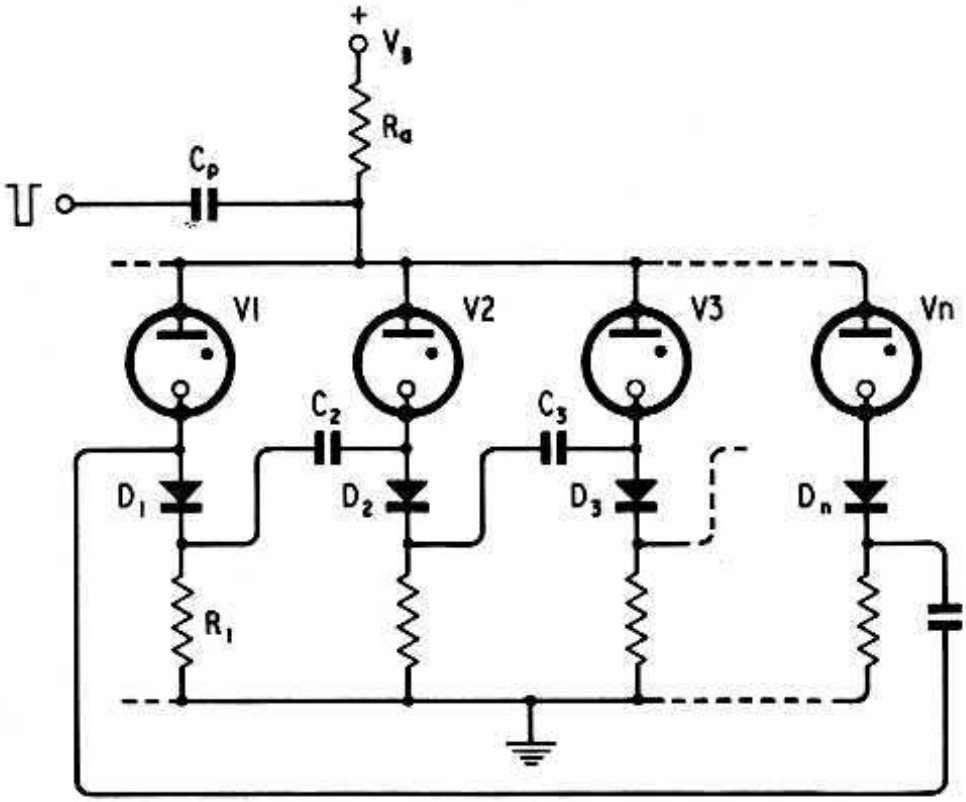


Fig. 7.9. A ring counting circuit with cathode coupling. Flood and Warman[6]

C_2 will charge up to iR_1. A negative pulse applied to C_p extinguishes V_1 and its cathode potential falls to zero. The capacitor C_2 retains its charge so that the cathode V_2 becomes negative and the rectifier D_2 is biased off. Thus the largest potential appears

across V_2, and at the end of the pulse, when the anode rises to V_B, V_2 will ignite preferentially. The pulse has thus transferred the conducting state from V_1 to V_2, and subsequent pulses will transfer it further along the chain. The maximum counting speed is of the order of 100 c/s. Alternatively the tubes may be interconnected by cadmium sulphide photoconductive cells according to the circuit of Fig. 7.10, where the light output of the conducting tube is used

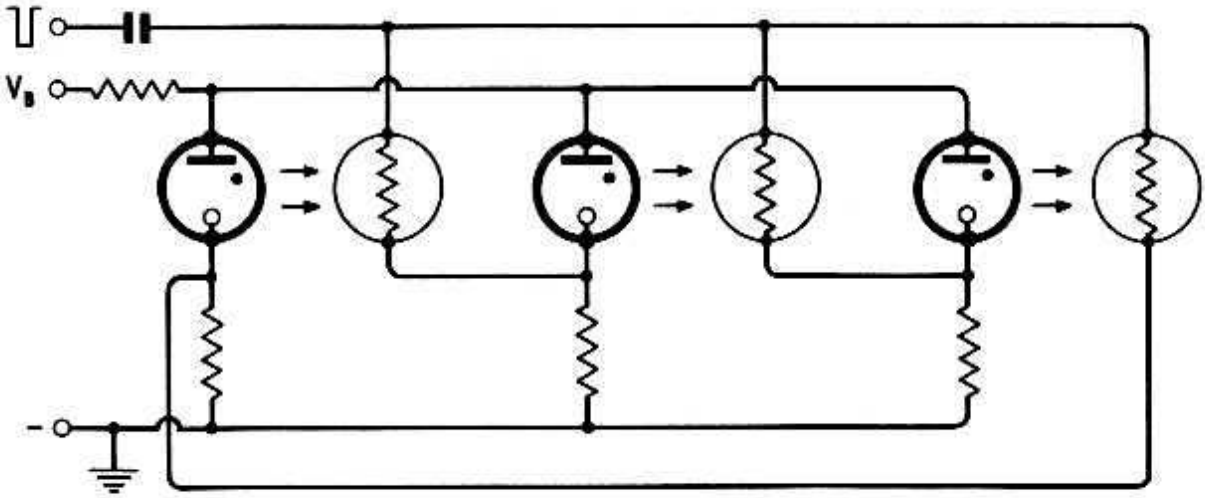


Fig. 7.10. A ring counting circuit with cadmium sulphide cell coupling

to bias the next tube in the ring via the CdS cell. The counting speed is limited by the inherent inertia of the CdS cell to a few counts per second. It has the advantage however of low cost and simplicity.

7.3.4. LOGIC AND MEMORY CIRCUITS

Switching diodes have been used in a number of logic applications, such as gates, shift registers, and memory matrices. They are attractive since they are cheap and simple to use, but they are limited to relatively slow speed applications.

The possibilities of using difference diodes as gates have been discussed by Flood and Warman.[6] The basic circuits are shown in Fig. 7.11. In Fig. 7.11 (a) which shows an 'OR' gate, a voltage pulse V_p, which exceeds V_s, applied to either input terminal will cause a tube to ignite. The voltage which then appears at the output is $(V_p - V_m)$. Providing this is less than the inverse breakdown of the other tube the input terminals remain isolated from each other. The 'AND' gate of Fig. 7.11 (b) operates on the pulse-plus-bias principle. The signal voltage applied to each input terminal must be less than V_s, but when a pulse is applied after the bias, the combined signal must be greater than V_s. Thus the tube will only

ignite when both signals are present. A similar circuit can be used as a 'NOT' gate, Fig. 7.11 (c). If a positive pulse, greater than V_s, is applied to input terminal A the tube will ignite and a voltage $V_p - V_m$ is delivered to the output. However, if a positive bias V_b is applied to terminal B such that $V_p - V_b < V_s$ breakdown cannot occur and no output will be seen. In a pulse-plus-bias circuit the bias must be less than V_m if the tube is to extinguish at the end of the

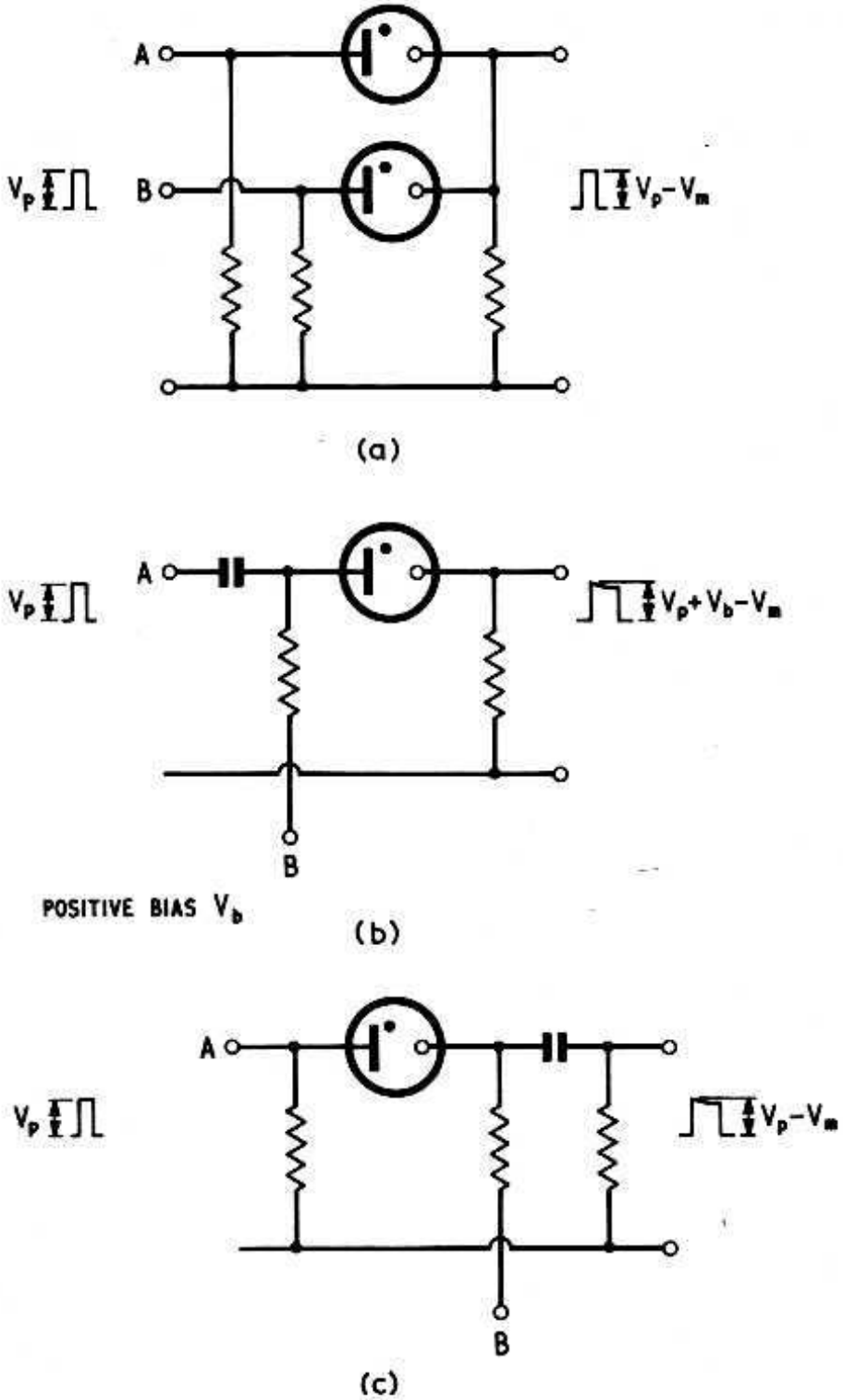
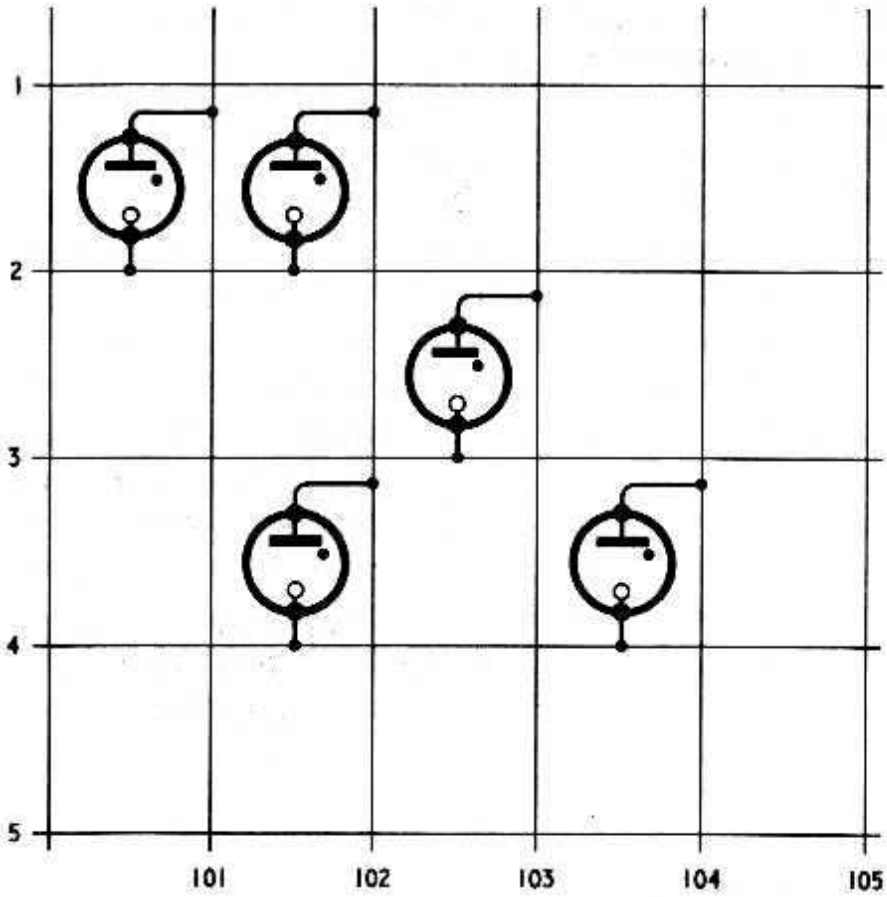
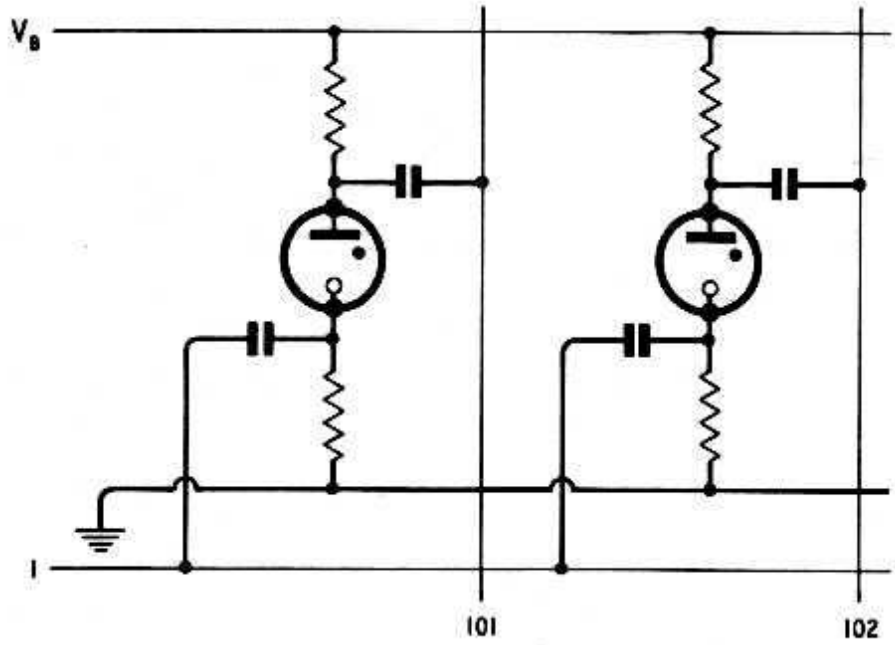


Fig. 7.11. Switching diodes used as logic gates
(a) basic circuit for an 'OR' gate
(b) basic circuit for an 'AND' gate
(c) basic circuit for a 'NOT' gate



(a)



(b)

Fig. 7.12. Switching diodes as cross wire connectors in a logic matrix. (a) circuit requiring a separate register and memory, (b) circuit in which register and memory are incorporated in the switch

pulse. More complex gating systems can be developed by combining the basic circuits.

Raphael and Robinson[7] have suggested the use of switching diodes in a logic matrix. The tubes are connected between cross wires as shown in Fig. 7.12 (a). If at any instant the potential of a vertical wire is raised to V and that on one of the horizontal wires lowered to $-V$ such that $2V > V_s$ then the diode connecting the two appropriate wires will ignite. Photoconductive cells can be employed to register the output. Alternatively the tubes can be connected to a d.c. supply with a voltage between V_m and V_s, and pulses applied to the horizontal and vertical wires of such an amplitude that only when both pulses are applied at the same time does a tube fire, Fig. 7.12(b). Thus a memory can be built into the matrix. Erasure is effected by lowering the h.t. supply.

Other circuits such as shift register, and trigger circuits, have been reviewed by Neale.[5] The use of diodes as indicators is discussed in Chapter 9.

7.4 Trigger Tubes

The chief problem associated with the use of switching diodes in logic circuits is providing isolation between input and output terminals, since they are connected to the same pair of electrodes. This problem does not arise with the three- or four-electrode tubes, known as trigger tubes,* which will be described in the following sections. Further, the introduction of an extra electrode (the 'trigger') increases the versatility of the glow discharge tube, for added to the switching properties is the possibility of amplification.

The potentialities of the trigger tube have been exploited for some time, and the development of both tubes and circuits had firmly established them in logic and control applications, where high speed is not essential, before switching diodes became available. Indeed, often a trigger tube will be used for purposes where a diode may offer a simpler and less expensive solution. The first trigger tube, the 313A, was described[8] in 1936 and was developed from stabiliser tubes of that period. It was virtually a double-cathode tube, both trigger and cathode being identical in geometry and coated with a barium layer. A discharge between the two cathodes (breakdown potential 70 V) lowered the breakdown potential between anode and cathode (breakdown potential

*The basic trigger tube has three electrodes, but extra electrodes are often included for auxiliary interactions. In particular, it is common to incorporate a priming electrode in tubes with sputtered metal cathodes to reduce the statistical delay.

approximately 175 V) and thus the tube could be used as a relay. Similar tubes having large area barium-coated cathodes, although not necessarily of the same geometry, were developed during the following 10–15 years by a number of manufacturers. These early tubes were employed in control circuits, replacing small thyratrons where elimination of the cathode heater was advantageous; the 313A was specifically developed for telephone switching.

The introduction of stabilisers with pure metal cathodes, cleaned by sputtering, was soon followed by the design of trigger tubes based on the same processing techniques. Considerable improvement on the stability and reproducibility of the characteristics was obtained by this method, allowing the tubes to be used in a much wider range of applications. Miniature tubes have been developed with close-tolerance trigger breakdown characteristics for industrial control applications, and inexpensive subminiature tubes for slow-speed logic circuits; Neale[5] lists approximately 50 types currently available in the U.K.

Because of the high voltage drop across the electrodes, of the order of 100 V compared with 10 V for thyratrons, trigger tubes are limited to relatively low-current applications, in the region of a few tens of milliamps mean.

7.5. The Principle of Operation of Trigger Tubes

The trigger tube can be considered as two diodes capable of interacting, one formed by the trigger–cathode gap, and the other by the anode–cathode gap. Each gap has its own breakdown and maintaining potential, but the breakdown of the anode–cathode gap can be influenced by the presence of a discharge in the trigger–cathode gap.

In modern trigger tubes the cathode is normally in the form of a plate, and the anode and trigger are wires or rods, often shaped or screened to ensure that breakdown occurs at one particular point. Usually the trigger is positioned to give the minimum breakdown potential in the cathode–trigger gap, whilst the anode is mounted on the same side of the cathode but further away to increase the breakdown voltage of the anode–cathode gap to well above the minimum.

The operation of the tube is best considered by referring to Fig. 7.13 (a) (the 'arrow' is the conventional symbol for the trigger). V_B is chosen to be between the breakdown and the maintaining voltage of the anode–cathode gap in the absence of a trigger–cathode discharge. To switch the tube the voltage on the trigger,

V_T, must first be raised until the breakdown potential, V_{Ts}, of the trigger–cathode gap is reached. A discharge will then pass between the trigger and cathode, the current being governed by the trigger resistance R_T. The discharge provides ions and electrons in the anode–cathode gap and, if the ionisation is sufficient, produces

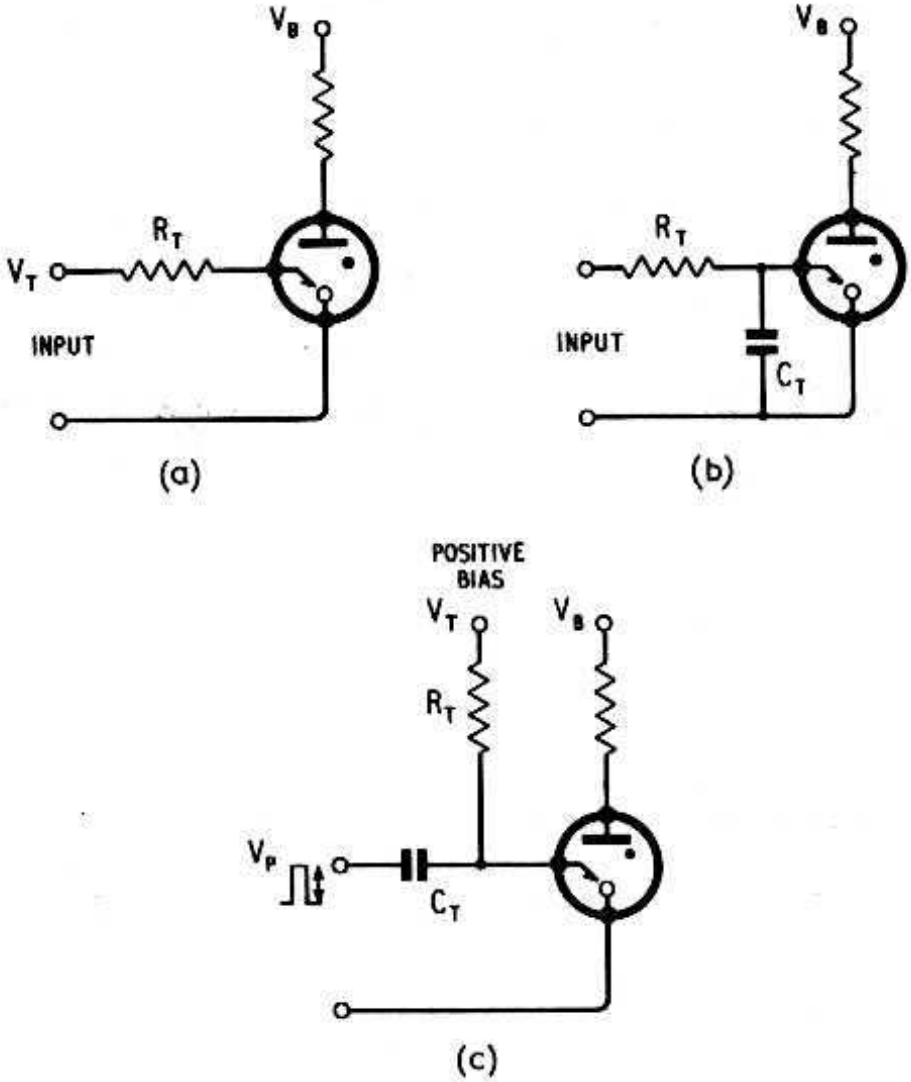


Fig. 7.13. Basic trigger-tube circuits (a) current triggering (b) capacitor triggering (c) Pulse-plus-bias triggering

space charge distortion in the cathode region. Providing the trigger discharge is in the right location the space-charge distortion lowers the breakdown potential V_{As} of the anode–cathode gap; this has been discussed in Section 2.5. The relation between anode–cathode breakdown and the trigger current is illustrated in Fig. 7.14, and is known as the transfer characteristic. If the trigger current is sufficiently large [at a value $i_{T(\text{trans})}$], V_{As} will be reduced below V_B and an anode–cathode discharge will be initiated. Thus the trigger conditions for operating the tube are

$$V_T > V_{Ts}$$

$$R_T < (V_{Ts} - V_{Tm})/i_{T(\text{trans})}$$

The trigger current required is quite small (of the order of 100 μA) and since the anode current can be 10 mA or more, current amplification is obtained.

As with all cold cathode tubes, initiation of the discharge is subject to time delays (see Chapter 5). These are the statistical and

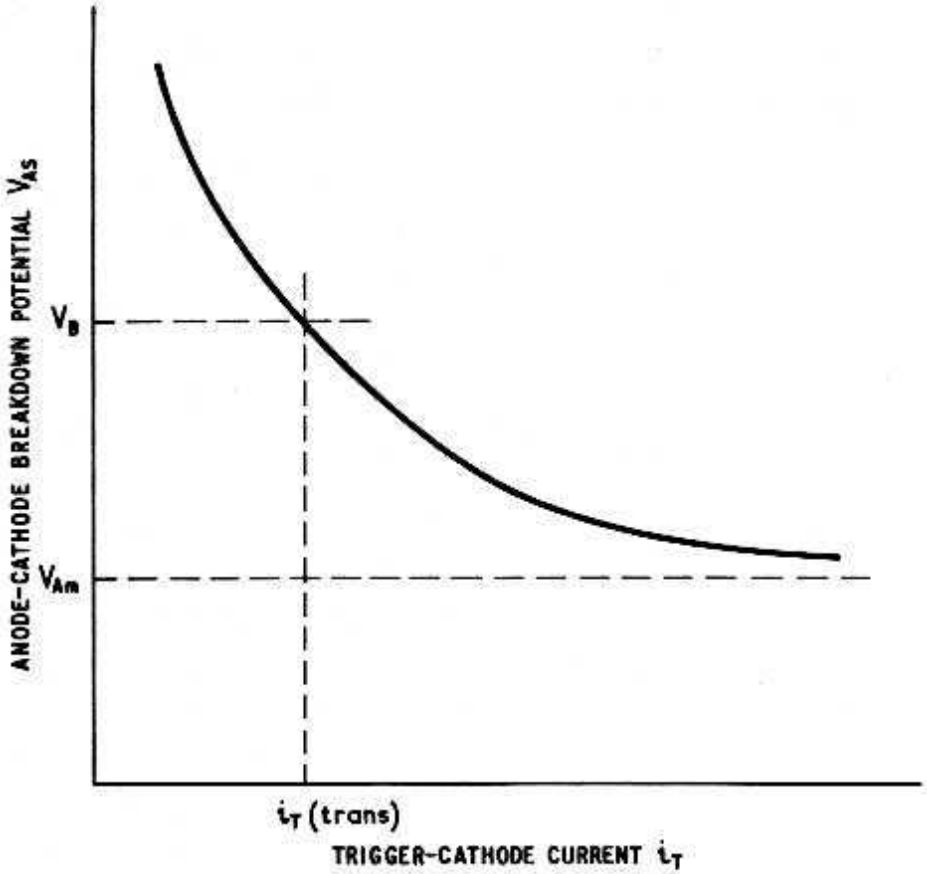


Fig. 7.14. Transfer characteristic of a trigger tube for current triggering

formative lags associated with the trigger–cathode breakdown, and also the formative lag in the establishment of the anode–cathode discharge. The latter is often referred to as the transfer time, although some authors apply the term to the sum of all three lags. These time delays dictate the length of time the trigger input has to be applied in order to initiate an anode–cathode breakdown, for the input must not be removed before the main discharge has reached a self-sustained state. For most applications the time delays should be kept to a minimum, and in particular the statistical delay must not be appreciably greater than the formative lags. Thus a fairly high level of priming is required in trigger tubes, much higher than is acceptable in stabilisers. In tubes having coated cathodes, the ambient light may well be sufficient, but for sputtered metal-cathode tubes a priming current must be drawn to an auxiliary electrode; an impractical level of

radiation would be required if radioactive priming were employed. The priming current is of the order of microamps and the electrode is therefore normally connected to a suitable supply via a high resistor $\simeq 10\,\mathrm{M\Omega}$.

The influence of the trigger–cathode current is restricted to the beginning of the anode–cathode discharge. Thereafter it has no influence, and the discharge can only be interrupted by lowering the anode voltage below V_{Am}. The trigger discharge must also be extinguished by lowering the trigger voltage below V_{Tm}.

7.6. Triggering Circuits and their Operation

Basically there are three methods of firing a trigger tube, and these are represented in the circuit diagrams of Fig. 7.13. The circuit of Fig. 7.13 (a) has already been discussed, and is known as current triggering. In the circuit of Fig. 7.13(b), a capacitor is connected between trigger and cathode. The trigger current charges up the capacitor until the breakdown voltage V_{Ts} is reached, when it discharges into the trigger–cathode gap to provide the necessary current. By this means the input impedance can be increased by some two or three orders of magnitude, though at the expense of the delay due to the charging time. The method is known as capacitor triggering. Since the tube itself will have a self-capacity between trigger and cathode, circuits 7.13 (a) and 7.13 (b) differ only in the value of the trigger resistance and capacitance. In most practical circuits the trigger capacitor will be large compared with the self-capacity of the tube. The third method of triggering, Fig. 7.13 (c), is to apply a pulse of sufficient amplitude and duration to the trigger via a capacitor. The trigger can be pre-biased to reduce the pulse requirements, and the method is known as pulse-plus-bias triggering.

Alternative methods of triggering are possible, such as applying a negative signal to the cathode or, in tubes designed for the purpose, a negative pulse to the trigger. In one tube, described by Tosswill,[9] a negative pulse is applied to an auxiliary cathode, with a remarkable increase in sensitivity compared with normal operation. The methods outlined above, however, are much more common. Capacitor triggering is widely used in control circuits where the tube is used as a relay or timer, whereas pulse-plus-bias is extensively used for logic circuits. These two methods are therefore described in more detail in the following sections.

7.6.1. CAPACITOR TRIGGERING

The author[10] has made a study of capacitor triggering with a high impedance input. The general principles, however, are applicable to lower impedance circuits. Consider first the trigger-cathode breakdown; as V_T is raised, current passes through R_T to charge C_T, Fig. 7.13 (b). Also there will be a pre-breakdown current flowing through the tube which will increase as the Townsend region is reached (to the order of 1 μA just prior to breakdown). Unless the total current i_T is sufficient, the pre-breakdown current will prevent the capacitor charging up to the breakdown potential. Therefore not only must V_T be above V_{Ts} but also the current must be above a certain critical level to cause a trigger–cathode breakdown. At breakdown the capacitor discharges into the gap and relaxation oscillations are observed. The value of i_T for onset of

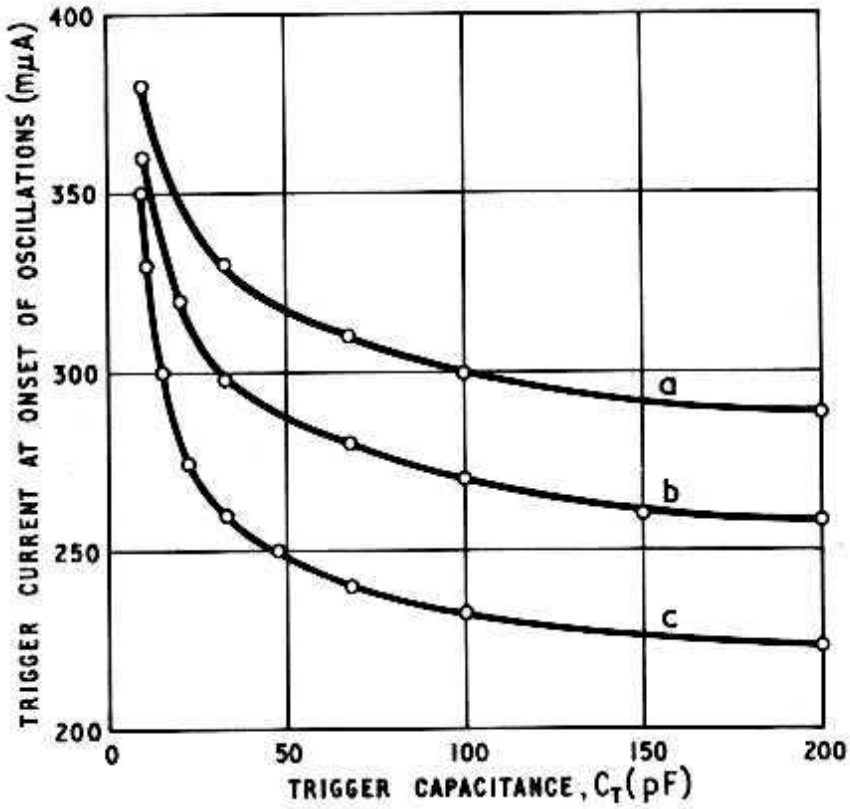


Fig. 7.15. Typical trigger current/capacitance curves for capacitor triggering with a high input impedance, with the priming current as parameter. Weston[10]
(a) $i_p = 87\mu A$
(b) $i_p = 43\mu A$
(c) $i_p = 1\mu A$

oscillations depends on R_T, C_T, and, rather surprisingly, on the priming current i_p. Fig. 7.15 shows some typical curves of i_T against C_T for an experimental tube. It was found that the characteristic obtained could be expressed by the equation:

$$C_T(i_T - i_c) = \text{constant} \tag{7.2}$$

where i_c, the asymptotic value of i_T for large capacitor values, depended on the priming current and geometry. The constant was a function of the trigger resistor. The results were in agreement with a simple theory based on consideration of the condition for stability during small fluctuations of current and voltage in the trigger–cathode gap.

Examining the anode–cathode breakdown as a function of the trigger–cathode discharge showed that a large reduction in anode breakdown voltage occurred in the region where the trigger current caused onset of oscillations. Thereafter, increasing the mean trigger current had only a small effect particularly at higher capacitor values (above 200 *p*F). This is explained by the fact that generally the anode–cathode gap ignites at the first capacitor discharge, and that increasing the current increases the frequency and not the

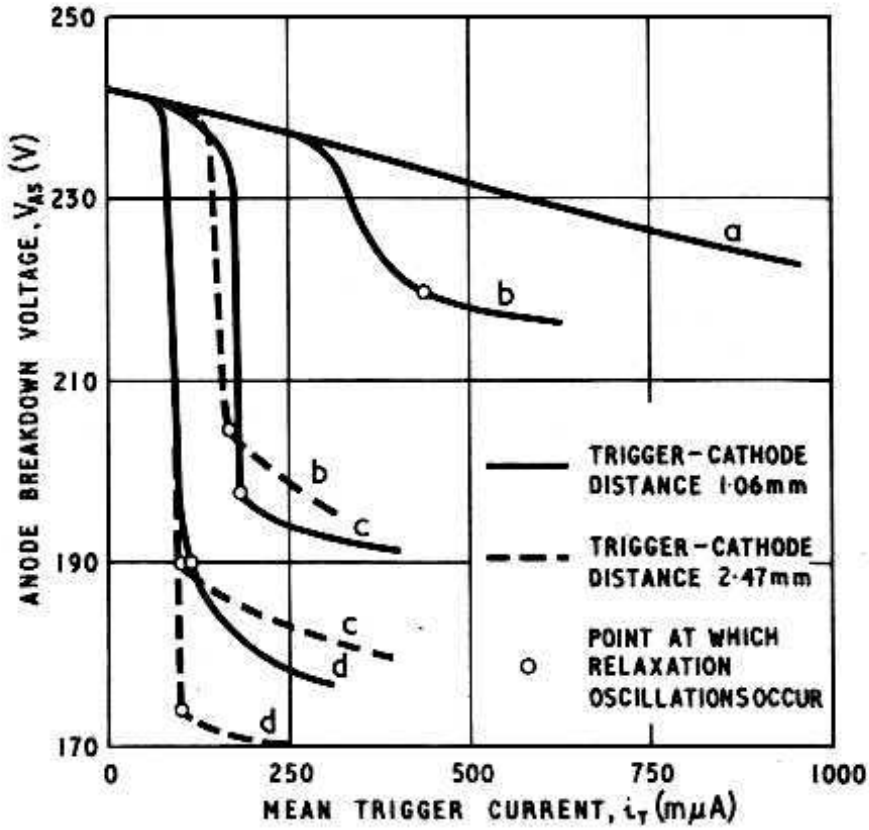


Fig. 7.16. Transfer characteristics of a trigger tube, with variable trigger cathode spacing, for capacitor triggering. Weston[10]

(a) $C_T = 0$
(b) $C_T = 10$pF
(c) $C_T = 22$pF
(d) $C_T = 48$pF

amplitude of the oscillations. Some transfer characteristics for an experimental tube are shown in Fig. 7.16. In the absence of a capacitor the curve (*a*) would be similar to Fig. 7.14 if the current range were extended.

Thus, providing the capacitor value is large enough to lower the breakdown potential below the supply voltage V_B, then the asymptotic value of i_T in Fig. 7.15 can be considered as a measure of the trigger sensitivity in terms of the minimum current required from the input circuit. This, however, may not be of practical use if it limits the repetition rate of the trigger tube. Of more interest is the minimum capacitance needed to lower the anode–cathode breakdown voltage below V_B. A transfer characteristic can then be obtained by plotting the breakdown potential V_{As} against trigger capacitance. In practice this produces a smooth curve of similar shape to Fig. 7.14 with C_T as abscissa.

In some applications, timers for example, a capacitance far in excess of the minimum value will be used. It is then advisable to connect a limiting resistor in series with the capacitor to prevent the discharge from causing damage to the tube.

7.6.2. PULSE-PLUS-BIAS TRIGGERING

In the pulse-plus-bias circuit, Fig. 7.13 (c), not only must the pulse amplitude V_p plus the bias voltage V_T be greater than the breakdown potential of the trigger–cathode gap, but they must provide sufficient transfer current for a long enough period to ensure anode–cathode breakdown. This imposes limitations on the minimum pulse duration, and also on the values of C_T and R_T. The value of the trigger resistor is limited, as before, by the pre-breakdown current in the tube. However, since V_T can be fractionally below the breakdown potential, the pre-breakdown current before application of the pulse is reduced considerably (by a factor of 100 or more) and higher values of R_T are possible. V_T should not be too low, since this would involve large pulses which, being differentiated, produce negative spikes at the end of the pulse with the possibility of spurious breakdown or negative current. It is generally recommended that V_T should not be less than the maintaining potential of the trigger–cathode gap. The upper limit V_T is set by the stability and reproducibility of the trigger–cathode static breakdown potential. The value of C_T dictates the maximum current flowing for a given pulse amplitude.[11] Although a steep leading pulse-edge may be applied to C_T the current in the tube can only build up slowly due to the inductive character of the impedance. At the same time C_T is discharging so that the voltage on the trigger is decreasing with time. The two effects give a current waveform of the shape shown in Fig. 7.17. This figure shows the

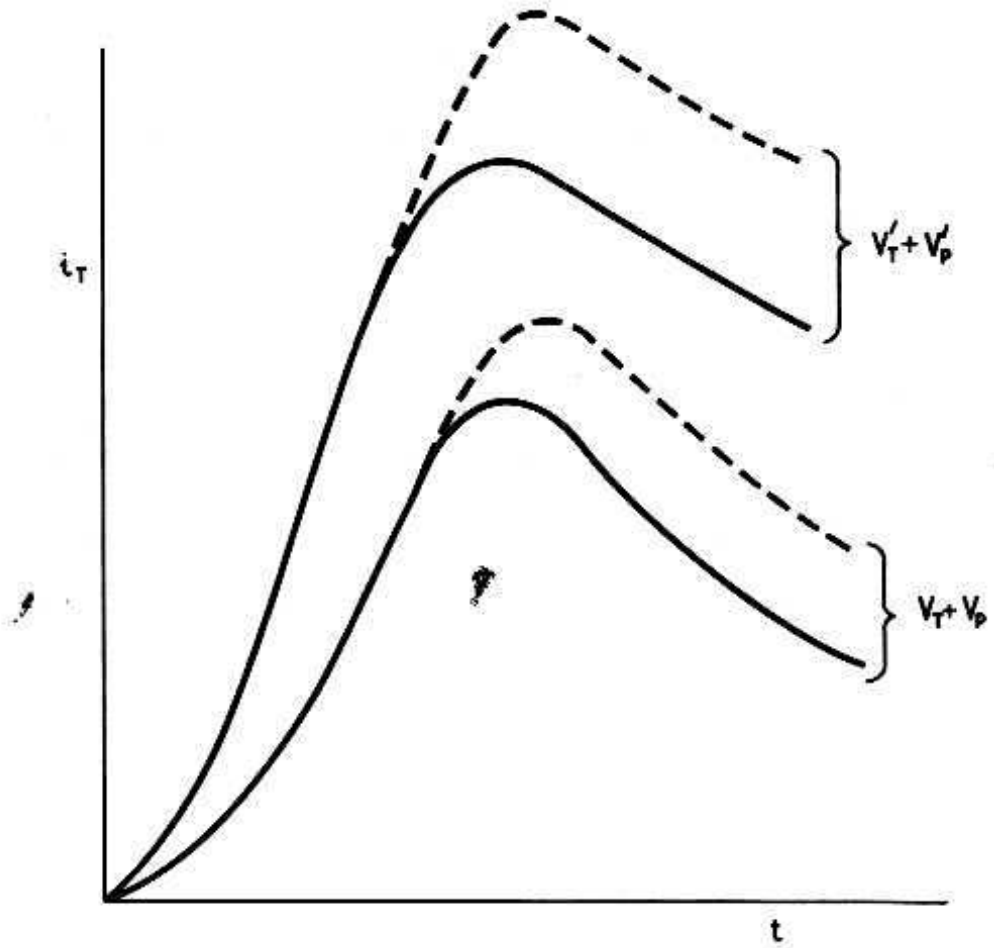


Fig. 7.17. Current waveform for the trigger–cathode gap for pulse-plus-bias triggering. $V_T^1 + V_p^1 > V_T + V_p$, and the dotted curve is for a higher capacitor value

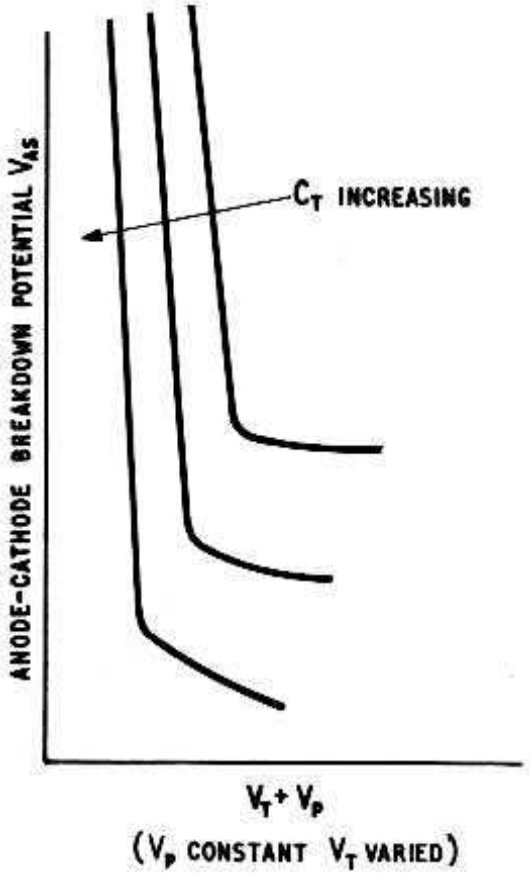


Fig. 7.18. Anode–cathode breakdown potential as a function of the pulse-plus-bias potential showing the effect of the input capacitor

effect of C_T, and also of $V_T + V_p$, on the current-time characteristics. Since the anode breakdown as a function of the peak current is similar to that of Fig. 7.14, smaller capacitors require higher pulses to fire the tube.

Plotting the anode–cathode breakdown as a function of the input voltage, $V_T + V_p$, for different capacitor values (Fig. 7.18) gives curves of a similar character to those obtained for capacitor triggering (Fig. 7.16). Thus, in this method of triggering it is also possible to give a transfer characteristic of V_{As} as a function of C_T. Further, providing R_T is large and the impedance of the pulse-producing circuit is low, the values of C_T pertaining to a given value of V_{As} do not differ greatly between the two triggering circuits.

The above considerations apply to a pulse duration which is long compared with the delay times. Clearly, if the pulse duration is less than the current build-up time (time to reach the peak, Fig. 7.17) the peak of the trigger current will be reduced. However, the formative lag is decreased if the potential is raised (Chapter 5), so that the larger the applied pulse the shorter can be its duration to give the same peak current. A curve of 'overvoltage' (volts above the static breakdown V_{Ts}) against pulse width for triggering the Z803 U is given in Fig. 7.19. The curve can be expressed as a roughly hyperbolic function $(V_T + V_p - V_{Ts})\, t_p$ = constant.

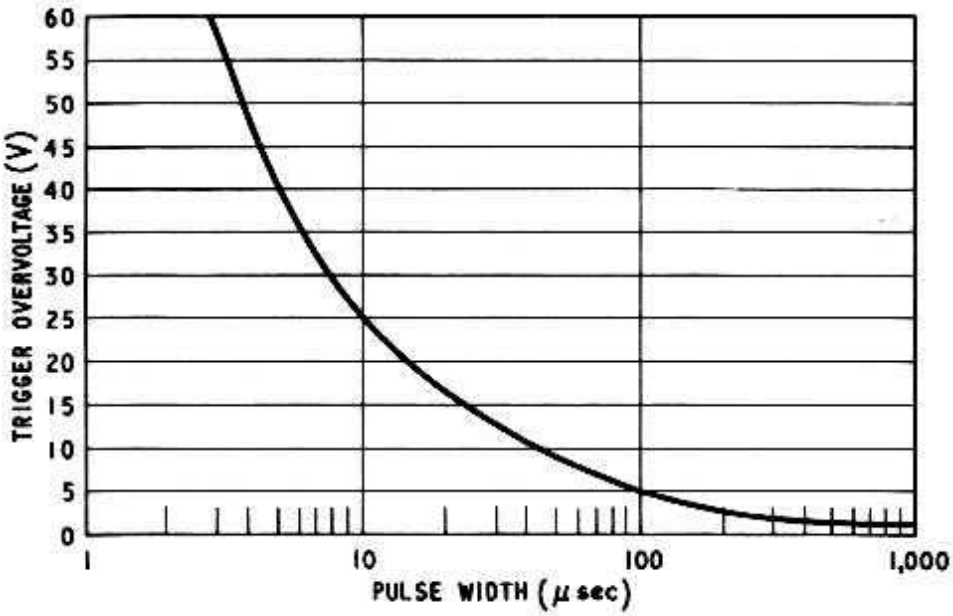


Fig. 7.19. The pulse overvoltage required to trigger a Z803U as a function of the pulse duration. Gimson and Crowther[12]

One disadvantage in the pulse-plus-bias circuit is that the bias must be applied some time before the pulse to allow time for the capacitor to charge. This can impose a limitation in some logic circuits.

7.7. Extinguishing the Discharge in Trigger Tubes

The discharge can be extinguished using the methods described in the sections on switching diodes, namely: interrupting the voltage supply, pulsing below the maintaining potential, self-quenching with a parallel condenser, or, where more than one tube is involved, making use of a common load resistor. In trigger tubes, however,

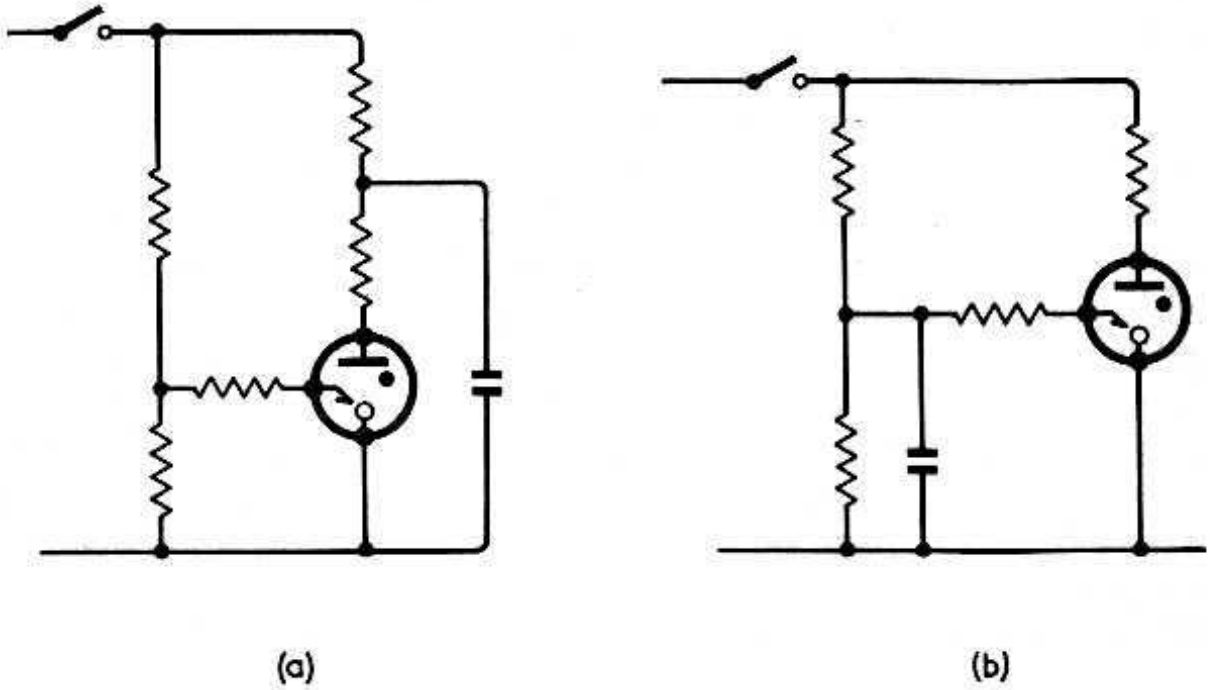


Fig. 7.20. Series switched extinguishing circuit (a) unsuitable and (b) suitable for close-tolerance tubes

both the anode–cathode and the trigger–cathode discharge must be extinguished. Further, since reverse current to the trigger can be harmful (this could occur if the trigger were taken negative during the anode–cathode discharge, see Section 7.9.1) it is necessary

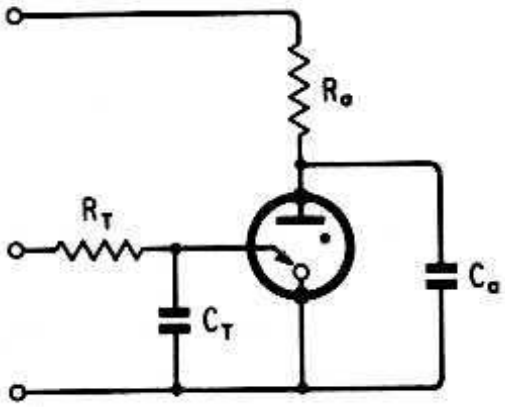


Fig. 7.21. Self-quenching extinguishing circuit

to extinguish the anode–cathode discharge at the same time or before the trigger–cathode discharge is extinguished. This does not present a serious problem, but certain precautions are necessary. For example, extinguishing the discharge by breaking the h.t.

supply would be unsuitable if an anode decoupling capacitor is connected across the tube; on the other hand a capacitor across the trigger input would ensure the correct sequence of extinguishing (Fig. 7.20).

Various extinguishing circuits suitable for control applications

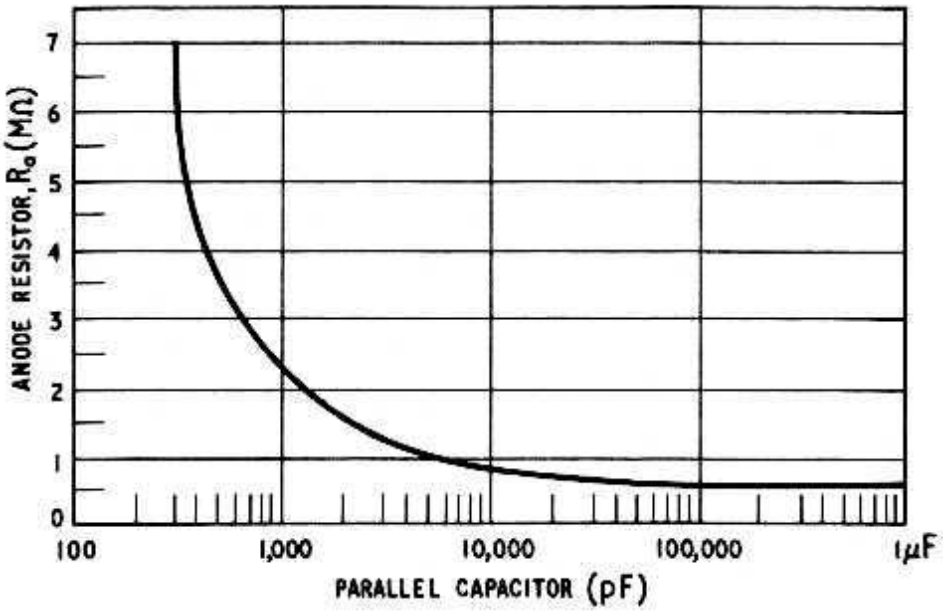


Fig. 7.22. Minimum value of anode resistor R_a in self-quenching circuit of Fig. 7.21, as a function of the capacitor for a Z803U. Gimson and Crowther[12]

have been given by Gimson and Crowther.[12] Of importance is the self-extinguishing circuit, where both the trigger–cathode and anode–cathode gaps are self-quenched, Fig. 7.21. The trigger circuit is the normal capacitor triggering arrangement, and providing R_T is greater than about 1 MΩ will always be self-quenching. The minimum value of R_a for various C_a values for self extinguishing the Z803U is given in Fig. 7.22.

7.8. The Design of Trigger Tubes

The design of trigger tubes falls roughly into two classes associated with the two major fields of application, namely control circuits and logic circuits. The first group includes such applications as flame failure control, timers, overvoltage protection and temperature control, and usually consists of miniature tubes with a high order of stability of the trigger-breakdown potential. The second group is used for such applications as machine tool control, counting circuits and business machines, where size and cost are of prime importance. These tubes are usually of sub-miniature flying lead construction. Clearly there will be exceptions where, say,

the characteristics of a tube meant primarily for control applications will be required for logic circuits and vice versa, and not all tubes will fall into the above categories. Also there are a number of design features common to both classes of tube.

Because of the number of parameters involved in the functioning of a trigger tube, the design is inevitably more complex than for stabilisers or indeed switching diodes. It is not to be expected that all the desirable characteristics can be met in one tube and compromises have to be made, so that a tube with a particular characteristic as one of its features may well have to be designed at the expense of others. This has led to a wide variety of designs with little similarity between tube types from different manufacturers. For this reason and also because the effect of gas filling, cathode material, and processing schedules has already been discussed in Section 6.3, the design parameters will not be discussed in such detail as in the case of stabilisers.

7.8.1. THE GAS FILLING

As with stabilisers, the gas composition and the pressure are largely governed by the desired breakdown and maintaining voltage levels. In some tubes, however, switching speed and repetition rate are of prime importance, and in this case the gas filling is chosen to give a fast ionisation and deionisation time.

Since the output voltage of a trigger tube is given by the difference between the anode supply voltage and the anode–cathode maintaining voltage, i.e. $V_B - V_{Am}$, and since V_B has an upper limit set by the anode–cathode breakdown potential V_{As}, a large potential difference is required between V_{As} and V_{Am}. On the other hand, it is desirable to keep the breakdown voltage of the trigger–cathode gap as low as possible from the point of view of the necessary input signal. These two requirements, although to some extent conflicting, can be satisfied by suitably positioning the electrodes, providing a large enough variation of breakdown potential with electrode spacing can be obtained. Thus, contrary to stabiliser requirements the gas must be chosen, not to give a 'flat' Paschen characteristic, but to give a Paschen curve in which the breakdown voltage rises steeply with electrode spacing. This can be achieved using pure gases such as neon, helium or argon (see Fig. 2.10). On the other hand, for reasons given in Section 6.3.1, Penning mixtures, which give rather flatter Paschen characteristics, are more suitable for stable and reproducible breakdown and maintaining voltages. Fortunately a compromise is possible by using Penning mixtures

in which the amount of gas with the lower ionisation potential is greater than 10%. Thus neon, with 10–15% of argon added to it, is quite a common gas filling in trigger tubes.

The required breakdown potentials similarly govern the pressure of the gas filling. For tubes with molybdenum cathodes where the maintaining potential in neon and neon-argon is around 100 V, an anode breakdown potential greater than 300 V is required if a reasonable output and tolerance on components and supplies are to be obtained. From the Paschen curves (Fig. 2.10) this implies a pd (d = electrode spacing) of 30–40 torr cm, and so if the electrodes are to be mounted within a miniature bulb, pressures in excess of 30 torr are needed; in sub-miniature tubes even higher pressures will be required. Since jumps in the maintaining potential-current characteristic are less important in trigger tubes, and since the circuit can often be arranged to operate the tube in the optimum mode for keeping the cathode clean, such pressures can give good reproducibility and stability on life. An advantage of the higher pressures and closer electrode spacing is the decrease in ionisation and deionisation times which result. For example with inert gases at 100 torr pressure a trigger tube working at repetition rates of 25 kc/s has been designed with normal tolerance components.[13]

A more common method of obtaining faster ionisation and deionisation is to add hydrogen to the inert gas (see Chapter 5), when similar repetition rates can be obtained at lower pressures. A trigger tube using such a gas mixture has been described by Hough and Ridler.[14] The disadvantage of using hydrogen is the difficulty of sustaining stable tube characteristics on life because of the activity of the hydrogen, and also the attendant higher values of breakdown and maintaining potentials. Nevertheless, such gas mixtures have been used successfully in several tube types and in particular in decade counting tubes (see Section 8.2).

7.8.2. THE CATHODE MATERIAL

The two types of cathode described in Section 6.3.2, that is cathodes coated with a low work-function layer, and pure metal cathodes prepared by sputtering, are also commonly used in trigger tubes.

The coated cathodes were employed in the early trigger tubes, but are still incorporated in modern tubes. They are used in some sub-miniature tubes suitable for logic circuits, where close tolerance on the trigger breakdown voltage is not too important but where a low voltage supply is desirable for economic and safety reasons. An example is the tube described by Domburg and Six[15] designed

as a speech passing switch for small telephone exchanges. With such cathodes the trigger breakdown potential is of the order of 70 V, and the anode breakdown 170–200 V. Since the maintaining potential is around 60 V the output is sufficient to trigger a second tube in cascade without the need of a bias potential. This is a valuable attribute for some logic circuits. Unlike stabilisers the glow is not continuously invested on the cathode, and providing the current is limited to a few milliamps the life of the coated cathode is not a major problem. Larger tubes for relay operation are also manufactured with barium-coated cathodes. Their main asset is that they are the only type of trigger tube that can be operated from low voltage a.c. mains supplies of 110–127 V. They are especially useful in applications having long standby periods between switching.

With a coated cathode it is impossible to attain the life stability or the narrow production spread obtainable with a sputtered pure metal cathode. Therefore, the latter type of cathode is more common in the majority of modern trigger tubes; normally the cathode metal is molybdenum. It is processed by sputtering in the same way as stabiliser cathodes, but since the gas is normally at a higher pressure pulse techniques are commonly employed. In tubes where the stability of the trigger breakdown potential is made a predominant feature, such as the Z806W, the maximum variation in V_{Ts} is only $\pm 2\%$ over ten thousand hours. The production spread of V_{Ts} is better than $\pm 4\%$. With neon-argon mixtures the trigger breakdown potential with a molybdenum cathode is of the order of 130 V, whilst the anode breakdown potential varies from 250–400 V according to the geometry. Since the maintaining potential is of the order of 110 V it is not always possible to connect tubes in cascade reliably without a trigger bias.

7.8.3. ELECTRODE CONFIGURATION

The position and geometry of the various electrodes control, more than any other factor, the performance of a trigger tube. This is especially true for the close tolerance sputtered-cathode tubes where secondary effects, such as spurious discharges to the sputtered film on the envelope, are observed, which are masked by the spread in characteristics in tubes with coated cathodes. Since breakdown rather than maintaining potentials are the main criteria in the functioning of a trigger tube, and since breakdown depends

on the field distribution between the electrodes, electrode shape as well as position is important to the design.

The cathode is normally in the form of a plate with its plane facing the anode, the dimensions being governed by the current carrying capacity. In the majority of tubes, the trigger electrode is in the form of a wire or rod, and is normally bent towards the cathode to ensure that breakdown occurs at the required point on the cathode and not to the leads. It is mounted close to the cathode, either between cathode and anode, or to one side. Its position and size governs the triggering performance.

As was explained in Section 2.4, the breakdown potential for a non-uniform electric field is not a unique function of pressure × electrode spacing. The minimum breakdown potential will depend on the field distribution and it can be depressed below the value pertaining to parallel-plate geometry by increasing the field at the cathode, for example by attaching a spike to the cathode plate. A decrease of 10% in the minimum breakdown voltage can be obtained by this method which is employed in the Z700U. Conversely, increasing the field at the trigger raises the minimum breakdown potential, and so for the conventional wire trigger construction the breakdown value will be somewhat higher than the minimum values given in Table 2.3 for parallel-plate geometry. The field at the trigger electrode not only depends on the diameter of the wire but also on the distance from the cathode, being lowered as the wire is brought closer to the cathode. Thus the minimum trigger breakdown potential that can be attained will also depend on the spacing, cf. Fig. 2.15.

The diameter and position of the trigger wire also affects the sensitivity of a trigger tube (the transfer characteristics). This is especially true for a trigger electrode disposed between cathode and anode, since the wire acts as a screen, so that only some of the ions and electrons will contribute towards the anode–cathode breakdown. The screening effect will depend on the wire diameter and also the extent of the glow, which in turn depends on the pressure. At high pressures the glow is very constricted and a finer trigger wire would be required than for the more diffuse glow at a lower pressure. Similarly a trigger discharge to the side of the cathode is less effective than one directly in the anode–cathode gap, particularly when the glow is constricted to a small area on the edge of the cathode. Thus the most sensitive tube, in terms of the trigger current required to ignite the anode–cathode discharge, will have a fine trigger electrode disposed between anode and cathode and will be filled to a low pressure. However, since the fine wire is not desirable from the point of view of breakdown

potential and rigidity of construction, nor is a very low pressure (Section 7.8.1), a compromise must be sought.

The actual shape of the trigger in terms of bends does not seem to have any pronounced effect on the transfer characteristics but the author[10] has found that increasing the cathode–trigger spacing

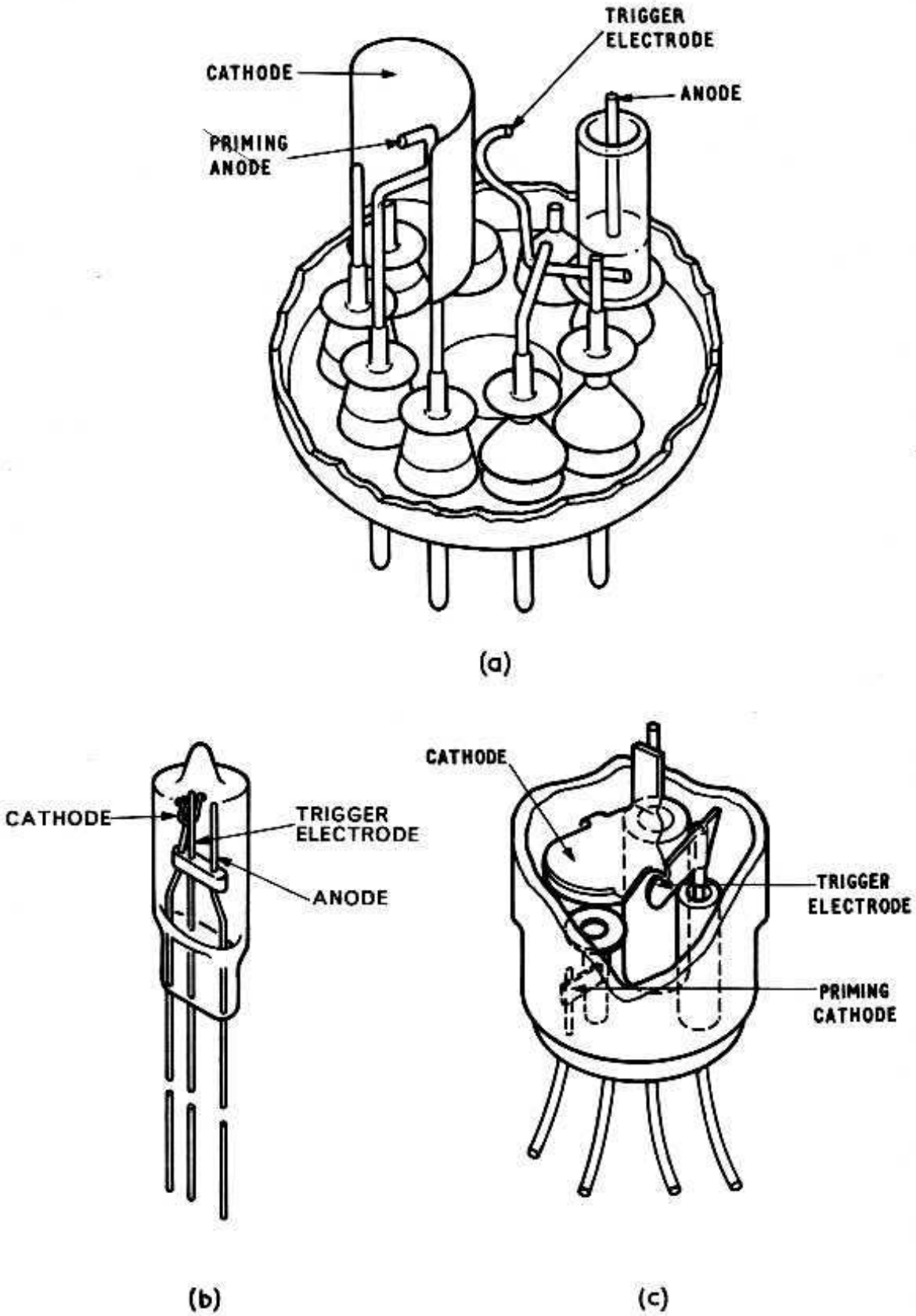


Fig. 7.23. The construction of typical trigger tubes; (a) the Z803U, (b) the GTR120W and (c) the Z700U

did increase the sensitivity when capacitor triggering was employed, Fig. 7.16. This may have been due to the increase in trigger breakdown voltage which would increase the voltage swing of the capacitor discharge. However the increase in V_{Ts} was only 8 V in 130 V, and it seems surprising that such a small increase in voltage could have produced the observed marked effect.

The anode is normally a straight wire screened so that the discharge only passes to the tip which is on a level with the cathode. The screening can be an insulated sleeve of glass or ceramic, or it can be a metal plate or tube which is electrically connected to an intermediate potential. The latter screening method increases the anode–cathode breakdown potential and is commonly employed in tubes designed to operate from 200–250 V r.m.s. mains. In the cheaper logic-application tubes a simpler construction is employed and screening of the anode is less common.

The priming electrode, when incorporated, can either be an auxiliary anode or cathode and will be mounted close to the appropriate electrode, but often remote from the main discharge. For example, a priming anode might be mounted behind the cathode plate (Fig. 7.23 (a)). The effect of the priming discharge on the trigger breakdown has already been mentioned, and its siting is very important in high sensitivity tubes.

The mounting of the electrodes in the envelope is governed by economic as well as technical considerations. For coated-cathode tubes, mica spacers can be employed to position the electrodes before mounting on to the foot. In sputtered tubes the positioning relies on rigidly fixing the electrodes on the foot 'lead-in' wires. For some precision tubes, the electrodes are held by a ceramic disk mounted directly on the glass foot. This prevents movement of the electrodes during sealing in when softening of the glass occurs.

Examples of typical trigger tube constructions are shown in Fig. 7.23. Fig. 7.23 (a) shows the electrode structure of the Z803U, a precision d.c. trigger tube for relay operation. The anode consists of a stiff nickel wire, partially surrounded by a glass tube, which increases the anode breakdown voltage. The trigger is in the form of a hook whose position relative to the cathode is closely controlled during assembly. The cathode is the largest electrode, made of molybdenum sheet bent in a U-shape. It is mounted on three supports for rigidity. The priming anode is mounted on the opposite side of the cathode to the trigger and anode and is also hook-shaped. The cathode is heavily sputtered during processing, and to prevent electrical leakage between electrodes small mica washers are mounted on the leads to prevent deposition of a continuous metal film over the foot. The whole structure is mounted

on a noval base and sealed into an envelope approximately 2·5 cm high. Fig. 7.23 (b) shows the GTR120W, an example of a low-priced sub-miniature switching tube with a barium-coated cathode. The cathode is a small 'getter' in the form of an iron tube containing barium. The trigger and anode are extensions of the lead-in wires held at the correct spacing by a glass bead. The sub-assembly is pinch sealed into a small envelope, and is particularly suitable for large scale mass production. The tube is filled with helium-oxygen mixture at a pressure of some 100 torr. When firing the getter, the oxygen reacts with the barium to form a barium-barium oxide coating on the cathode which is moderately stable. Fig. 7.23 (c) shows the electrode structure of the Z700U which is also a sub-miniature switching tube, but with a sputtered molybdenum cathode giving a much smaller spread, and greater stability of the voltage characteristics. The cathode is a molybdenum plate having a spike directed towards the trigger electrode, which is a flat metal strip. The anode is in the form of a disk mounted below the cathode, with a priming cathode close to it. The cathode and trigger leads are screened by ceramic tubes. The whole structure is mounted on a 'button' foot.

7.9. Trigger Tube Performance

The static breakdown potentials, triggering performance and recommended operating conditions are given for the individual tube types by the manufacturer. However, there are one or two factors affecting the performance and life of trigger tubes which are not always appreciated by the user, and are sometimes not made clear in the manufacturer's data. One such factor, which has come to be known as hysteresis, is the depression of the trigger breakdown voltage after passing a large anode–cathode discharge current. The effect is similar to deionisation in that the breakdown potential slowly recovers, but with a time constant which is large compared with that due to deionisation. Hysteresis is present to some degree in all trigger tubes, but is only important in close-tolerance tubes where trigger breakdown is used as a reference voltage. The effect in the Z803U can be seen from Fig. 7.24 where the variation in trigger breakdown voltage is plotted as a function of the time after extinction of an anode–cathode discharge at various currents. The phenomenon has been studied by Smith,[16, 17] who demonstrated that the depression of the breakdown voltage was caused by heating of the cathode as a result of the discharge. The temperature of the cathode did not itself change the break-

down potential, but caused the gas density ρ to vary across the gap, lowering it close to the cathode. Under these conditions, which give a non-uniform E/ρ distribution, where E is the field, the effective value of the second Townsend coefficient γ was found to increase beyond the value determined under uniform E/ρ conditions for the E/ρ applying at the cathode. The value of V_{Ts} is thus

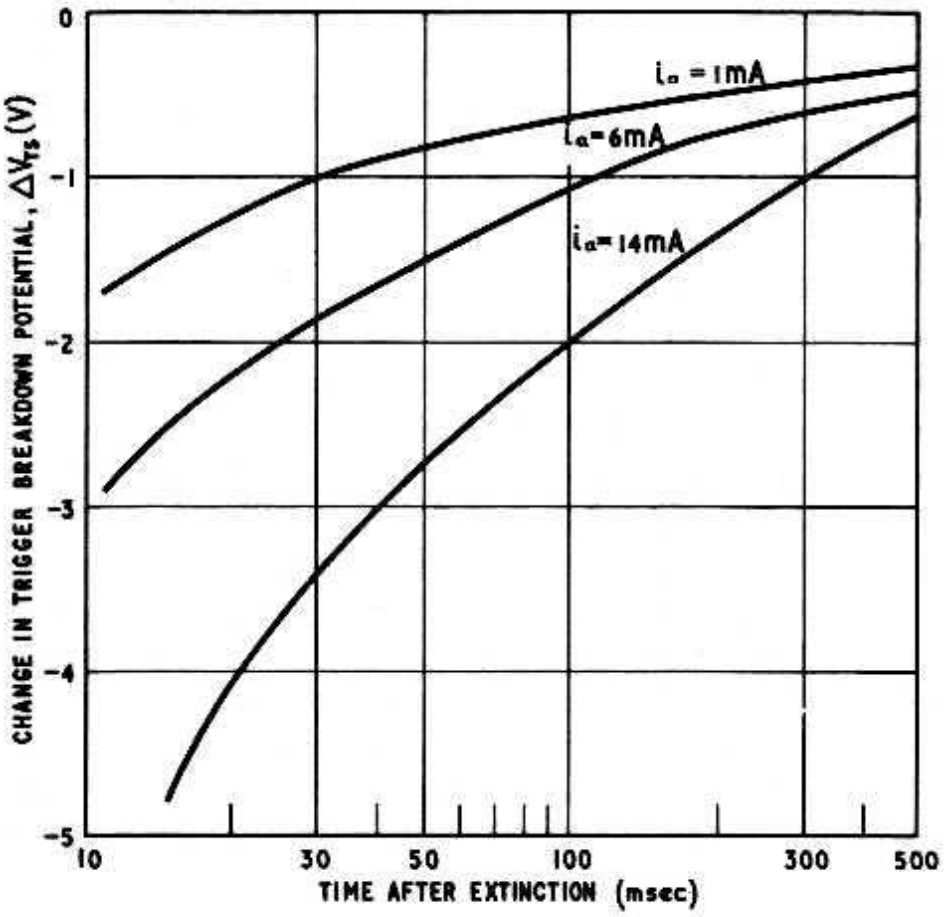


Fig. 7.24. Variation in trigger breakdown potential after passing a discharge current in the anode–cathode gap as a function of time after extinction, for a Z803U

lowered (cf. Chapter 2). η also depends on E/ρ, but the measured change in η after a discharge was insufficient to explain the observed breakdown depression.

The effect can be minimised by ensuring good thermal conduction from the cathode via the supports, and also by operating at a pressure below that giving the minimum breakdown potential. However, in the latter case, the voltage change may alter sign with time after extinction for certain anode–cathode currents.

The impedance of the glow discharge has already been discussed. It is relatively low, of the order of kilohms, and for normal trigger tube application does not present any problems. However, to obtain the high breakdown potential between cathode and anode, the electrode spacing is such that when the discharge is passing, a

glow usually occurs at the anode with a corresponding rise in potential, i.e. an anode fall (see Section 3.3).

Although the anode fall presents negligible impedance, it can give rise to oscillations, in the audio frequency range, of several volts amplitude superimposed on the maintaining potential. Normally these are not very worrying, although they could cause instabilities in the circuit under exceptional circumstances. However, there are some applications, for example small telephone exchanges, where audio oscillations would be quite unacceptable. For such applications, special tubes have been designed in which the anode fall is eliminated.[15]

7.9.1. THE LIFE OF A TRIGGER TUBE

As with stabilisers, the life of a trigger tube during operation is characterised by a slow drifting of the relevant discharge potentials until one of them, usually the trigger breakdown potential, becomes unacceptable. Normally this will only occur after thousands of hours of passing anode current. The extent of the drift depends on a number of factors, but is fundamentally caused by either contamination of the cathode around the triggering area or erosion of the cathode as a result of sputtering. Contamination can be due to gases evolved from the electrodes and/or the envelope, or to deposition of sputtered material from other electrodes, particularly the trigger. The latter can be the more serious cause of contamination, but will normally only occur if the tube suffers inverse breakdown between the anode or trigger and the cathode, or if appreciable negative trigger current is drawn during anode-cathode conduction. Care must therefore be taken in the design of the circuit to avoid either of these contingencies. The effect of contamination can be minimised by ensuring that the glow is invested in the triggering area during the main discharge, so that the relevant portion of the cathode is continually cleaned by sputtering. To this end, the manufacturer will sometimes advise taking a negative current to the trigger during anode–cathode conduction in order to anchor the discharge[11] in the triggering region, especially when the current is insufficient to cause complete coverage of the cathode surface by the glow. However, it is dangerous to apply this principle to all tubes, and generally the tubes should be run at a current large enough to give complete cathode coverage, or at least coverage in the triggering area (the manufacturer usually quotes minimum safe currents for this purpose). Providing the manufacturer's recommendations are followed, the

tube should not drift outside limits due to contamination, and the tubes will finally fail as a result of sputtering. For the coated cathode, sputtering will eventually remove the active layer, and the respective voltages will rise to those representative of a discharge with a cathode of the substrate metal. With pure metal cathodes two effects are observed; firstly whiskers form on the trigger, reducing the trigger-cathode spacing and in extreme cases forming a short circuit, and secondly part of the cathode, often in the area of the trigger, erodes away completely forming a hole, thus increasing the trigger–cathode gap. Clearly the life will be related to the rate of sputtering, and hence to the current density and duty cycle during the operation of the tube. Several authors have suggested an inverse-power law relation between the life of a glow discharge tube and the direct current. Recently the behaviour of trigger tubes during life has been studied in more detail by Bond.[18, 19] He found that in general the breakdown potentials drifted whilst the transfer characteristics remained substantially constant, but that if any deterioration of the transfer characteristics did occur then it was strictly in step with deterioration of the breakdown potentials. Thus the trigger–cathode breakdown potential could be taken as the criterion for determination of life. Tests were carried out with tubes operated from a half-wave rectified 50 c/s a.c. supply and also under d.c. conditions. The average changes in V_{Ts} during life for Z803U tubes operated in the former mode at various peak and mean currents are shown in Fig. 7.25. The pattern of behaviour is similar at all currents, except the highest; V_{Ts} first drifts slowly by approximately 0·5% of its initial value, then decreases steadily going outside the specified limits. Further running of the tube results in a marked and rapid increase in V_{Ts}. It is considered that the initial drift is due to either cleaning or contamination, whilst the fall and final rise is a result of sputtering (the fall being due to whisker formation on the trigger, and the rise due to erosion of the cathode). The results at high mean currents suggest that there is an upper mean current between 75 and 100 mA at which contamination, presumably due to overheating, becomes serious. Taking end of life as the time at which V_{Ts} had fallen by -2% of its initial value, a plot of this life time against peak current on logarithmic scales gave a straight line, Fig. 7.26. From the slope, the life was roughly inversely proportional to the fourth power of the peak current (cf. sputtering rate measurements, Section 4.4). A similar plot was obtained for d.c. operation. Life tests were also reported on two other trigger tubes having sputtered molybdenum cathodes, namely the Z806W and the Z700U. Both showed similar behaviour during operation.

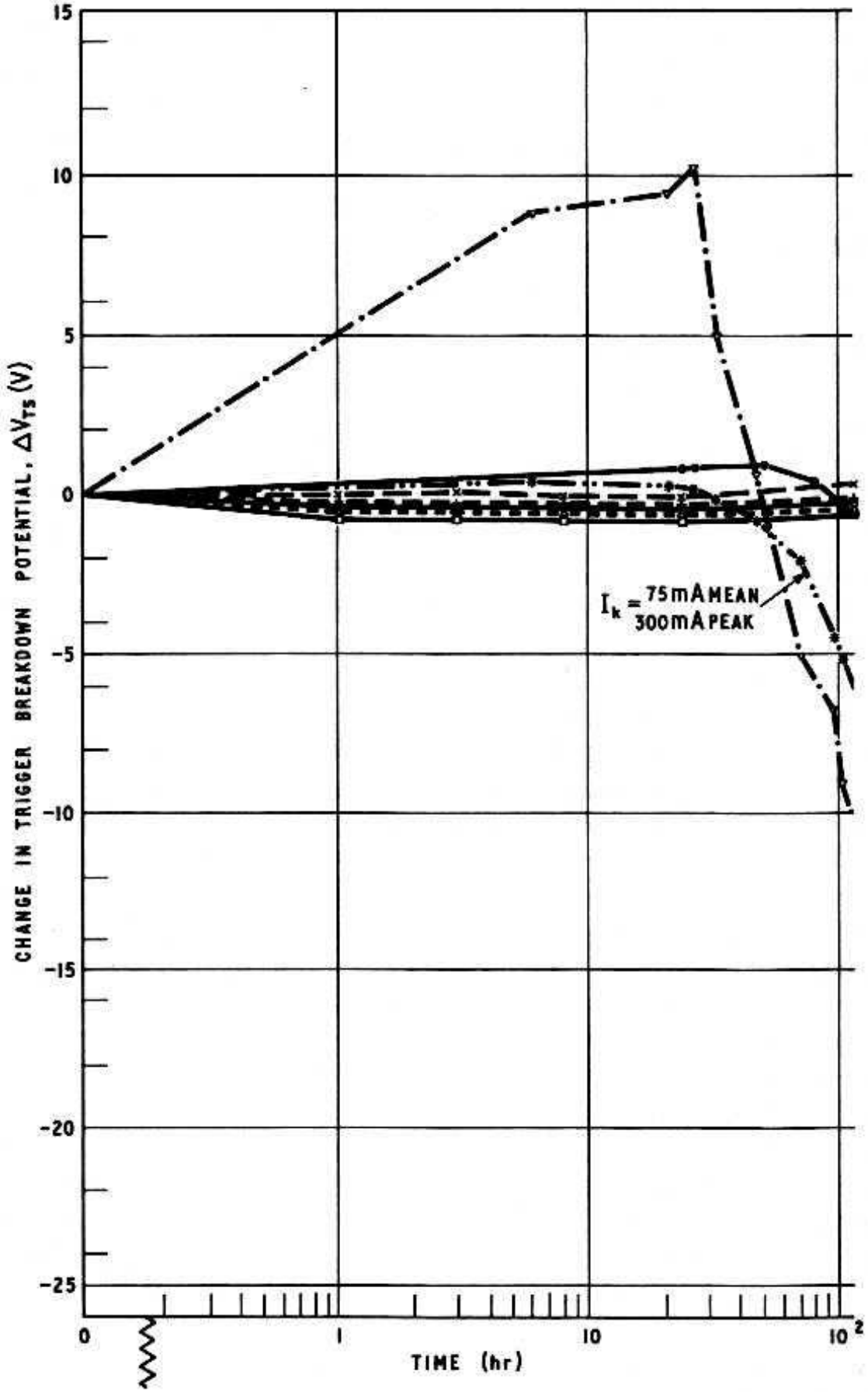
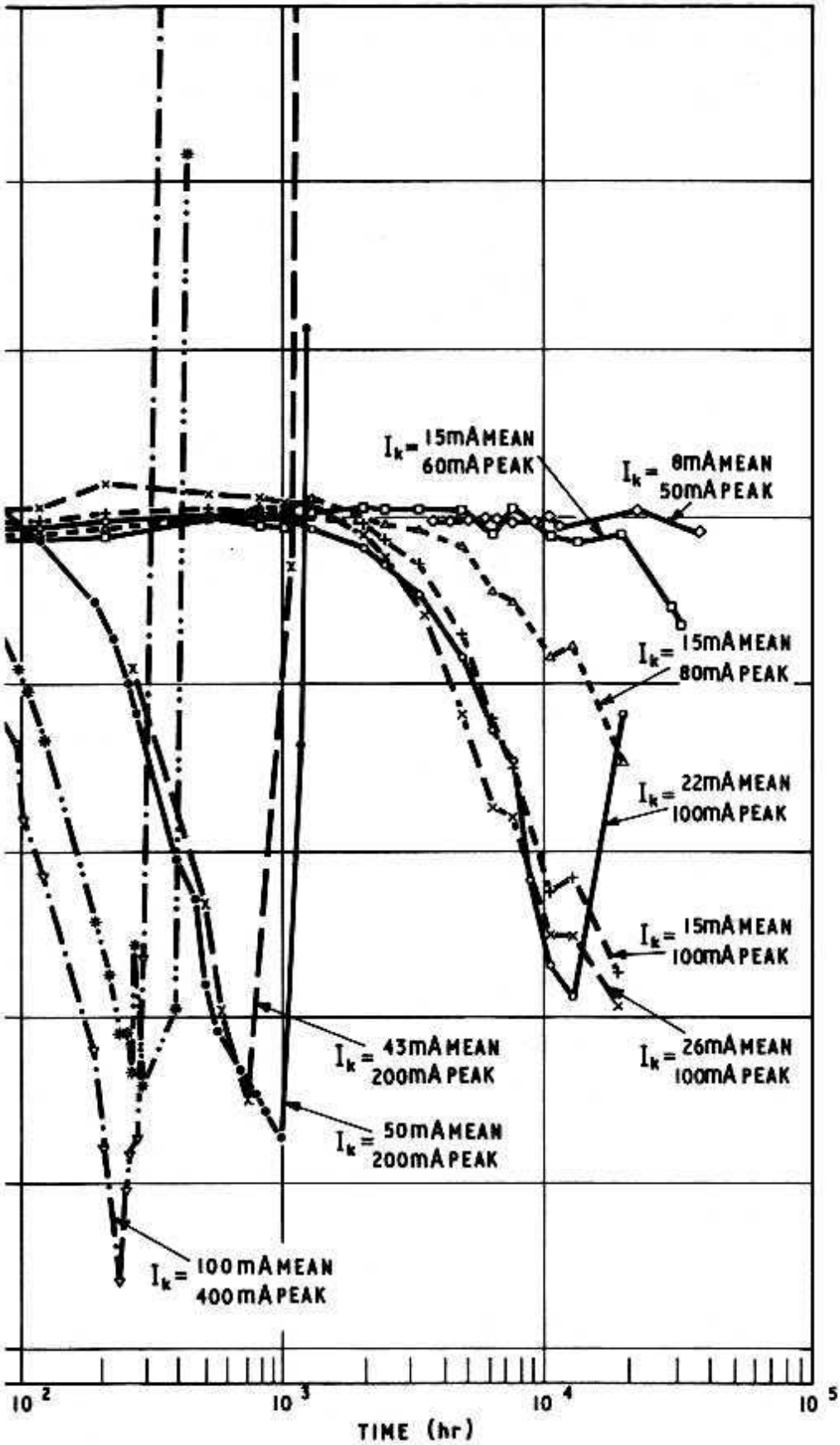


Fig. 7.25. Average variation in trigger breakdown potential of Z803U's during a.c.



operational life as a function of time, with the current ratings as parameter. Bond[18]

The Z806W is a modern trigger tube which maintains its characteristics with greater precision and has a longer life than the Z803U; comparison with Z803U of life as a function of peak current is shown in Fig. 7.27. Tests under 'standby' conditions were also made, and at room temperature the variations over thousands of

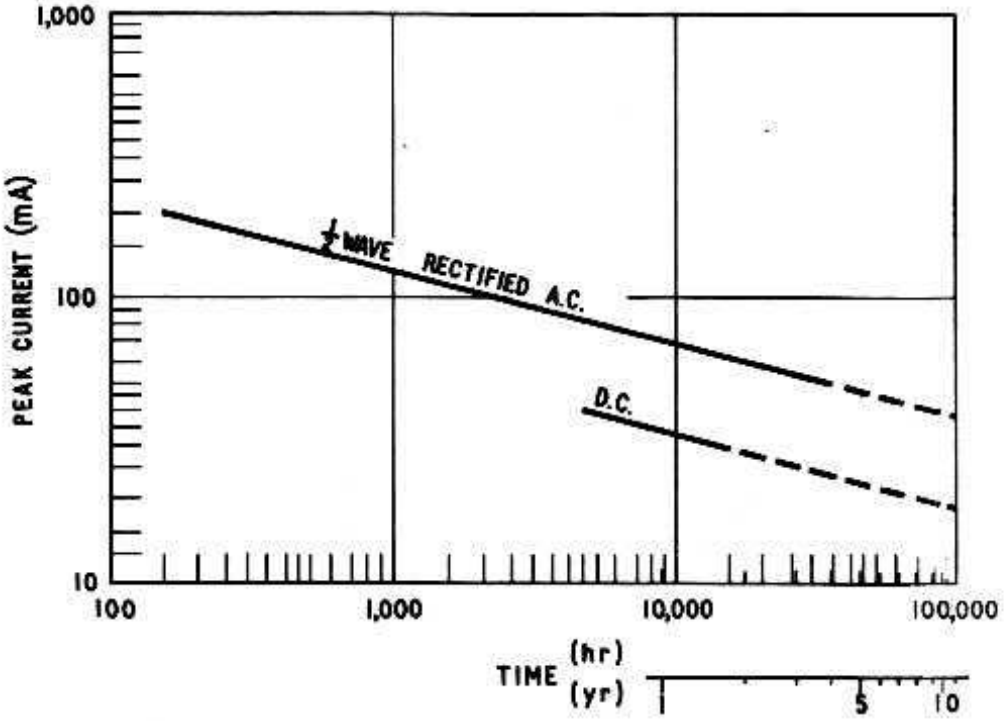


Fig. 7.26. Relationship between the peak current and life expectancy for the Z803U. (The 80% points of the distribution have been plotted i.e. 80% of the tubes can be expected to have a life in excess of that shown.) Bond[19]

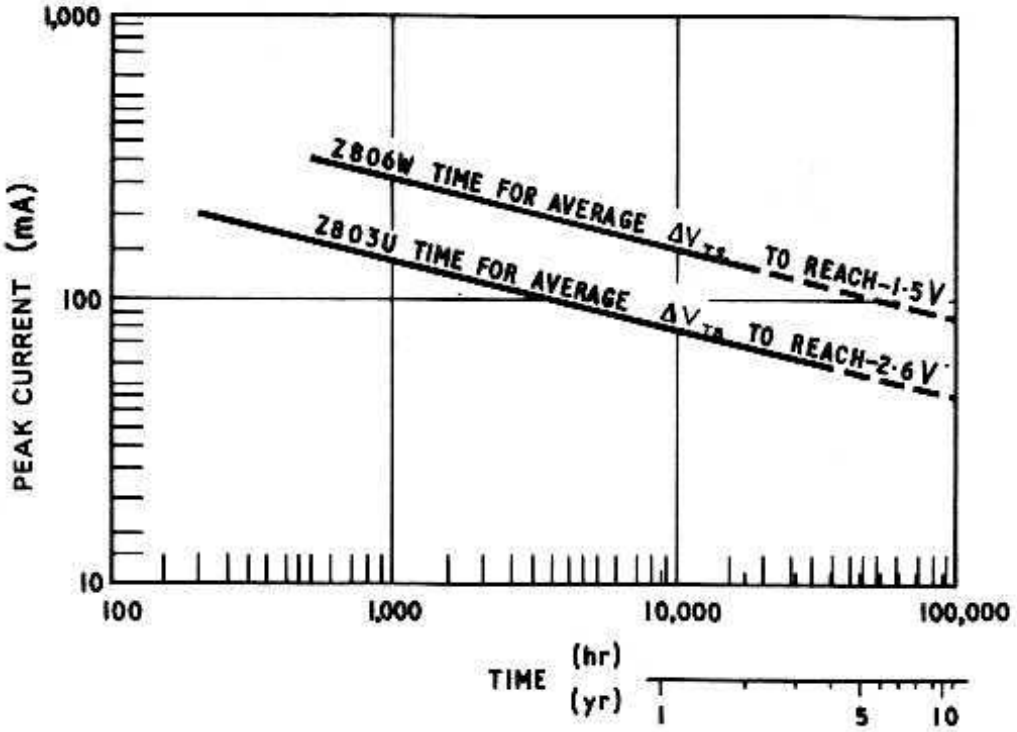


Fig. 7.27. Relationship between the peak current and life expectancy for the Z806W and Z803U ($\frac{1}{2}$ wave rectified a.c.) Bond[19]

hours (30,000 in the case of Z803U) were less than $\pm 2\%$. Because of the similarity of the pattern of behaviour between tubes and at various currents, and because 'standby' has little effect on the trigger breakdown voltage, it is possible to predict the life expectancy under any operating conditions, providing the conditions do not upset the mechanism described; for example overheating the tube or allowing inverse breakdown would invalidate the calculations.

Although the pattern of behaviour was slightly different for the coated-cathode tube Z900T, the general conclusions relating to life expectancy could also be applied to this tube type, i.e. that the performance at the rated current could be assessed from accelerated life tests.

7.10 Trigger Tube Circuits and Applications

The basic techniques of firing and extinguishing a trigger tube have already been given. The application of these techniques to the two classes of circuits—control, relay, and timing circuits, and logic circuits—will be discussed in this section.

7.10.1 CONTROL, RELAY AND TIMING CIRCUITS

In the majority of uses under this heading the trigger breakdown voltage is employed as a reference voltage. A change of state of some input quantity is arranged to raise the trigger voltage above the breakdown level, and is detected by conduction in the tube.

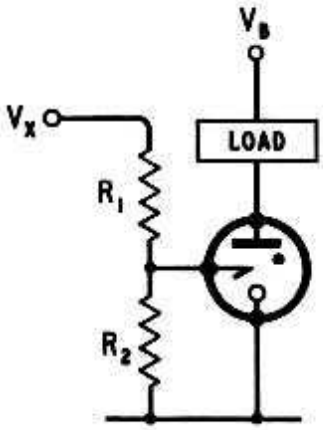


Fig. 7.28. Basic detector circuit

Typical examples are overvoltage detectors, flame-failure detectors, and temperature regulators. Timers also use V_{Ts} as a reference voltage and can be used for process control, welding control and delay switches.

The basic detector or protection circuit is shown in Fig. 7.28. If R_1 and R_2 are fixed resistors, the circuit can be used as an overvoltage detector in which a fraction of the supply voltage to be monitored is compared with the trigger breakdown voltage. R_1 and R_2 are adjusted so that when the input V_X exceeds a predetermined level,

$$V_X \geqslant V_{Ts}\left(\frac{R_1+R_2}{R_2}\right)$$

the tube ignites and remains conducting. The current which then flows through the anode load can pull in a relay to operate an alarm and/or protect other equipment on the input line. By careful design a practical circuit can be made to respond to overvoltages of less than 10 μsec duration. Using a fixed value of V_X and R_1 and replacing R_2 by a light-sensitive resistor, the circuit can be used to detect failure of a flame in, say, oil-burning boilers. R_1 and V_X are chosen to give a voltage at the trigger of less than V_{Ts} with the flame alight. When the flame fails R_2 increases and thus V_T, so that the tube fires. The conditions for operation are therefore:

$$R_2 \text{ (flame on)} < R_1\left(\frac{V_{Ts}}{V_X - V_{Ts}}\right) < R_2 \text{ (flame out)} \qquad (7.3)$$

By making R_1 a temperature sensitive resistor and fixing R_2 and V_X the tube can be used to detect when a predetermined temperature has been reached.

The disadvantage of this type of detector circuit is that it is not 'fail-to-safe' i.e. failure of the circuit does not give the same indication as the fault being detected. This problem can be overcome by using a.c. on the anode so that the discharge will be extinguished every half cycle. A basic flame-failure circuit using a photocell is shown in Fig. 7.29. The anode supply voltage to the Z803U is a halfwave-rectified (to prevent inverse breakdown) unsmoothed 50 c/s voltage with a peak value within the anode voltage operating range of the tube. In the presence of the flame, current through the photocell charges the capacitor C_2 via R_1 to V_{Ts}, so that the tube fires during each cycle of the mains. The capacitor C_1 smooths out the photocell fluctuations due to flickering of the flame. The current through the trigger tube is thus a series of pulses at mains frequency. To prevent the relay chattering, C_3 is connected across it. In the event of a failure of the flame the photocell current will cease, and C_2 will not charge up to V_{Ts}. The relay will then become de-energised and the contacts used to initiate appropriate action. The

relay will also be de-energised by circuit faults so that the system is 'fail-to-safe'.

This basic circuit only operates over a limited range of light levels, since reliable operation depends on the rate of charging of C_2 giving trigger breakdown coincident with the anode voltage

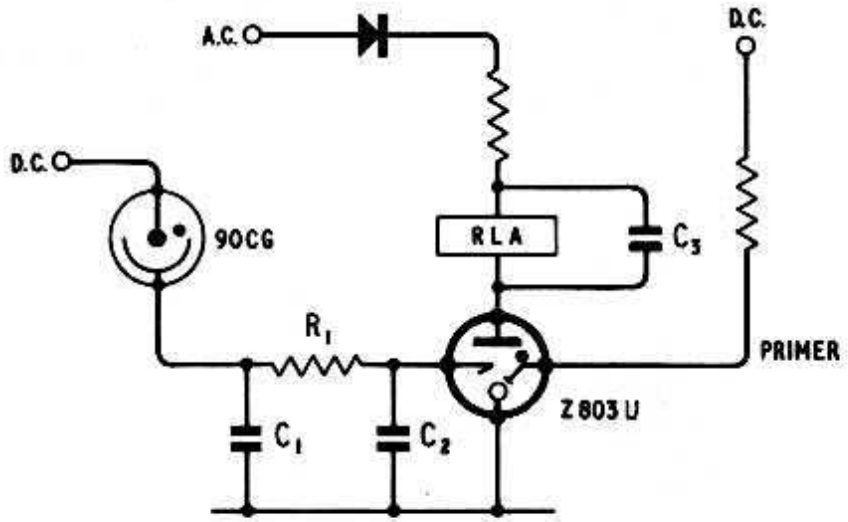


Fig. 7.29. Basic flame-failure circuit which is fail-to-safe

frequently enough for the average current to hold in the relay. To increase the range Crowther[20] describes an alternative circuit where negative feedback is applied. This effectively synchronises the charging circuit to the anode pulses.

Clearly the application of unsmoothed rectified a.c. to the anode can be employed with similar input circuits to those already suggested for a d.c. anode supply. The advantage is that the circuit can be self re-setting, i.e. when the input rises above a critical level the relay pulls in, but when it falls again the relay will drop out. Thus if one of the resistors in the input potentiometer is temperature sensitive the circuit can be used as a temperature controller, and in practice controllers have been designed to switch on and off with 1°C change in temperature. In some tubes designed to have a high inverse breakdown between anode and cathode, such as the Z804U and Z805U, the rectifier is not required. It is worth noting that the input impedance for such detector circuits is greater than $10^{10}\Omega$ up to the point of trigger breakdown, and, with a suitable series resistor, can be above 1 MΩ during conduction.

An important application of the trigger tube to control circuits is in timers, and several papers have appeared in the literature on their design and performance.[21, 22] The basic timer circuit is shown in Fig. 7.30. The capacitor C_T is charged by way of the resistor R_T, and its potential rises exponentially from zero towards V_X until it reaches V_{Ts}. When V_{Ts} is reached the tube strikes, and the current

through the load, which can be either a resistor or a relay, is used to activate other electronic equipment. The timing period is given by the expression:

$$t = C_T R_T \log_e \left(\frac{V_X}{V_X - V_{Ts}} \right) \tag{7.4}$$

The accuracy and stability of the timer depend on the variations of C_T, R_T, V_X and V_{Ts}. It is seen, however, that variations in either

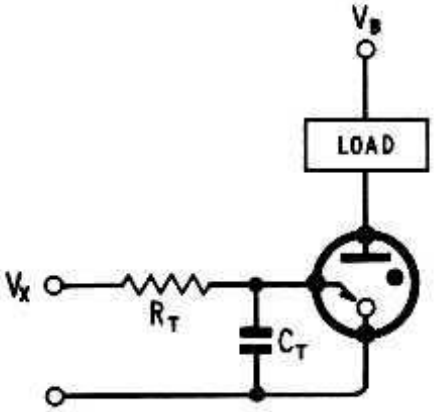


Fig. 7.30. Basic timer circuit

R_T or C_T will cause proportionate changes in t, but changes in V_X or V_{Ts} will have a greater effect. Light[21] showed that the fractional change in t due to fractional changes in V_X or V_{Ts} could be expressed by the following equations:

$$\frac{\Delta t}{t} = \beta \frac{\Delta V_{Ts}}{V_{Ts}} \tag{7.5}$$

and

$$\frac{\Delta t}{t} = -\beta \frac{\Delta V_X}{V_X} \tag{7.6}$$

where

$$\beta = \frac{V_{Ts}/V_X}{(1 - V_{Ts}/V_X) \log_e \dfrac{1}{(1 - V_{Ts}/V_X)}}$$

He plotted β against V_{Ts}/V_X and showed that it was always greater than 1 but approached 1 as V_{Ts}/V_X tended to zero. For highest accuracy, therefore, V_{Ts}/V_X should be as small as possible. However, there are practical limits to the upper value of V_X which can be employed, and in general values of V_{Ts}/V_X between 0·4 and 0·8 are used. With a paper capacitor, a 1% stability evaporated film resistor, and a supply voltage stabilised by a reference tube, the long term stability (1 year) of a typical timer, with say a Z803U tube, is of the order of 2 or 3 per cent. A short term stability of better than 0·1 per cent over an eight-hour period is obtainable.

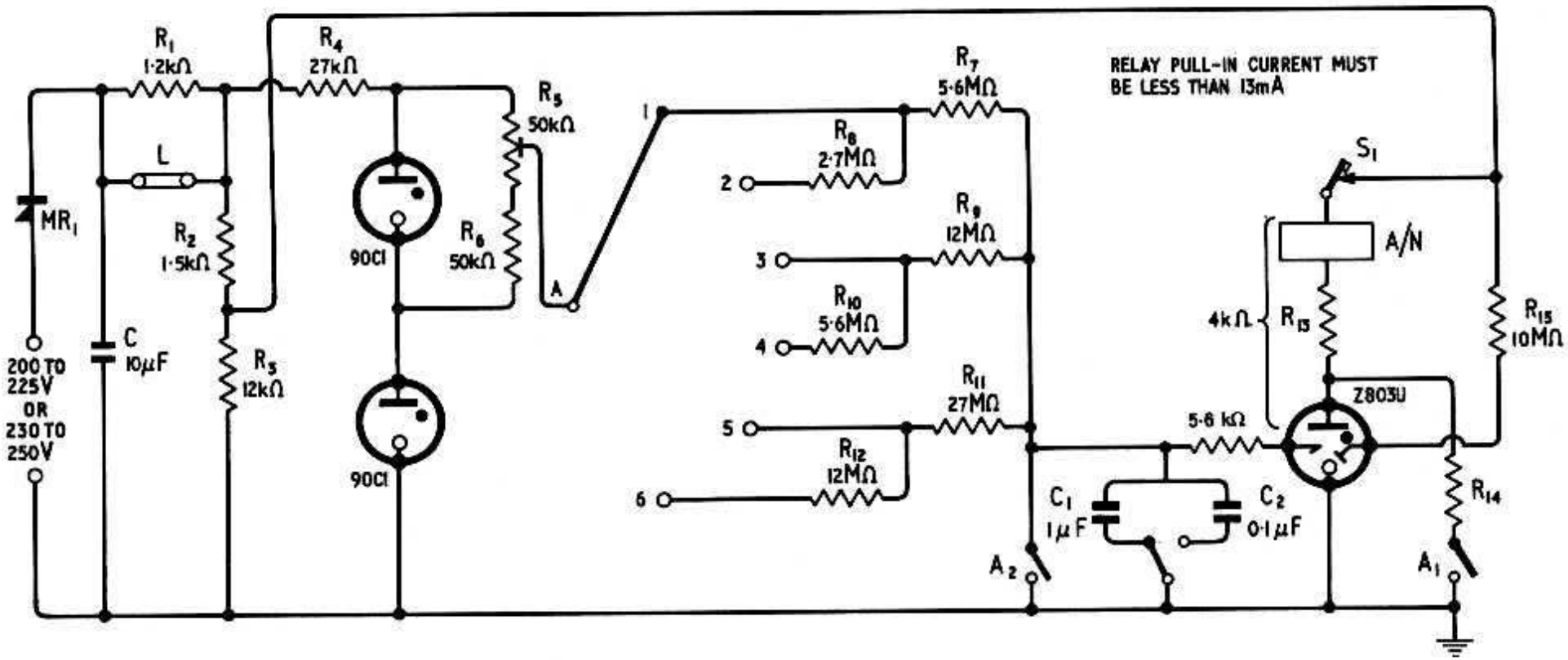


Fig. 7.31. Circuit for a general purpose timer using a Z803U. Gimson and Crowther[12]

Timing periods from tens of milliseconds up to minutes are attainable with suitable choice of R_T and C_T.

A general purpose interval timer working off the mains is shown in Fig. 7.31. Timers, where the time interval is inversely proportional to a power law of the input voltage, i.e. $t \propto 1/V_X^n$, have been described by Hercock and Neale[22] for timing photographic exposures.

7.10.2. LOGIC CIRCUITS

The application of logic circuits covers a wide field which includes scientific computing, industrial computing and control, telephone exchanges, and nucleonic scaling. Where the logic requirements do not have to be met at high speeds, a number of advantages can be obtained by the use of trigger tubes. The principal one is that the tubes themselves perform a considerable amount of the logic as well as acting as a memory, pulse amplifier, and pulse forming circuit. It is beyond the scope of this book fully to describe circuits for the many applications, and all that can be attempted here is to outline the basic gating and switching functions that can be achieved with trigger tubes. A more detailed treatment can be found in books by Neale[5] and Durand.[23]

Fig. 7.32 illustrates three methods of producing an 'AND' gate. The first of the gates, Fig. 7.32 (a), is the familiar pulse-plus-bias circuit (cf. Section 7.6.2), which forms the basis of most counting circuits. It has the one disadvantage that the bias must be applied some time before the pulse to allow time for the capacitor to charge. This is overcome in the circuit of Fig. 7.32 (b) where the two pulses can be applied simultaneously. The combined amplitude of the two pulses must, of course, exceed V_{Ts}. These two circuits combine two inputs; in the circuit of Fig. 7.32 (c) any number of inputs may be used. If no signal appears at 2 and 3 the diodes are not backed off and the signal on 1 is virtually shorted through the diodes. If, however, 2 and 3 are backed off by suitable positive signals, the signal on 1 is sufficient to fire the trigger tube. The necessary circuit conditions are:

$$\frac{V_{Ts}-V_p}{V_B-V_{Ts}} < \frac{R_1+R_2}{R_1+R_2+R_3} < \frac{V_{Ts}}{V_B} \tag{7.7}$$

assuming that $R_1 \ll R_d$ where R_d is the back resistance of the diodes. In a practical circuit account must be taken of the tolerances on voltages and resistors.

The circuits of Fig. 7.32 (a) and 7.32 (b) can also be used as 'OR' gates if each of the applied signals is greater than V_{Ts}. However, an 'OR' gate can be greatly facilitated by use of a trigger tube having two symmetrically placed trigger electrodes, such as the Z700W. In this case, a pulse of sufficient amplitude applied to either trigger

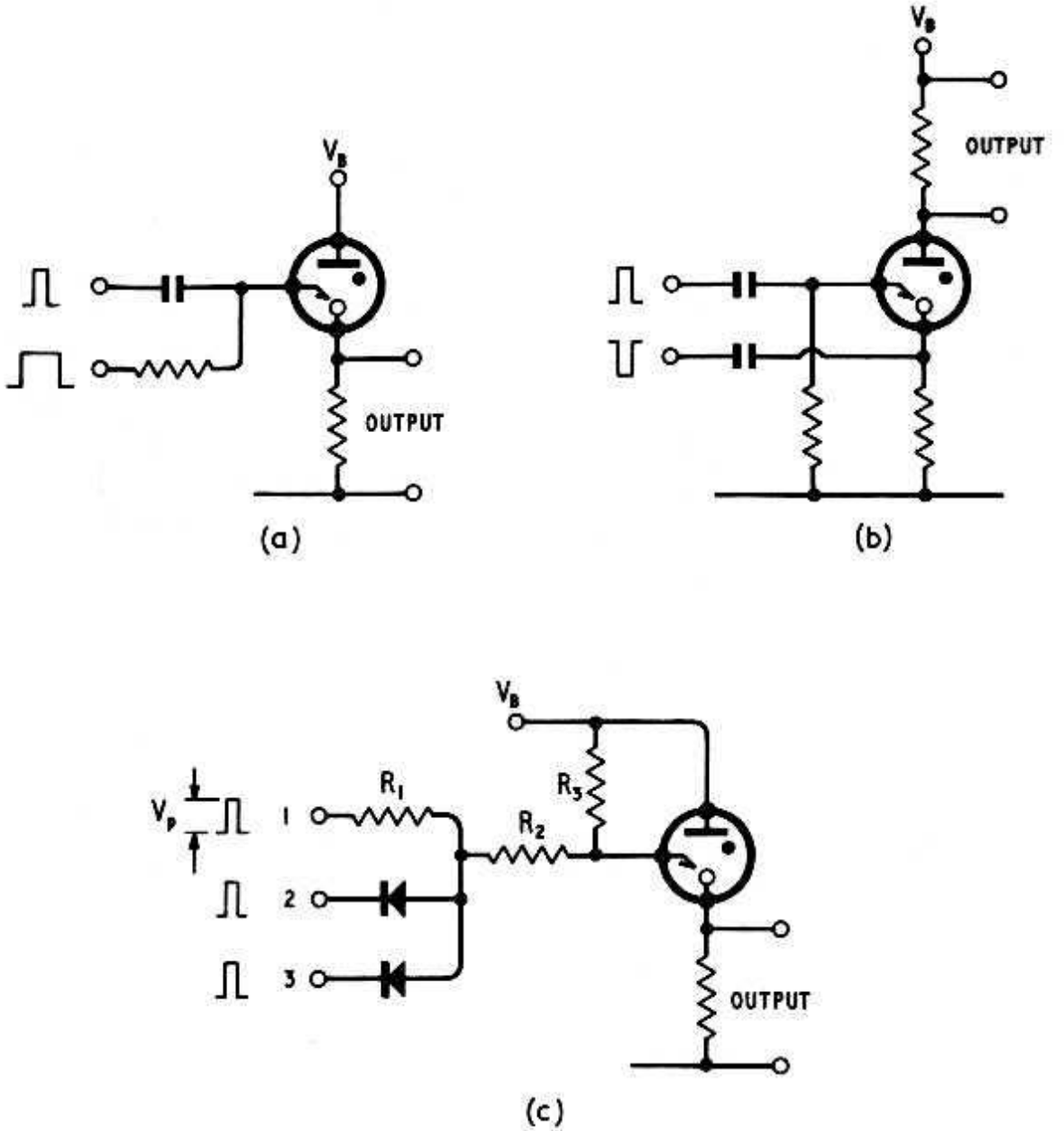


Fig. 7.32. Three circuits for obtaining logic 'AND' gates, employing (a) pulse-plus-bias, (b) double pulse, and (c) diode logic

will fire the tube, Fig. 7.33 (a). To cover more than two inputs a diode gating system on the input can be employed as in Fig. 7.33(b).

Two 'NOT' gates are illustrated in Fig. 7.34. They are effectively the inverse of the 'AND' gates of Figs. 7.32 (a) and 7.32 (b); the positive pulse to the trigger must now exceed V_{Ts}. The inhibiting

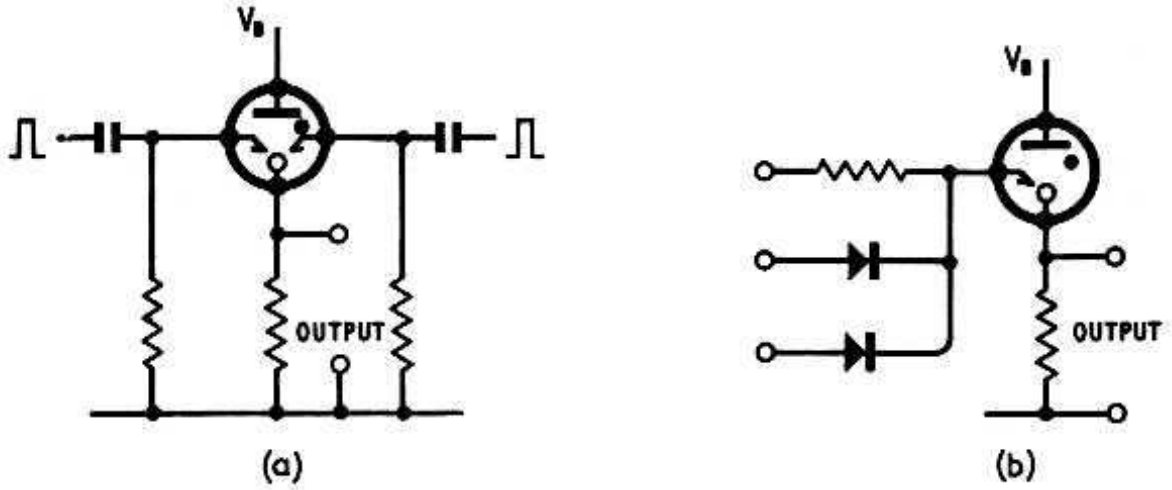


Fig. 7.33. Two 'OR' gates; (a) employing a double-trigger tube and (b) a diode logic circuit

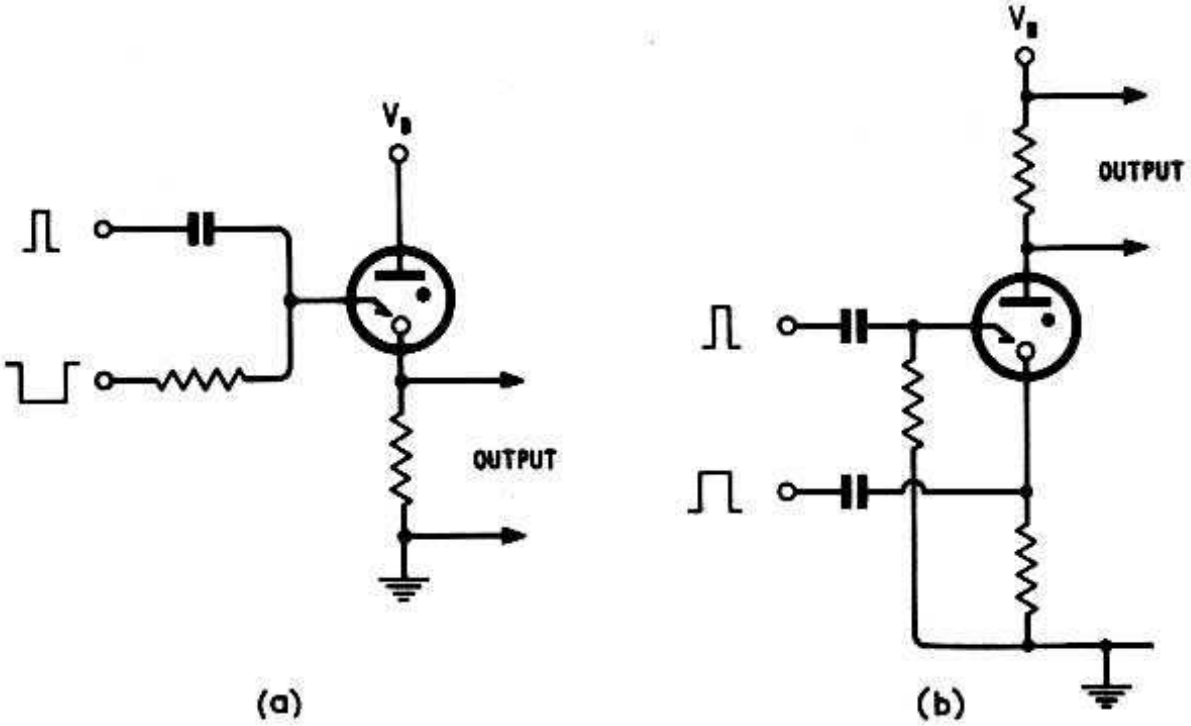


Fig. 7.34. Two 'NOT' gates, (a) employing pulse-plus-bias (b) double pulse

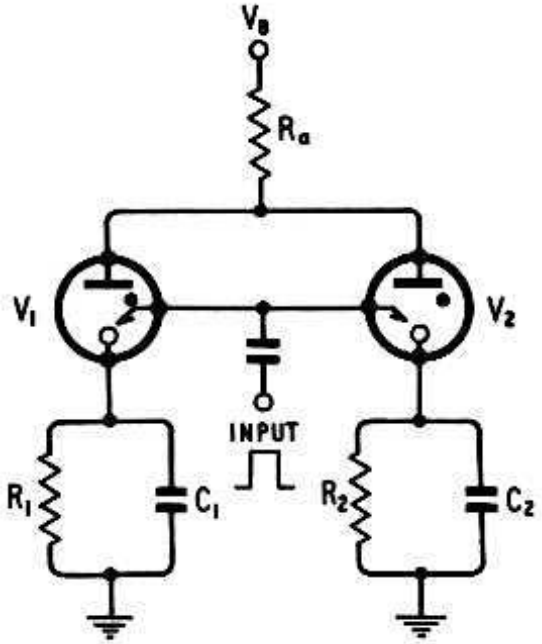


Fig. 7.35. Basic common-anode binary counter

pulse should be of $0{\cdot}5\ V_{Ts}$–$0{\cdot}7V_{Ts}$ amplitude. In addition it should be noted that combinations of the 'AND' and 'NOT' gates can be employed in conjunction with the 'OR' gate of Fig. 7.33 (a).

Trigger tubes can also be employed in another important logic operation, that of counting. A basic binary counter is shown in Fig. 7.35, using a common anode load. Assuming V_1 is conducting, then the cathode potential will be at V_k, and the anode at V_A where

$$\left.\begin{aligned} V_A &= V_{Am}+V_k \\ V_A &= V_B - iR_1 \end{aligned}\right\} \qquad (7.8)$$

and

$$V_k = iR_1$$

The anode–cathode potential on V_2 is equal to V_A, and R_1, R_2 and V_B are chosen such that V_A is within the recommended anode voltage range of the trigger tubes. When a pulse of greater than V_{Ts} is applied to the common trigger line, tube V_2 will ignite. Initially, its cathode will be held at earth by capacitor C_2 so that

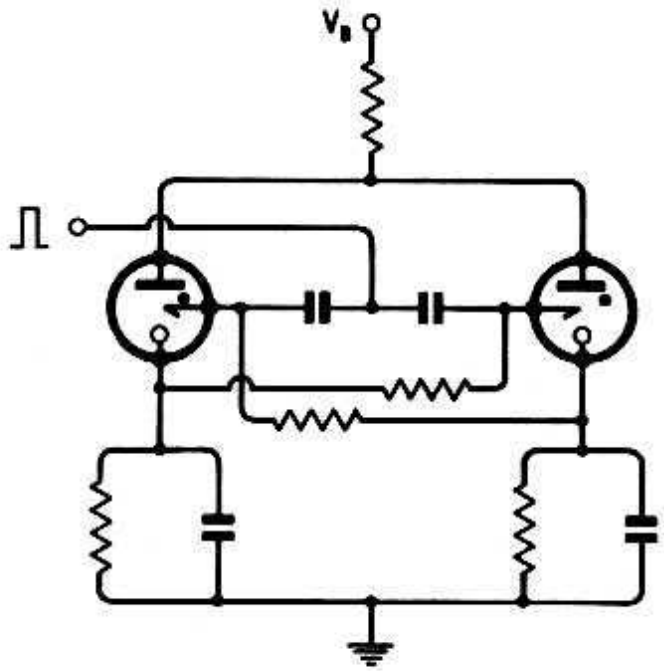


Fig. 7.36. Common-anode binary counter with pre-bias

the common anode connection is lowered to V_{Am}. Since C_1 initially holds the cathode of V_1 at V_k, $V_{Am}-V_k$ appears across V_1 and the tube will be extinguished. The cathode potential of V_2 will rise to V_k exponentially with a time constant $\tau_2 = C_2\ R_a\ R_2\ (R_a+R_2)$, whilst that on the cathode of V_1 will fall to zero with a time constant

$\tau_1 = R_1C_1$. At any time t:

$$\text{Voltage across } V_1 = V_A - V_k \exp\left(\frac{-t}{\tau_1}\right) - V_k \exp\left(\frac{-t}{\tau_2}\right) \tag{7.9}$$

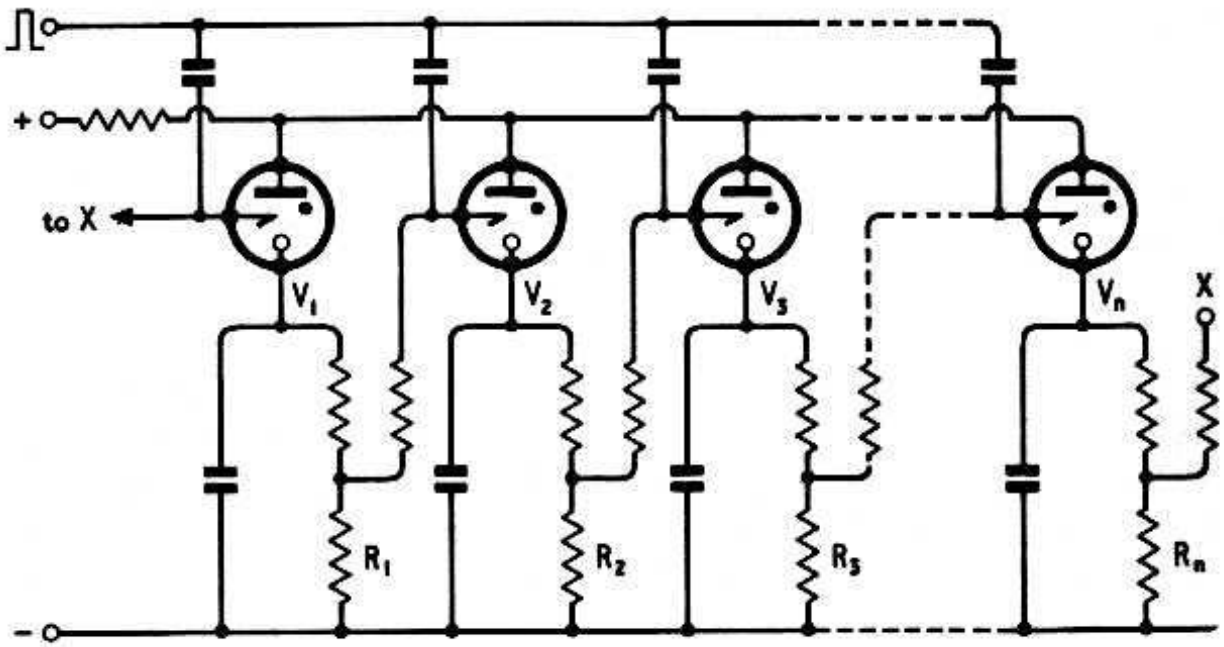


Fig. 7.37. A ring counting circuit with cathode coupling to provide the pre-bias

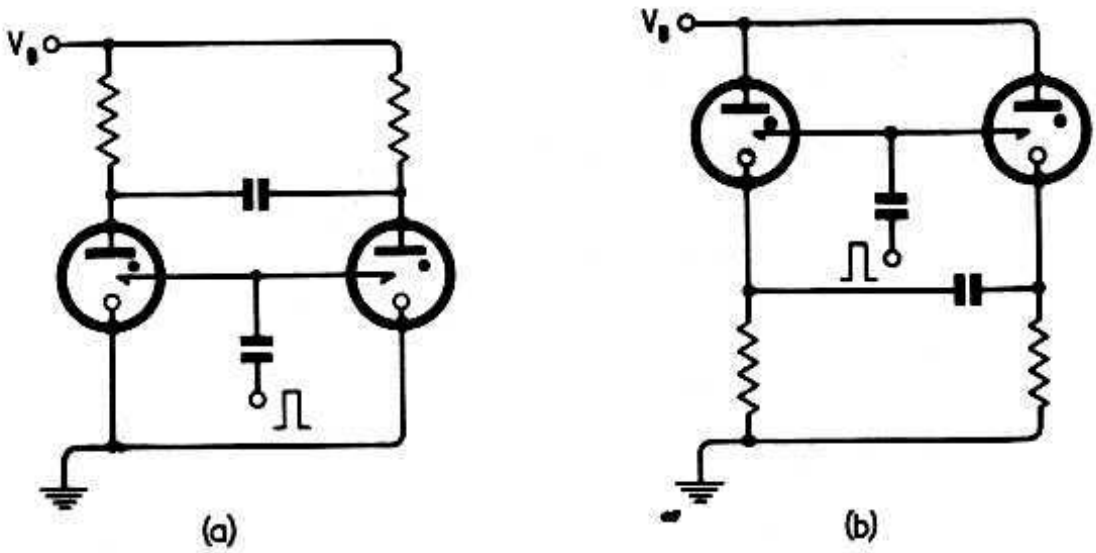


Fig. 7.38. Capacity-coupled binary counting circuits; (a) anode coupled, (b) cathode coupled

The two time constants must be selected to give the tube sufficient time to deionise. Once the anode potential of V_1 has returned approximately to V_A the circuit is ready to receive the next input pulse which will ignite V_1 and extinguish V_2. The output can be

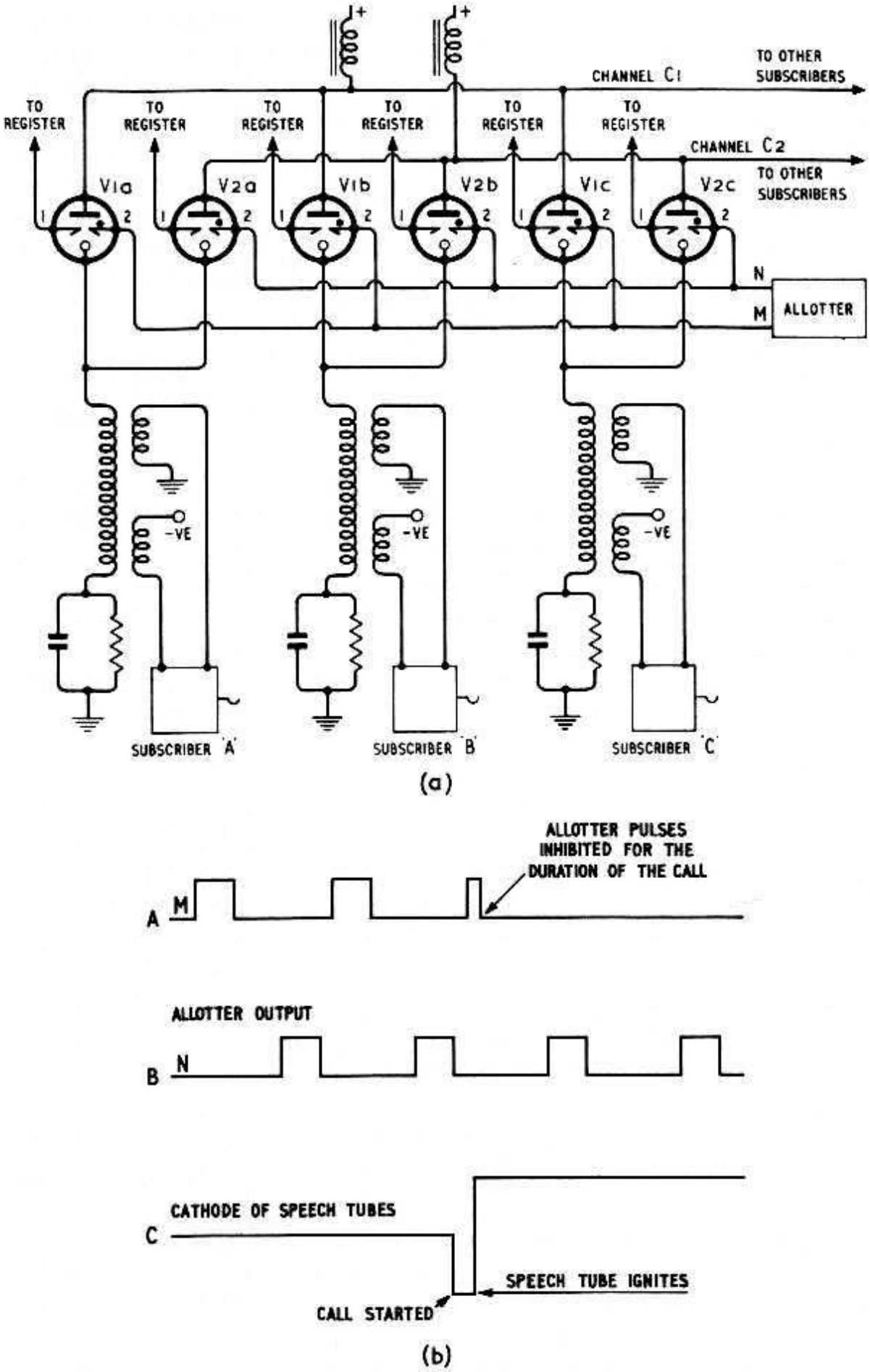


Fig. 7.39. A small private automatic telephone exchange using trigger tubes (a) basic speech passing circuit for three subscribers, (b) associated waveform in the speech passing circuit

taken from across either cathode load. The counter can be activated by a pulse of less than V_{Ts} if the cathode of one trigger tube is connected via a resistor to the trigger of the other to provide a pre-bias. This is illustrated in Fig. 7.36. Such a circuit can be adapted to construct a ring counter with as many tubes as required (for example a decade counter using ten tubes). The basic circuit for such a ring counter is given in Fig. 7.37. Suppose V_1 is conducting, then the voltage drop across R_1 raises the trigger potential on V_2 to just below V_{Ts} whilst the triggers of all the other tubes are at earth potential. Applying a pulse to the common trigger line of less than V_{Ts} will fire V_2 but none of the other tubes. V_1 will then be extinguished and a bias applied to the trigger of V_3. Thus a second pulse will move the glow from tube V_2 to V_3 and so on round the ring.

Alternative methods of connecting two tubes to form binary counting circuits are shown in Fig. 7.38, where capacity-coupled anodes or cathodes are employed. These circuits have the advantage over that of Fig. 7.35 in that they can be designed as bistable, monostable or astable circuits, whereas the common-anode circuit can only readily be designed as a bistable circuit. Further, they can be arranged to give an output pulse which is more suitable for driving a secondary binary stage than can be obtained using the circuit in Fig. 7.35.

The use of these basic logic functions, which are all achieved without the addition of other active devices is well illustrated in a small all trigger-tube private telephone exchange. A simplified diagram of the speech circuit for three subscribers is shown in Fig. 7.39(a). The principle of the exchange is that the trigger tubes act as switches in the speech path. If a tube is non-conducting, it has a very high impedance and is equivalent to an open switch, whilst once it is conducting the tube has a low effective impedance and can be regarded as a short circuit. The exchange consists of a number of common points or channels C_1, C_2 etc. to which each subscriber is connected via a trigger tube. Thus each subscriber has as many speech tubes as there are channels (two channels are shown for example in Fig. 7.39). A connection is made for instance between subscribers A and B by firing tubes V_{1a} and V_{1b}, and speech can flow between them. If at the same time subscriber C wishes to make a call to a further subscriber he must employ channel C_2 by igniting V_{2c}. From this brief description a number of logic requirements can be set out:

1. One of the subscriber's speech tubes must be ignited if:
   (a) the subscriber lifts the handset, or
   (b) another subscriber calls the subscriber in question.

2. It must not be possible to ignite a speech tube if:
   (a) the channel to which it is connected is occupied
   (b) the called subscriber is engaged.

Consider first 1 (a)—when the subscriber lifts his handset to make a call a negative bias is applied to the common cathode line, but

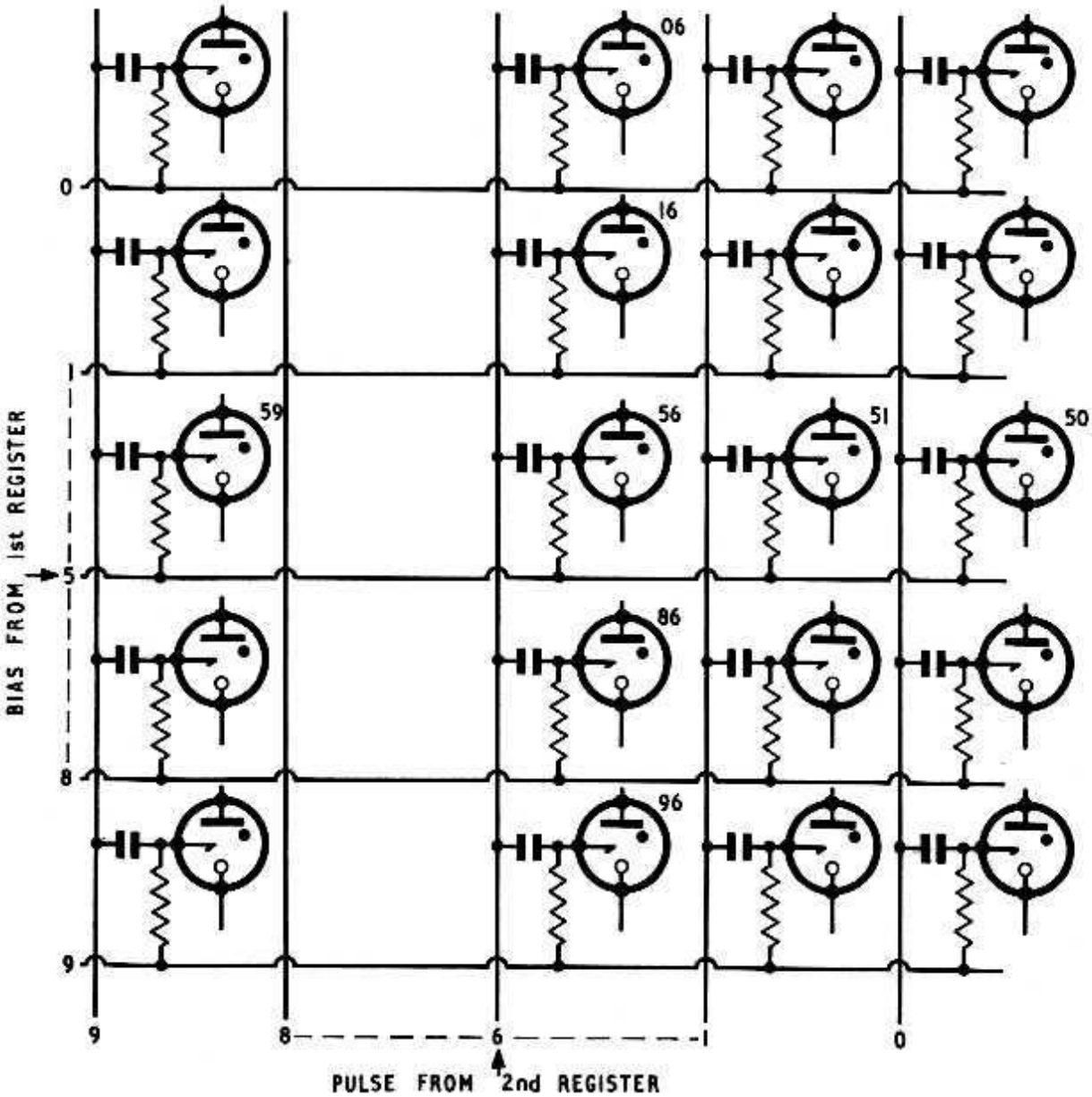


Fig. 7.40. Connections for the register triggers of the speech passing tubes of one channel in the automatic telephone exchange

which is insufficient to cause breakdown on any tube. Now the 'No. 2' triggers of all the tubes in a given channel are common (lines N and M of Fig. 7.39) and taken to a circuit called the allotter. This circuit is arranged to feed a pulse on to each of these lines in rotation as illustrated in the waveforms A and B of Fig. 7.39 (b). These pulses are again insufficient to cause ignition except when a negative bias is also present at the cathode. The sum of the two

voltages will then cause the appropriate tube to strike. This is an 'AND' gate of the type shown in Fig. 7.32 (b). Providing only one subscriber has lifted his handset at this instant, only one tube in the whole exchange will be thus energised and it will ignite. The channel selected is now occupied and in order to prevent another subscriber entering the channel the allotter pulses to that channel are immediately inhibited for the duration of the call (waveform A, Fig. 7.39 (b)).

Once a subscriber has selected a channel, say C_1, he must now, by dialling, select and ignite the speech tube connected to C_1 of the subscriber he wishes to call. In a 100-line exchange he will dial two pulse trains consisting of say 5 and then 6 pulses. These pulses are counted and recorded in register circuits of the type shown in Fig. 7.37. Now all the 'No. 1' triggers of a particular channel are formed into a matrix as shown in Fig. 7.40. On completion of the first pulse train, all subscribers between 50 and 59 will receive a bias, and on completion of the second pulse train subscribers 06, 16 etc. to 96 will receive a pulse. Only subscriber 56, however, will receive both a pulse-plus-bias which will ignite his speech tube. Again an 'AND' gate is employed, this time the type shown in Fig. 7.32 (a). Referring back to the diagram of Fig. 7.39, it will also be seen that each speech tube is an 'OR' gate since it can either be ignited by a signal received from the register on trigger 1, or by a combination of negative bias and allotter pulse at trigger 2. Finally the 'NOT' gate of Fig. 7.34 (b) is employed to prevent a further subscriber breaking into an already established call. For instance if subscriber A is conversing with B via C_1 and subscriber C calls B on channel C_2, the speech tube V_{2b} of subscriber B must be inhibited. This is achieved by allowing the conduction in the channel C_1 tube V_{1b} to cause a positive bias at the common cathode point by the RC network, see Fig. 7.39(a).

The ringing tone and engaged signal can also be generated by cold cathode tubes, when used as relaxation oscillators.

7.10.3. STABILISER CIRCUITS

Goulding[24] first pointed out in 1952 that trigger tubes with closely controlled trigger breakdown could be used in self-quenching circuits to provide voltage stabilisation. The basic circuit is shown in Fig. 7.41. Essentially a fraction $R_2/(R_2+R_1)$ of the output voltage is applied to the trigger electrode. If V_{out} exceeds $V_{Ts}\,(R_1+R_2)/R_2$ the tube will ignite and, providing R_a and C_1 are chosen so that the anode–cathode circuit is self-quenching, a series of current pulses

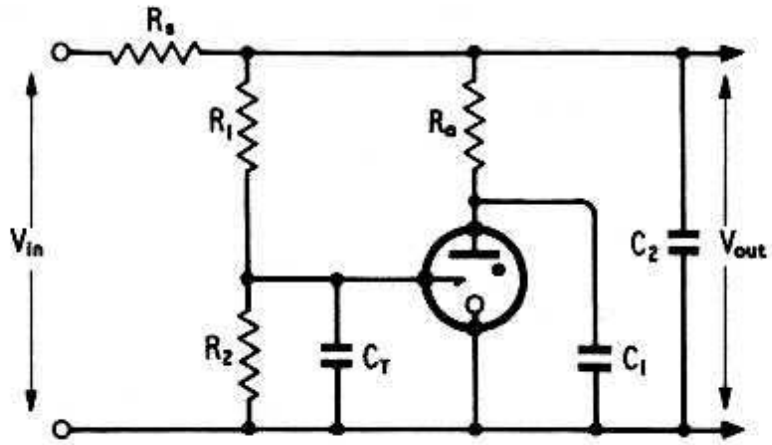
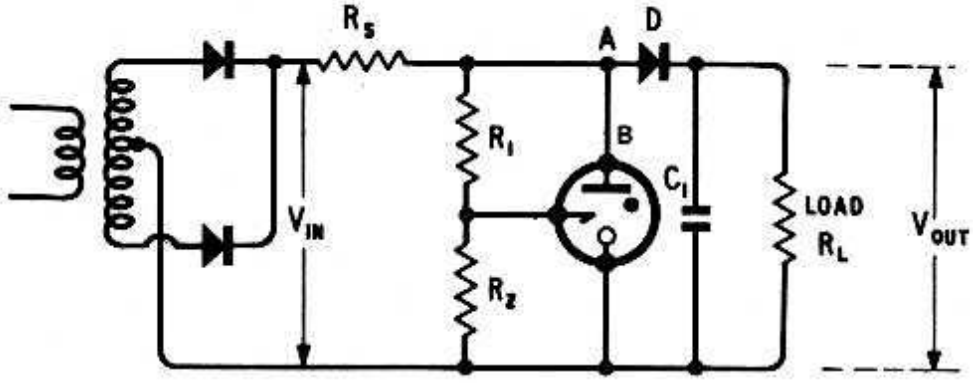
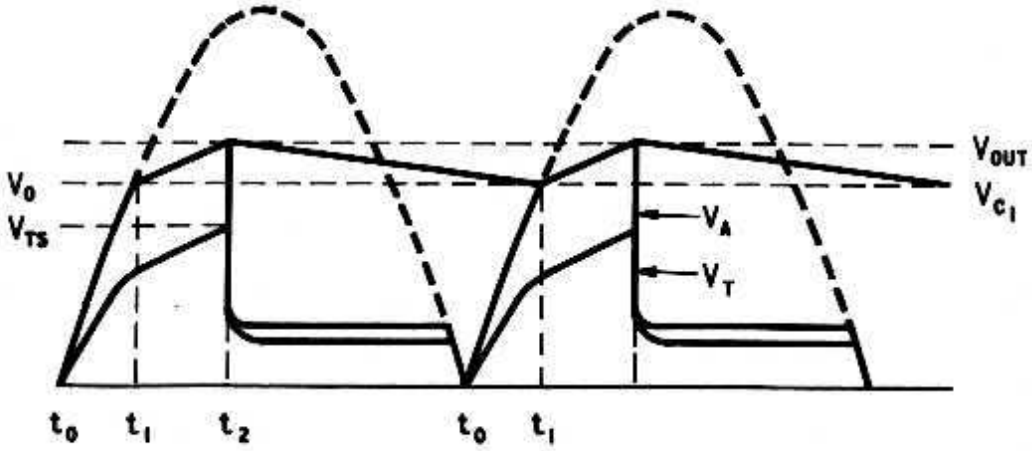


Fig. 7.41. Basic voltage stabilisation circuit operated from a d.c. supply



(a)



(b)

Fig. 7.42. Voltage stabilisation circuit for currents in the mA range operated from a.c. (a) basic circuit (b) voltage waveforms

will be conducted through the tube causing C_2 to discharge until V_{out} has fallen to $V_{Ts}(R_1+R_2)/R_2$. Providing the input current is greater than the output current, the tube will stabilise at V_{out} slightly in excess of $V_{Ts}(R_1+R_2)/R_2$. The advantage of the trigger tube stabilising circuit over diode stabilisers is that higher voltages can be obtained, and also the output voltage can be altered by varying the values of R_1 and R_2.

The above circuit operates from a d.c. supply and is limited to currents of the order of tens of microamps. An alternative circuit working from a.c. or unsmoothed rectified a.c., which can stabilise with load currents in the mA range, is shown in Fig. 7.42 (a).[25, 26] Fig. 7.42 (b) shows the waveforms of this basic circuit; at t_o, with the tube extinguished, V_A follows the positive-going input voltage whilst V_T will be $R_2/(R_1+R_2)\ V_{in}$. At t_1 the potential V_o of the reservoir capacitor C_1 is reached; the diode D is then opened and C_1 is charged via the resistor R_s. When V_T reaches V_{Ts} at t_2, trigger breakdown followed by anode breakdown occurs; this is arranged by suitably choosing R_1 and R_2 so that V_{Ts} is not reached before V_A is within the anode voltage operating range of the tube. The voltage of the anode at which this occurs is the maximum value to which C_1 is charged. When the tube becomes conducting the anode potential falls to V_{Am} and the diode D is blocked off. The output circuit is now isolated from the trigger tube circuit and C_1 commences to discharge through the load until t_1 in the following positive half-cycle. Meanwhile at the end of the half-cycle the trigger tube is extinguished. The circuit therefore stabilises at the maximum voltage reached by the anode before breakdown which depends only on R_1, R_2 and V_{Ts}.

$$V_{out} = \frac{R_1+R_2}{R_2} V_{Ts} \tag{7.10}$$

The ripple on the output is dependent on C_1 and the load, and, since the mean output voltage falls as the ripple increases, regulation will also depend on these and only to a lesser extent on the input voltage.

There is a limit to the value of C_1, since it must be fully charged during the charging period, and R_s, in the basic circuit, has a lower limit set by the maximum tube current. An improvement can be made by inserting an anode resistor R_a in the anode lead (across AB in Fig. 7.42a). However, there is still a limitation since the point A must not rise above the breakdown potential when the tube is conducting i.e.

$$i_a R_a + V_m < V_{out}$$

or, in terms of R_s and approximate potentials,

$$R_a < R_s \frac{V_{out} - V_m}{V_{in} - V_{out}} \tag{7.11}$$

R_s is a more fundamental variable than R_a, since it determines the operating range of current in the circuit when considered in conjunction with V_{in}, V_{out} and V_m. The interdependence of these factors and the determination of circuit values are discussed by Neale[5] and Atkinson and Hill [26]. In general, trigger-tube stabilisation is not recommended where the highest stability is required. On the other hand a stability of $\pm 2\%$ up to a current of 20 mA, at voltages from 200 V to 350 V is fairly easily obtained. To obtain higher voltages two such circuits can be stacked together to obtain a stabilised output from a voltage doubler. Oxley,[27] however, describes a circuit in which the output of a voltage doubler is stabilised by a single trigger tube. This is made possible by earthing the anode, a change which also makes extra smoothing of the output possible.

REFERENCES

1. LOTHALLER, W. E. and VAN VLODROP, P. H. G., *Electron. Appl.*, **23**, 89, (1962/3).
2. WRIGHT, D. H., *Symposium 'Cold Cathode Tubes and their Applications'*, Cambridge, (1964) [I.E.R.E. conference proceedings **Number 3**]
3. DOUGLAS, A., *Electron. Eng.*, **31**, 672, (1959)
4. V. DE KERCKHOFF, H., *Symposium 'Cold Cathode Tubes and their Applications'*, Cambridge, (1964) [I.E.R.E. conference proceedings **Number 3**]
5. NEALE, D. M., *Cold Cathode Tube Circuit Design*, Chapman & Hall, London, (1964)
6. FLOOD, J. E. and WARMAN, J. B., *Electron. Eng.*, **28**, 528, (1956)
7. RAPHAEL, M. S. and ROBINSON, A. S., *Electronics*, **29**, 162, (July 1956)
8. INGRAM, S. B., *Bell Lab. Record*, **15**, 114, (December 1936)
9. TOSSWILL, C. H., *Philips Tech. Rev.*, **18**, 128, (1956/7)
10. WESTON, G. F., *Proc. Inst. Elec. Engrs.*, **111**, 1546, (1964)
11. ANDEWEG, J., KERR, G. and POORTER, T., *Electron. Appl.*, **25**, 45, (1964/5)
12. GIMSON, K. F. and CROWTHER, G. O., *Electron. Eng.*, **29**, 536, (1957)
13. KERR, G., *Rad. Electron. Engr.*, **28**, 207, (1964)
14. HOUGH, G. H. and RIDLER, D. S., *Electron. Eng.*, **24**, 152, (1952)
15. DOMBURG, J. and SIX, W., *Philips Tech. Rev.*, **15**, 265, (1954)
16. CROWTHER, G. O. and SMITH, J., *Electron. Eng.*, **33**, 728, (1961)
17. SMITH, J., *Proc. Intern. Conf. Ionisation Phenomena Gases, 5th*, Munich, N. Holland Publishing Co., Amsterdam, **Vol. I**, 449, (1962)
18. BOND, M. E., *Electron. Eng.*, **34**, 798, (1962)
19. BOND, M. E., *Symposium 'Cold Cathode Tubes and their Applications'*, Cambridge, (1964) [I.E.R.E. conference proceedings **Number 3**]
20. CROWTHER, G. O., *Mullard Tech. Commun.*, **3**, 234, (1957/8)
21. LIGHT, L. H., *Mullard Tech. Commun.*, **2**, 37, (1955/6)
22. HERCOCK, R. J. and NEALE, D. M., *Proc. Inst. Elec. Engrs., Part II*, **99**, 507, (1952)

23. DURAND, M., 'Les Thyratrons a Cathode Froide', *Tube Electroniques*, **Vol. XVII,** Philips Eindhoven, (1963)
24. GOULDING, F. S., *Electron. Eng.*, **24,** 493, (1952)
25. KERR, G. and VAN VLODROP, P. H. G., *Electron. Appl.*, **21,** 25, (1960/61)
26. ATKINSON, E. P. T. and HILL, C. F., *Electron. Eng.*, **37,** 788, (1965)
27. OXLEY, A. J., *Electron. Eng.*, **36,** 548, (1964)

CHAPTER 8

Stepping Tubes

8.1. General Principle of the Stepping Tube

The stepping or counting tube is a more recent device than either the stabiliser or trigger tube, the first tubes being commercially available at the end of the 1940s.

Essentially, within the tube, a glow discharge can rest in one of several stable positions, and is caused to transfer from one stable position to another in a predetermined sequence by means of an input signal. Normally the tube contains a common anode and ten cathodes giving the stable positions or discharge gaps. Since the discharge to each of these cathodes should ignite and extinguish under identical conditions, the cathodes are of the same geometry and symmetrically disposed with respect to the anode; they are usually positioned in a circle around an anode disk. With such a geometry the discharge moves round one position in the circle with each input signal so that one revolution registers a count of ten. An electrical output can be taken from a resistor in series with an appropriate cathode and/or the position of the glow can be detected visually through an end window. The tube can therefore be used in any decade scaling or counting application not requiring very high speeds. It has the advantage over trigger-tube scalers in requiring a smaller number of components with the consequent lower cost and higher inherent reliability.

The operation of the tube depends on the fact that the discharge in the conducting gap directly influences adjacent gaps by lowering their breakdown potential, i.e. the adjacent gaps are primed in a similar way to that in which the anode-cathode gap is primed by the trigger discharge in a trigger tube. The priming ensures that adjacent gaps are more likely to ignite than those further away if the potentials of several of the non-conducting cathodes, relative

to that of the current carrying cathode are lowered simultaneously. However, in order that the discharge should transfer in some definite sequence, for example clockwise round the circle, it is necessary for the degree of priming to be asymmetric about the conducting cathode. This can be achieved by suitably shaping the cathodes so that the discharge is invested on a preferred section, for example close to the next cathode in the sequence.

It is possible by suitable external circuitry to effect transfer of the discharge using just the ten shaped cathodes,[1] but in general the directional properties become more reliable, and the circuit simpler, by introducing transfer cathodes between the main cathodes, to which the input signal is applied. Several designs of decade counting tube based on a combination of shaped cathodes and a single transfer cathode between the main cathodes, were described in the literature around 1950.[2–5] The most successful tube, however, designed at about that time was one in which the electrodes were not shaped, but in which two transfer or guide cathodes were interposed between each pair of adjacent main cathodes. Transfer in the correct direction was effected by two consecutive pulses applied to the respective guide electrodes. The tube, described by Bacon and Pollard,[6] consisted of thirty rod cathodes (one main or rest cathode and two guides for each digit) equally spaced around a disk anode with their lengths perpendicular to the anode plane (Fig. 8.1 (a)). It was given the proprietary name of 'Dekatron'. This design of tube with rod cathodes has survived almost unchanged and is the commonest type in use today. Several versions are available having slight differences in construction and gas filling, and are marketed by more than one manufacturer. The principle of operation of this two-pulse stepping tube is described in more detail in the next section.

Initially the counting speeds of such tubes were limited to a few kc/s. By choosing a gas filling with an admixture of hydrogen the speed has been increased some tenfold (up to 50 kc/s), but it became clear that no further improvement in counting speed could be expected from a double-pulse transfer system.

In 1952 Acton[7] designed a tube based on a similar construction that could be driven by a single pulse. The tube had an extra guide cathode, making forty rod cathodes in all, and was filled with a helium–hydrogen gas mixture. Although it represented a higher speed tube at that time, it has an upper limit of only about 20 kc/s. However, it has advantages in some applications and is still available commercially. More recently, single-pulse tubes of smaller dimensions have been designed[8] capable of speeds up to 1 Mc/s. The tubes have single transfer cathodes and both rest and transfer

cathodes are 'overlapped' to give the preferred counting direction.

Another recent design has arisen from the desirability of directly coupling the output from the stepping tube to a numerical glow-discharge display tube without intermediate amplification. The tube, described by Reaney,[9] is a modification of the Dekatron in which ten auxiliary anodes are introduced.

Although the basic principles of the various designs are similar, each tube type has a slightly different mode of operation requiring different circuit arrangements. However, the double-pulse stepping tube is by far the most widely used today, and also the one which has been most thoroughly investigated. Therefore, most of the information in this chapter will be devoted to this type, and only brief reference made to other designs.

8.1.1. THE PRINCIPLE OF OPERATION OF THE DOUBLE-PULSE STEPPING TUBE

In the thirty cathode construction illustrated in Fig. 8.1 (a) the rest cathodes $K_o, K_1, \ldots K_9$ are either taken out separately for electrical connections (selector tubes) or K_o is connected separately and K_1, $K_2, \ldots K_9$ are connected to a common lead-out (counting tube); the latter arrangement allows the assembly to be mounted on a conventional 8-pin base. The relative guide cathodes are all interconnected, i.e. all the first guides, guides *A* or *GDA*, in the forward counting direction are connected together and all the second guides, guides *B* or *GDB*, are interconnected. To understand the principle of operation, consider the three digit positions K_o, K_1, and K_2 with the related circuit shown in Fig. 8.1 (b); the subscripts on the guide electrodes are designated by the rest cathode they follow. In this circuit the rest cathodes are at earth, and the guides positively biased by 40 V. A supply voltage of about 500 V is applied to the anode via R_a which limits the current to a few hundred microamps, so that about half of the rod above the ceramic spacer is covered by the cathode glow. Assuming the glow is invested on K_1, then the anode will be at the maintaining potential ($\sim$ 190 V for an inert-gas filled stepping tube). The discharge to K_1 will prime the other cathodes in the tube, reducing their breakdown potential. The degree of priming will depend on the distance from K_1, so that the breakdown potential of the cathodes will be greater the further they are from K_1, a typical curve of the breakdown potentials on neighbouring cathodes is shown in Fig. 8.2. If now the potential of *GDA* is lowered, a value, which is negative with respect to the potential of K_1, will be reached where breakdown occurs between GDA_1

and the anode. Due to the effect of the priming this will occur some twenty volts before a discharge can be ignited on GDA_o, or some forty volts before a discharge can be ignited on any other guide A cathode (see Fig. 8.2). The glow discharge to GDA_1 will result in an increase in current drawn to the anode through R_a, thus lowering

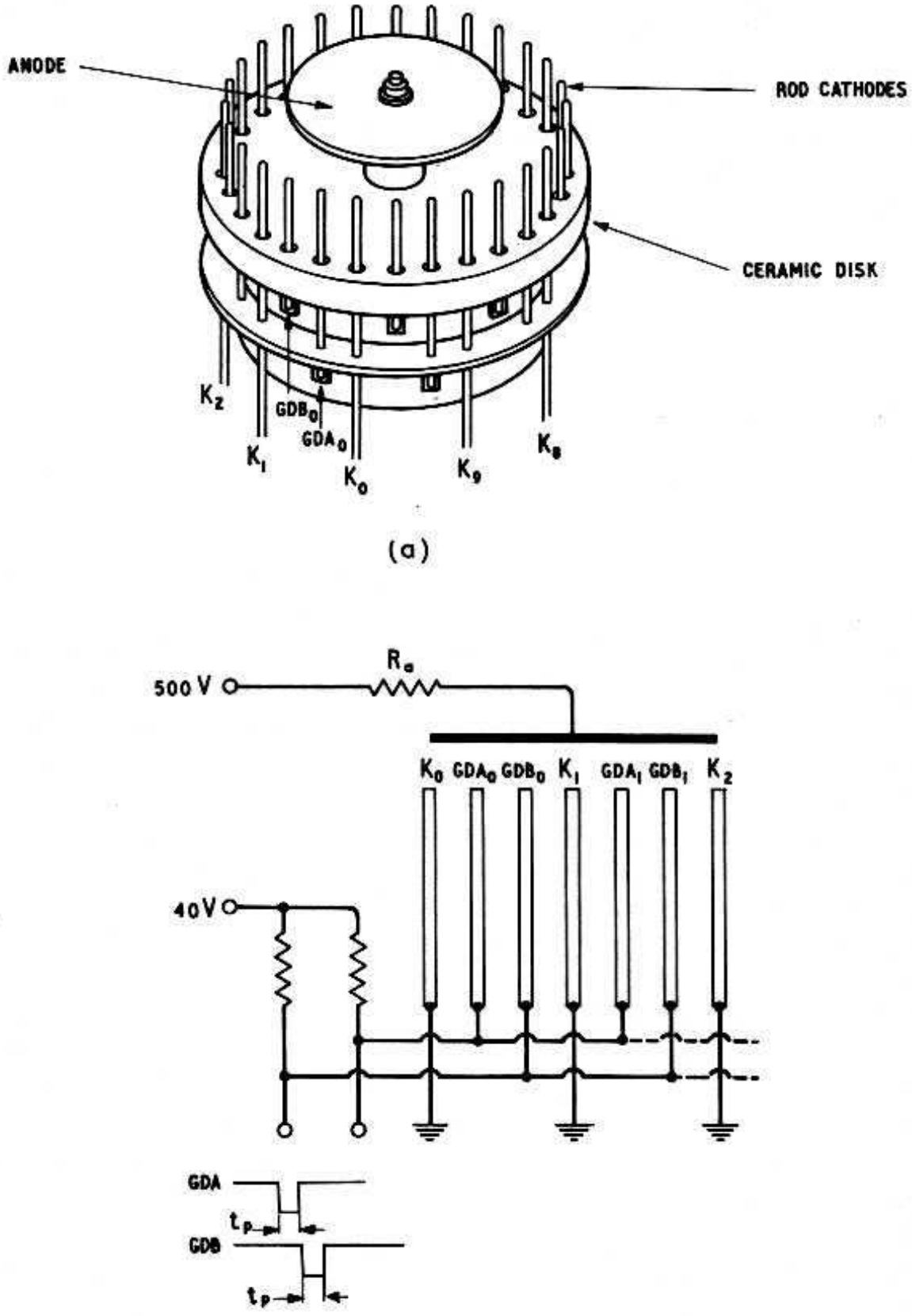


Fig. 8.1. The two-pulse stepping tube (a) basic design of the electrode structure (b) circuit for a three-digit count

the anode potential below V_m, and K_1 will be extinguished. After a time t_p, when the K_1-anode gap has sufficiently recovered as a result of deionisation, the potential on GDA can be raised again, but not above earth lest K_1 should restrike. If, however, at the same time as the voltage on guides A is raised the potential on guides B is lowered, then, because the breakdown potential distribution will now be symmetrical about the minimum at GDA_1,

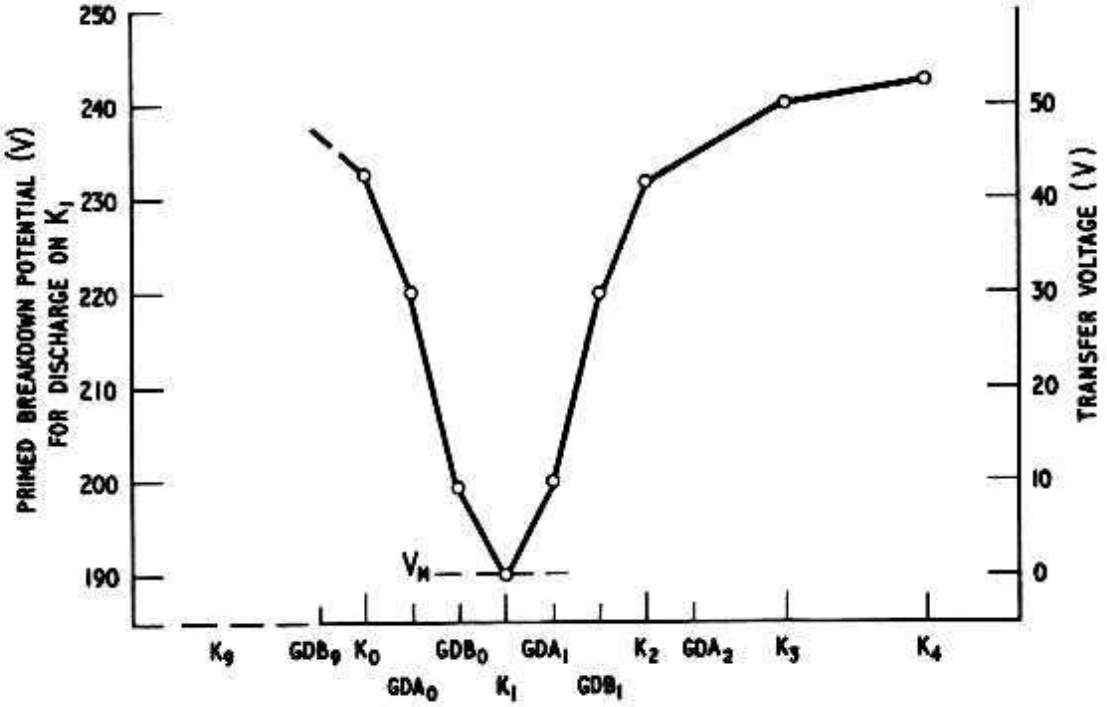


Fig. 8.2. Transfer characteristic for a two-pulse stepping tube, expressed as the breakdown potentials between anode and the various cathodes, when a discharge current of 340 μA is passing to K_1

GDB_1 will ignite, with the subsequent extinction of GDA_1. Again after time t_p the voltage on guides B can be raised. GDA_1 is now at +40 V whilst K_2 is at earth, thus when GDB_1 returns towards the 40 V bias line the K_2-anode gap will ignite in preference to all other anode-cathode gaps. According to Fig. 8.2 this will occur when the anode potential has risen to 200 V, i.e. when GDB is 10 V positive. The discharge will therefore have transferred from K_1 to K_2 by the application of sequential negative pulses to guides A and B. Clearly by repeating the pulses the discharge will move on from K_2 to K_3 and so on, the position of the glow registering the number of applied double pulses. If the pulses were applied in the reverse order i.e. the first pulse to guides B and the second pulse to guides A, then the discharge would transfer in the opposite direction, from K_1 to K_o. Thus the tube can be used as a bi-directional counter.

The duration of the pulses, i.e. the time t_p, is determined by the deionisation or recovery time of the previous gap, and sets the

limit to the counting speed. However, since adjacent gaps are primed by the discharge the degree of deionisation is not as high as required by a trigger tube, and speeds of 4 kc/s can readily be achieved even with an inert gas filling. The pulse amplitude is related to the pulse duration; the larger t_p is, the smaller can be the amplitude. However, there is a lower limit set by the d.c. characteristics of the tube, illustrated by the breakdown voltage distribution to the cathodes, Fig. 8.2.

In the above illustration of the mechanism square-wave pulses are implied, but this is not essential and other pulse shapes can be employed. The circuit requirements are discussed in Section 8.3. An output of about 20 V can be obtained by connecting a resistor in the required cathode lead.

Normally these stepping tubes have ten rest or main cathodes and are suitable for decade scaling and computing, however tubes with twelve main cathodes are also available in the U.K.

8.1.2. PRINCIPLE OF OPERATION OF OTHER TUBE TYPES

In the single-pulse 'Dekatron' in which there are four cathodes per digit,[7] the mechanism can be compared to that of a tube having a single transfer cathode where all the cathodes are shaped. The directional property, equivalent to shaping the cathode, is achieved by connecting adjacent cathodes in pairs via a *CR* network. This is illustrated in Fig. 8.3 where three digit positions are shown.

Assuming the discharge is on K_1, then to transfer the discharge, the potential at *G* is lowered below zero. This lowers the potential on both the guide *A* and *B* cathodes, and, as a result of the priming effect of the discharge, GDA_1 will ignite, and K_1 will be extinguished. However, because of the resistance R_1 (of the order of 200 KΩ) the potential on GDA_1 will rise as the capacitor C_1 charges up, and the discharge will transfer automatically to GDB_1 (since *GDB* is then at a lower potential than K_1). If the potential at *G* is now raised the discharge will transfer to the nearest earthed cathode which is GDC_1. Again the capacitor C_2 charges up and automatic transfer to K_2 is accomplished. Thus the discharge has moved from K_1 to K_2 by the application of a negative pulse to *G*. To transfer in the opposite direction the connections to *GDA* and *GDC* must be interchanged.

The cathode pairs of the single-pulse 'Dekatron' (*GDA* plus *GDB* or *GDC* plus *K*) could be replaced by a single cathode of equivalent width if the discharge could be made to transfer automatically from one edge to the other. This can be achieved by

several methods, but most conveniently by exploiting the hollow cathode effect (cf. Section 3.6). The discharge to a hollow or concave cathode surface is more efficient than to a plane surface and a discharge will prefer to go to a hollow area. Thus if a cathode is fabricated with a plane surface in contact with a hollow region, a glow

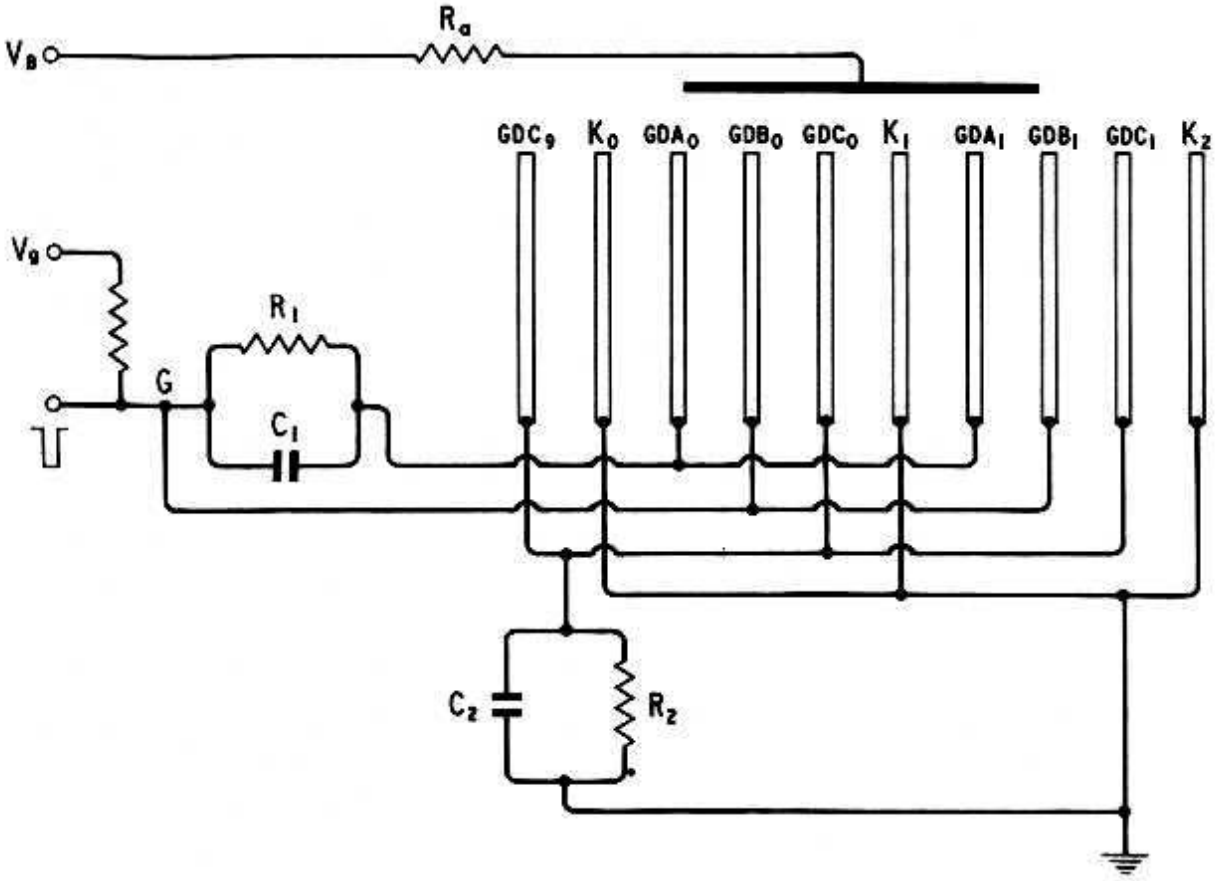


Fig. 8.3. Circuit for a three-digit count in the single-pulse Dekatron

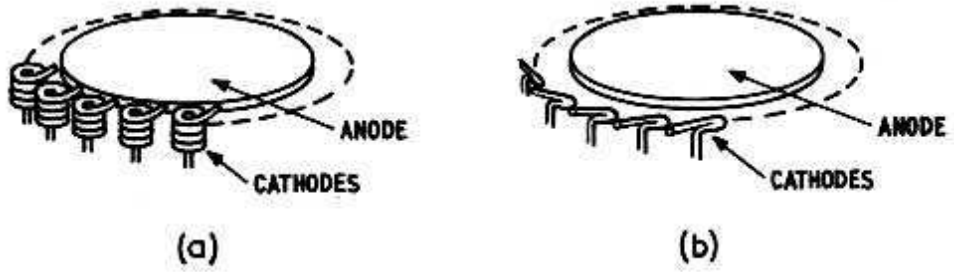


Fig. 8.4. Schematic diagram of electrode structures, using shaped cathodes, suitable for single-pulse operation: (a) the 6167 described by Peck,[4] (b) the EZ10 described by Apel[8]

discharge initially invested on the plane surface will transfer into the hollow area. A practical embodiment of the effect in a counting tube is illustrated in Fig. 8·4(a), which shows the essential electrode construction of a tube described by Peck.[4] In the static condition the glow rests in the hollow formed by the inside surface of the spiral. The tail of the next cathode in the counting direction is

therefore closest to the discharge and will have a breakdown potential lower than the adjacent cathode in the opposite direction. On application of appropriate potentials, the glow will transfer first to the tail of the next cathode, but will then move along into the hollow thus priming the following cathode. In a more recent design Apel[8] obtains the directional properties by tilting the cathodes so that the cathode in the forward direction is immersed in the discharge whilst that in the reverse direction is outside the discharge. This is illustrated in Fig. 8.4(b). The hook shape ensures that the discharge goes to the top of the cathode and not to the support. To ensure that the discharge does not step backwards as a result of residual ionisation alternate rest cathodes can be connected together (i.e. odds connected and evens connected), and a parallel *RC* combination connected in each of the two common leads. Thus at the end of the transfer pulse there is a residual bias on the previous cathode. Under these conditions counting rates up to 1 Mc/s can be achieved.

By connecting alternate digit cathodes and also alternate guide cathodes, Apel[10] showed that a tube could be stepped reliably using a single transfer cathode even if the cathodes were not directionally shaped or overlapped. The circuit arrangement is shown in Fig. 8.5 (a). The alternate digit cathodes are driven by two antiphase rectangular output voltages from a bistable multivibrator. These voltages are also applied to the guide cathodes, via differentiating circuits C_1R_1 and C_2R_2. The waveforms are shown in Fig. 8.5 (b). Assuming that K_1 is conducting, then all neighbouring cathodes will be positively biased. An input to the multivibrator switches it into its second stable position. The voltage on K_1 and all other odd cathodes rises whilst that on the even cathodes falls. At the same time a positive pulse is applied to GD_o whilst GD_1 receives a negative pulse. The combination of the negative pulse to GD_1 and the drop in voltage on K_2 ensures that the discharge transfers from K_1 to K_2 and not in the opposite direction. The counting direction can be reversed by interchanging the connections to the odd and even guides. Reliable counting at speeds up to 1 Mc/s are claimed for the hydrogen-filled subminiature tube ECT 100 designed to operate with this type of circuit. The output is taken from auxiliary electrodes placed near the digit cathodes. The tube is not suitable for visual read-out.

By the use of auxiliary output electrodes higher powers can be taken, and the GCA10G has been specifically developed using this technique to provide sufficient output to drive a numerical display tube.[9] The tube has essentially the same structure as the double-pulse stepping tube, and uses the same transfer system, but it has

ten auxiliary anodes positioned between the rest cathodes and the central main-anode disk. A study of the *i-V* characteristics of these auxiliary anodes, which act as probes in the discharge, shows that conduction to the auxiliary anode associated with the cathode which is glowing occurs some 40 V before conduction to the adjacent auxiliary anodes. Thus an output of 40 V can be taken from the auxiliary anode with currents of the order of 2 mA. Further, the

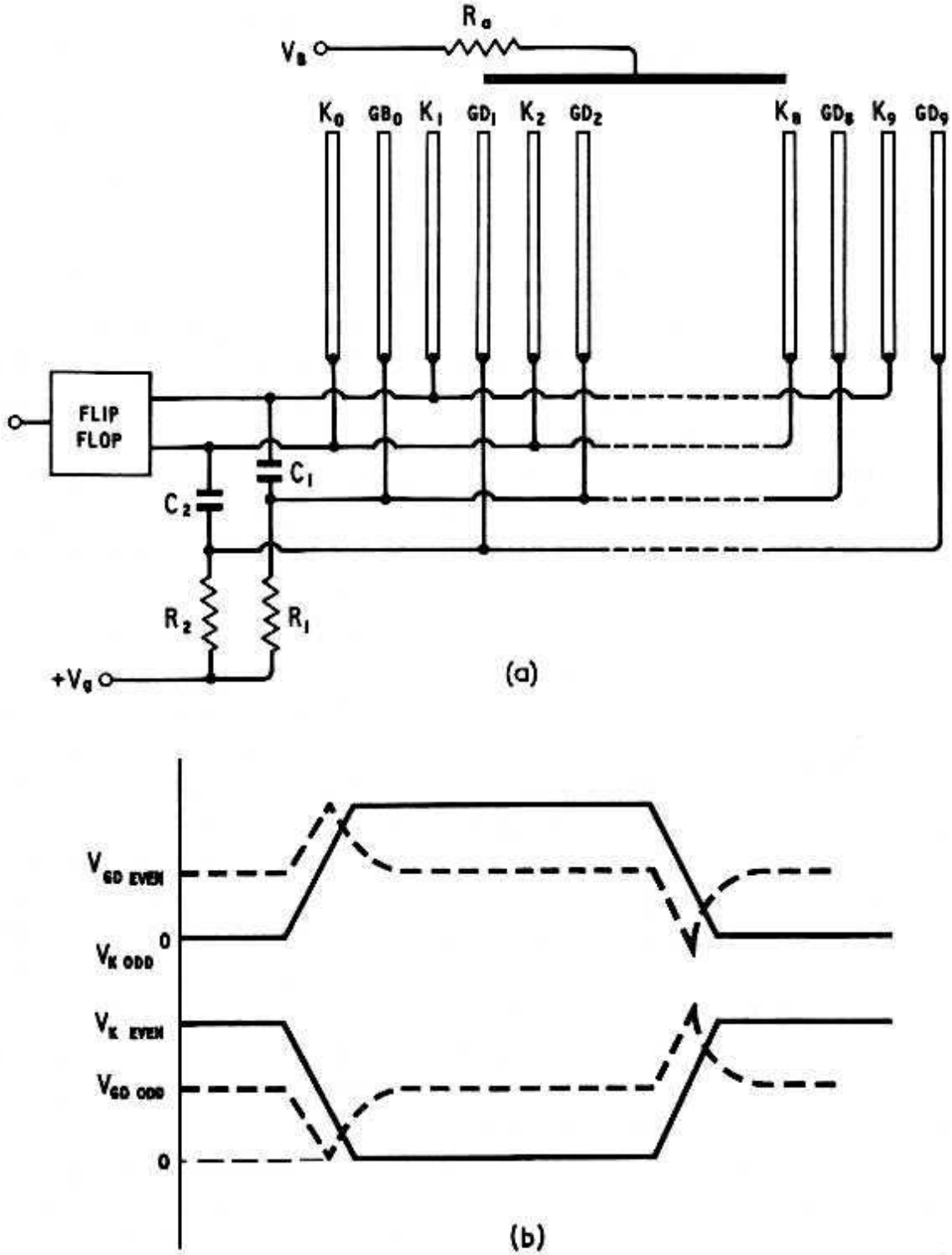


Fig. 8.5. (a) Drive circuit for the ECT*100, a stepping tube with a single transfer cathode and without directional-shaped electrodes. (b) Voltage waveform at the cathodes*

anode can make large voltage excursions without upsetting the transfer mechanism. Coupling of the tube to a numerical indicator is described in the next chapter (Fig. 9.25).

8.2. The Design of a Stepping Tube

A stepping tube is considerably more complex than a trigger tube, and when considering the tube design a number of requirements have to be taken into account, some of which are conflicting in terms of glow-discharge parameters. The main requirements for a stepping tube are as follows:

1. The tube should step with an input signal which is as low as possible.
2. The stepping should be reliable—there should be no possibility of transference in the wrong direction, of failure to step, or of stepping more than one position per input pulse.
3. The tube should allow for reasonable tolerances on the circuit components and supply voltages. In the interests of economy the tube and circuit should be simple.
4. The tube should give a clear visual output of the digit position. (This may not be essential in some applications.)
5. It should be possible to take an output from the required digit of comparable amplitude to the input signal.
6. The tube should be capable of functioning at any speed up to its maximum over a useful life.

If the requirements (1) and (5) are considered, then ideally one would like the input signal to be less than the tube output, so that tubes could be connected in cascade to form a counting chain without the need of an intermediate amplifier. Although tubes have been constructed where this is just possible, it cannot be attained readily in a relatively simple design. This is partly because the desirable parameters for a low input signal conflict with the second requirement for reliable stepping (see Section 8.2.1) and in general the latter requirement is of more importance.

Similarly there is a conflict between reliability and the need for visual output. A factor important to reliability is the symmetry of the tube, not only in the accurate positioning of the electrodes, but also in the breakdown and maintaining potential of all the discharge gaps; i.e. the second Townsend coefficient γ should not vary from cathode to cathode. As has already been pointed out (Section 6.3.2) the only satisfactory method of achieving reproducible and stable values of γ is to sputter the cathode thoroughly and to cover the envelope with an opaque deposited layer. In the case of the stepping

tube this is not possible since a viewing window in the glass envelope must be provided.

Thus in the design of a stepping tube compromises have to be sought to provide the most satisfactory tube, and not all the requirements can be met. Fortunately the imperfections of the stepping tube can be overcome by suitable circuits in most applications, and in a later section the circuit design will be discussed in some detail.

8.2.1. THE GAS FILLING

The gas pressure and composition are chosen mainly for the 'transfer' characteristics of the tube. The term here is used in a rather different sense to that applying to trigger tubes, since in the stepping tube one is interested in the primed breakdown potential of the various cathode-anode gaps as a function of their position relative to the conducting cathode, and the current in the conducting gap is of less importance. Thus Fig. 8.2 represents the transfer characteristics for a stepping tube. Alternatively the difference between the breakdown potential of the various cathode–anode gaps and the maintaining potential of the conducting gap can be plotted. This voltage difference represents the minimum voltage relative to the conducting cathode which must be applied to any given cathode to transfer the glow on to that cathode. It is referred to as the transfer voltage and is given by the righthand scale on Fig. 8.2.

Since the A and B guide electrodes are respectively connected to common leads so that corresponding guides follow the same voltage excursions, ideally the transfer voltage for adjacent cathodes (for example GDA_1 if the discharge is on K_1) should be zero or very small, whilst that for the next-but-one cathode (GDA_o for K_1 conducting) should be high. Unfortunately such an idealised characteristic is not possible, nevertheless the curve of Fig. 8.2 can be varied considerably by altering the gas pressure and composition. For satisfactory functioning of the stepping tube the transfer characteristics must be neither too 'flat' giving poor discrimination between the corresponding guides in the two directions, nor too steep requiring large input pulses to step the tube. The voltage limits of the curve are set by the maintaining potential at the lower end, and by the unprimed breakdown potential at the upper end, and clearly the difference between these two potentials must be large relative to the spread between individual values for a success-

ful stepping tube. In practice $V_s - V_m$ is arranged to be greater than 100 V which restricts the choice of gas, making Penning mixtures particularly unsuitable (cf. Section 6.3.1). It also requires that the pressure times the electrode spacing is high enough to give a breakdown potential well above the Paschen minimum.

However, a more important factor affecting the transfer characteristics is the negative glow thickness, since this is a measure of the area of influence of the discharge as a priming agent. The glow thickness is related to the cathode fall distance d_K, which from Section 3.2 was seen to follow the similarity principles, the product pd_K being directly related to the cathode fall voltage V_c, and inversely related to the current density (both of these are dependent on the gas composition). The current density appears to show the greatest influence when considering different gases; for instance, in the case of Penning mixtures, although V_c is lower than for the main constituent gas, the current density in the normal glow at the same pressure is also much lower so that d_K is relatively large. Thus the glow is more extensive for Penning mixtures than for pure gases, giving a flat transfer characteristic around the minimum. It is therefore difficult to obtain adequate discrimination between the corresponding guides in the two directions in Penning mixtures even with 10-20% admixtures of argon, unless pressures above 100 torr are employed, and then it is normally difficult to clean the cathodes.

The pure gases would be suitable for stepping tubes from the above considerations, especially neon, since satisfactory transfer characteristics at a reasonable pressure are combined with good visual indication. However, the counting speed with pure neon is limited to a few hundred cycles per second due to the lengthy deionisation time. The most favourable inert gas mixture in terms of deionisation time is obtained by admixing a small percentage of neon (2–5%) in helium, see Section 5.4.2. Fortunately the helium-neon mixture also gives good transfer characteristics in a stepping tube at pressures of the order of 40–80 torr, and has a good visible glow. It is, therefore, the most suitable inert-gas mixture for stepping tubes, being used in the majority that are capable of counting up to 5 kc/s. For higher counting speeds inert gas–hydrogen mixtures are used. These can be helium-hydrogen or neon-hydrogen. The former, although claimed to have the faster deionisation time, gives rather poor light output, having a pale mauve negative glow. Neon-hydrogen on the other hand, gives the bright orange glow of the neon and for this reason is more favoured. Although hydrogen is an active gas which 'cleans up' more readily than most gases, especially in the ionised state, by suitable pro-

cessing to saturate the electrode structure and glass envelope, lives comparable to those of inert-gas filled stepping tubes can be attained.

8.2.2. THE ELECTRODE MATERIALS AND GEOMETRY

Since the anode material plays no part in the discharge it is chosen for its mechanical properties and cost; nickel is a suitable material and is commonly used.

The cathode material, on the other hand, affects the breakdown and maintaining potentials through the second Townsend coefficient γ, and therefore must have good electrical as well as mechanical properties. Because the spread in breakdown and maintaining potentials from one anode–cathode gap to another must be relatively small if the transfer characteristics around the tube are to be consistent, coated cathodes are quite unsuitable for stepping tubes. Only pure metal cathodes which can be cleaned by sputtering are satisfactory, although because of the need for visual output the cleaning of the cathodes cannot be as thorough as for stabilisers and trigger tubes. From the discussions of cathode materials for stabilisers and trigger tubes (Sections 6.3.2 and 7.8.2) it might be supposed that molybdenum would prove the best material for the rod cathodes. However, with inert gases this is not so, for although good initial characteristics can be obtained with molybdenum cathodes, they do not behave well on life, particularly when the glow is invested on one cathode for any length of time. The reason is that the discharge to the conducting cathode appears to contaminate adjacent cathodes and is manifest in an increase in the breakdown potential of the adjacent cathodes (i.e. a decrease in γ) with time, raising the transfer voltage until the tube fails to step on the application of the input pulse. The effect is analogous to the contamination of a trigger tube cathode when inverse trigger current is taken. It is not clear why molybdenum should be more susceptible to 'adjacent cathode contamination' than other metals such as nickel commonly used in discharge tubes.* It has been suggested that the actual sputtered molybdenum deposited on to the molybdenum cathode will change the γ of the surface, but this is difficult to understand in view of the fact that the cathode surface finish has little effect on the discharge characteristics.

*In the 1 Mc/s tube ECT 100 molybdenum cathodes are used with hydrogen gas. Very little sputtering takes place, and adjacent cathode contamination does not appear to be a problem. Apel[10] carried out life tests up to 10,000 hrs. with the glow resting on one cathode without affecting the counting behaviour.

Another possibility is that gases are released by the ion bombardment of the conducting cathode which react with the other cathodes in close proximity to form a surface layer, and that in the case of molybdenum either more gas is released or the reaction has a greater effect on the characteristics.

Adjacent cathode contamination occurs with all cathode materials and is the main cause of tubes miscounting. It is, however, less marked with nickel, where the tube can be run with the discharge stationary on one cathode for several hundred hours before the breakdown potential on the adjacent cathodes will have risen sufficiently to prevent the tube stepping. Nickel is therefore the commonest material used for the rod cathodes. The contamination is easily removed by passing a discharge to the affected cathode for a few seconds. Because the rate of restoration is so much more rapid than the contamination, tubes that are regularly cycled, even at only one pulse per day, will show considerably longer lives than those in which the discharge is stationary for long periods; this is discussed in the next section.

To minimise the spread in breakdown and maintaining potentials it is essential to clean the cathodes thoroughly during manufacture, and to prevent contamination as far as possible from gases evolved from the tube parts during operation. Clean cathodes are achieved by a careful chemical etching before assembly, and by sputtering after the tube has been sealed from the pump. Although the sputtering process is limited, it is possible, because of the geometry, to clean the cathodes and deposit an opaque layer on the glass envelope close to the cathodes, without seriously reducing the light transmission of the viewing window. During this process the glow must be invested on all the cathodes at the same time. In the large envelope tubes there is a considerable area of glass apart from the window which cannot be covered by the sputtered material, and it is common to insert a barium getter in such tubes.

There is normally a final ageing process in the manufacture of the tube in which the glow is cycled round the cathodes, spending an equal length of time on each.

The transfer characteristics depend critically on geometry, and essential to the stepping tube is the accurate symmetrical positioning of the cathodes around the anode disk. The method of achieving this is to use ceramic spacers with fairly close tolerances, a typical construction being illustrated in Plate 8.1. The anode disk, 15 mm diameter, is supported on a central wire about 7 mm above the main ceramic spacing disk. This ceramic disk has the 30 holes for positioning the cathodes in a circle of 20 mm diameter, chosen to allow the glow position to be readily distinguished. The holes are a

close fit on the 0·4 mm diameter cathode wires, but are counterbored to prevent electrical leakage via any sputtered film which may be formed on the ceramic. Beneath the main ceramic, the cathode wires are welded on to spiders which interconnect the appropriate guide electrodes. The spiders are also positioned by ceramic disks. The whole assembly holds the cathodes rigidly in position for mounting on the foot. In this operation, the position of the electrodes must be lined up with the foot or base pins to ensure the digit positions are in the right place when plugged into the valve socket.

8.3. Circuit Design and Performance for the Double Pulse Tube

8.3.1. THE INPUT DRIVE

As Kandiah has shown[11] there are several combinations of pulses and bias potentials which can be employed to drive a stepping tube; for example a positive pulse can be applied to the cathode carrying the discharge to effect the transfer. However, the common method is to apply a pair of negative pulses to the guide cathodes which are normally held at a positive bias, as already described in Section 8.1.1. The problem is then one of interpreting the published data of the guide drive requirements for practical circuit waveforms.

Published data are usually based on the use of rectangular drive pulses where the voltage and the period for which the discharge rests on each of the thirty cathodes are precisely defined.* It is rather expensive to produce pulses of this shape, and unless the tube is operated near its limits less exacting waveforms derived from low cost circuits can be used.

The closest approach to negative rectangular pulses can be obtained from the output of a blocking oscillator; typical output waveforms are shown in Fig. 8.6. The pulses depart from the 'ideal' negative rectangular waveform in two respects which must be taken into account in the circuit design. Firstly the trailing edge of the guide *B* pulse is relatively long owing to the decay time of residual energy in the transformer. Since the discharge remains on the guide *B* cathode until its potential has risen above that of the main cathode, it is important that the total period, rather than the time for which the pulse amplitude is maximum, be considered as the pulse width. Secondly, when the blocking oscillator switches from guide *A* to the guide *B* period, a gap equal to the rise time of

*Normally the maximum counting speed of the tube, i.e. 5 kc/s for the Z504S, can only be achieved with square-wave pulses.

the guide A pulse occurs between the end of the guide A pulse and the arrival of the guide B pulse. If this gap is greater than 5μ sec it could allow the anode potential to rise sufficiently to restrike the previously primed main cathode. A practical transistor circuit designed to drive a Z504S at speeds up to 4 kc/s has been described

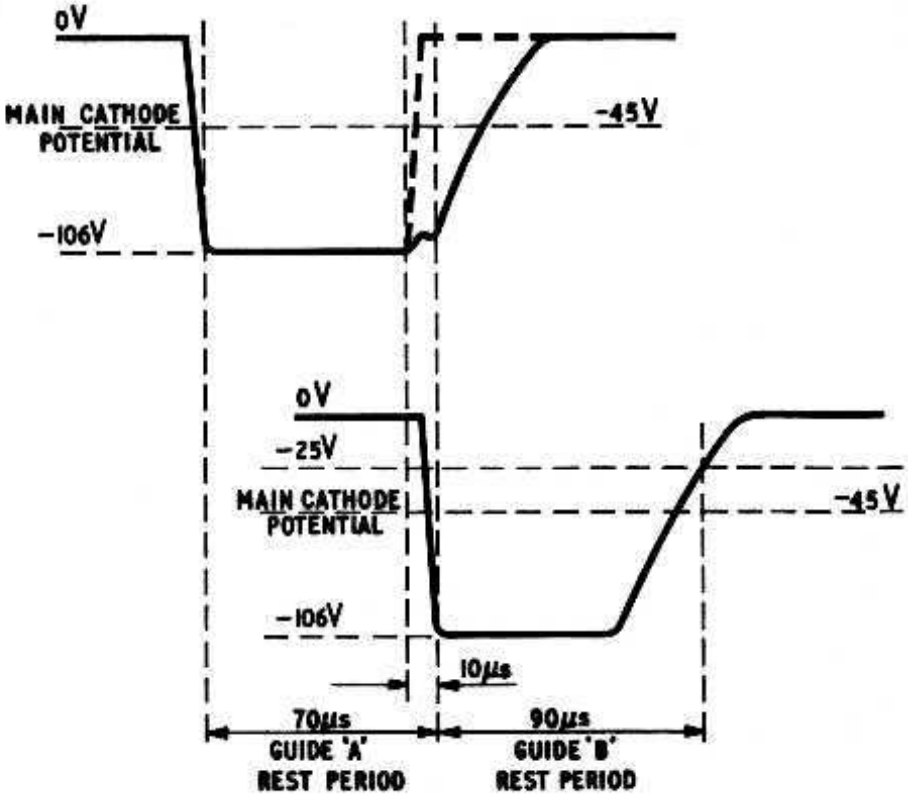
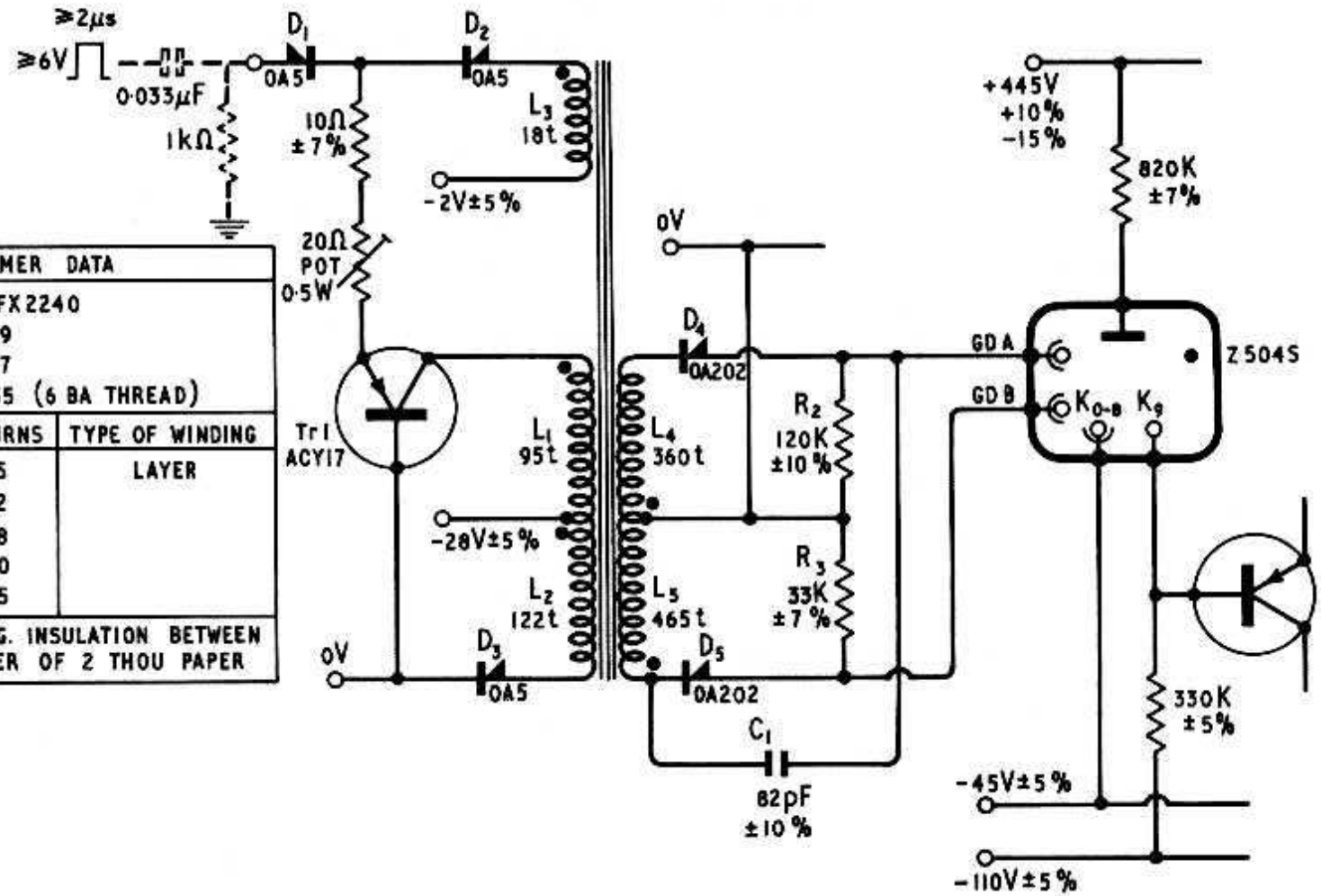


Fig. 8.6. Output waveforms from a typical blocking oscillator drive circuit, measured at the guide cathodes of the stepping tube

by Chappell and Jeynes[12] and is shown in Fig. 8.7. In this circuit a gap of approximately 10μ sec occurs between the guide pulses at the transformer output terminals. The gap is avoided at the guide cathodes, however, by inserting capacitor C_1 between the guide A cathodes and the guide B transformer winding so that the guide A potential is held negative for the duration of the 10μ sec switch-over period. The dotted line of Fig. 8.6 indicates the end of the guide pulse at the transformer, and the solid line of the pulse at the guide A cathodes. Alternative circuits which differ slightly from the above have also been described in the literature.[13]

In a number of applications, particularly in stepping tube coupling circuits, only one rectangular pulse is available. In this case the input pulse can be applied direct to the guide A cathodes and also to an integrating circuit, which forms a delayed pulse for the guide B cathodes. A typical drive circuit using a double triode, and the associated guide waveforms, are shown in Fig. 8.8. In operation the discharge steps to the guide A cathodes when the rectangular pulse is applied. At the same time the guide B potential



TRANSFORMER DATA		
TRANSFORMER CORE FX2240		
BOBBIN TYPE DT 2179		
TAG PLATE DT 2227		
CLAMPING NUT DT 2155 (6 BA THREAD)		
WINDING ORDER	No TURNS	TYPE OF WINDING
L_1	95	LAYER
L_2	122	
L_3	18	
L_4	360	
L_5	465	
ALL WINDINGS 40 S.W.G. INSULATION BETWEEN WINDINGS 1 LAYER OF 2 THOU PAPER		

Fig. 8.7. 0–4 kc/s blocking oscillator drive circuit

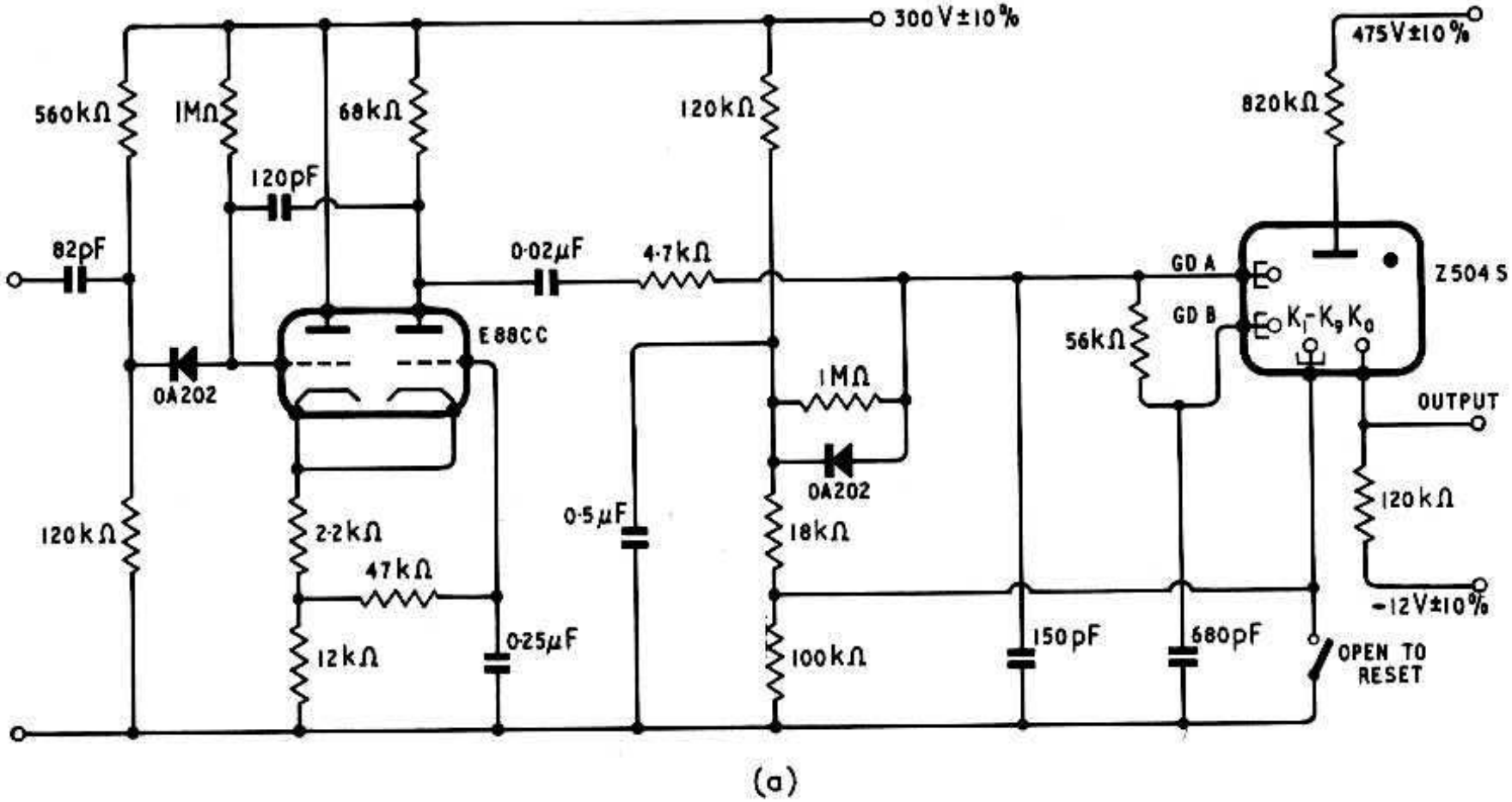


Fig. 8.8. Rectangular and integrated pulse drive; (a) drive circuit using a double triode to produce the rectangular pulse (b) Voltage waveforms at the guide cathodes

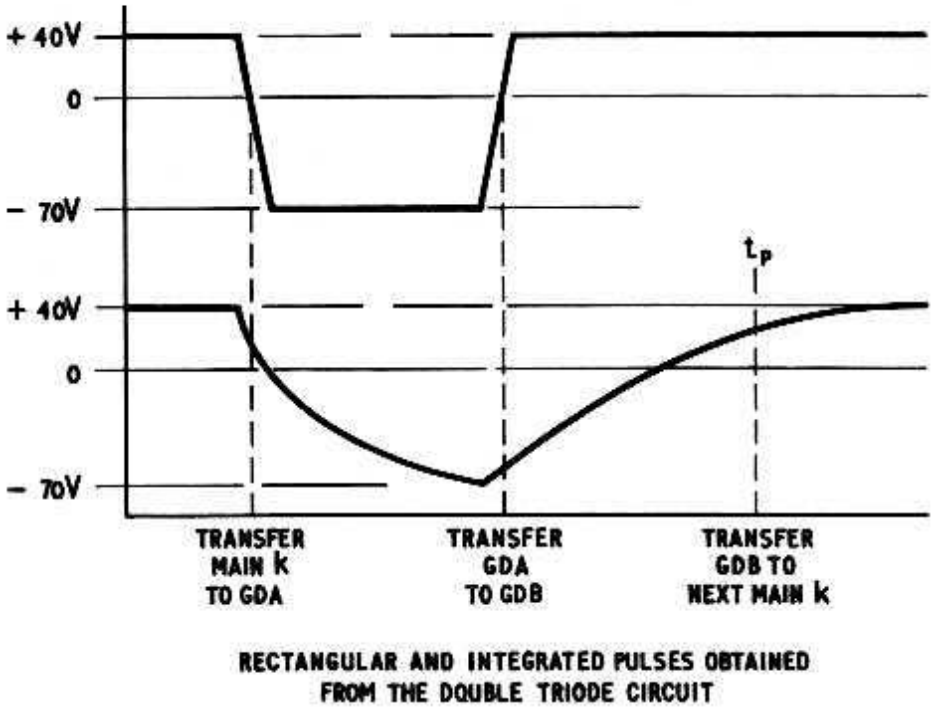


Fig. 8.8. (b)

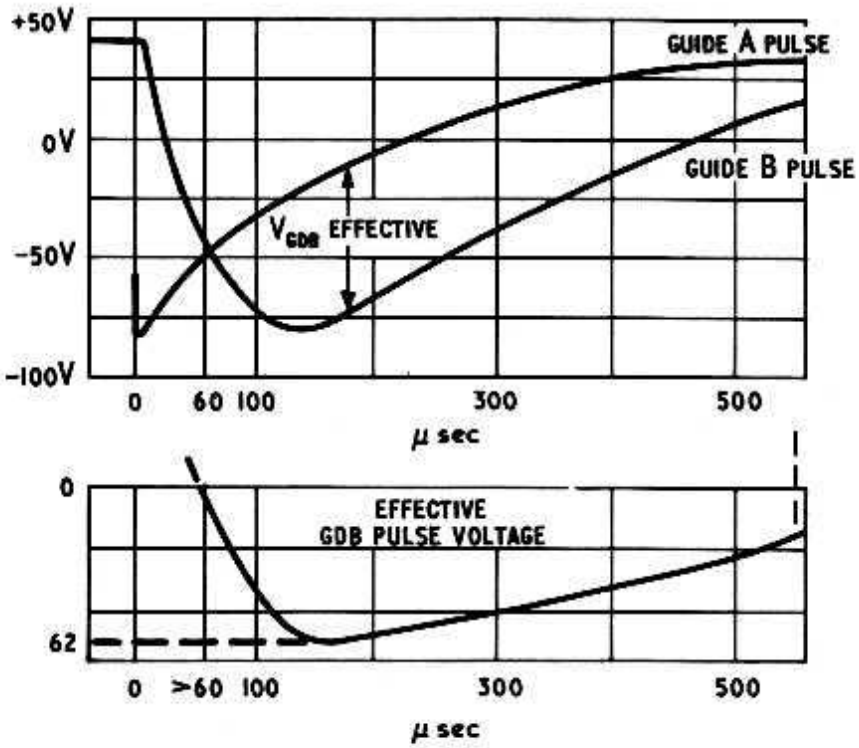


Fig. 8.9. Voltage waveform at the guide cathodes for differentiated and integrated pulse drive

starts to fall exponentially towards the potential of the guide *A* cathodes. At the end of the rectangular pulse the discharge transfers to the guide *B* cathode which is already at the required negative potential. The potential of the guide *B* cathodes then rises as the integrating network (56 kΩ and 680 pF) charges up again until,

at time t_p in Fig. 8.8b when the guide B potential is positive, the discharge can step on to the next main cathode. It is important to note that the effective width of the guide B pulse is from the end of the guide A pulse to t_p. This method of operation can be used in circuits designed for speeds up to 3 kc/s (for Z504S) where exact control of the output pulse shape is not required. Also, as there is

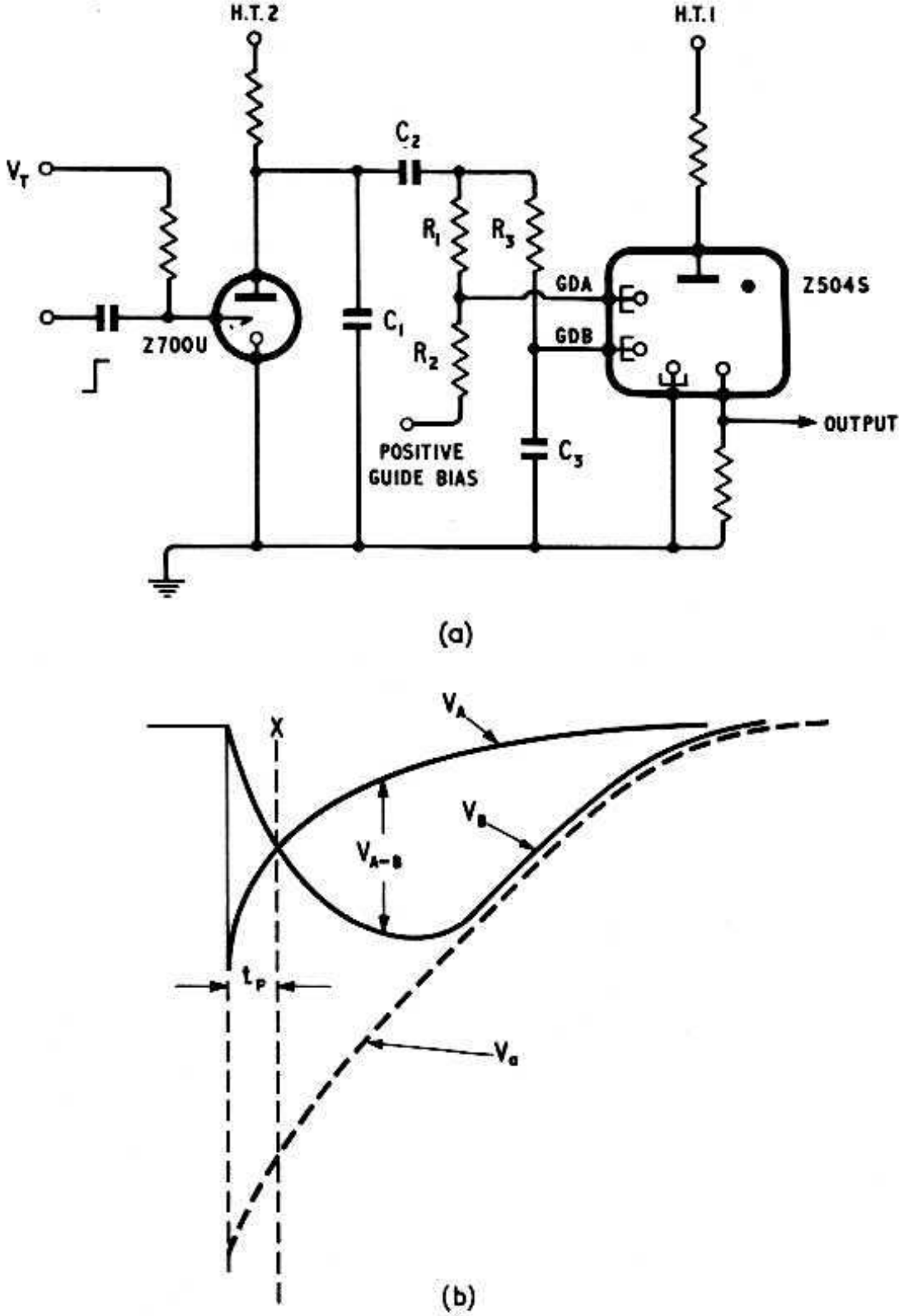


Fig. 8.10. Differentiated and integrated pulse drive: (a) drive circuit using a trigger tube suitable for speeds up to 40 c/s; (b) Voltage waveforms, illustrating the derivation of the guide B pulse

no possibility of a gap between the guide pulses, close control of the input pulse shape is not usually necessary.

The final and most common pulse drive is illustrated in Fig. 8.9. Basically the input pulse to the guide *A* cathodes is a differentiated negative pulse and the input to guide *B* cathodes is an integrated negative pulse. These pulses can be obtained by suitably amplifying a single voltage or current step function. To ensure that the discharge remains on the guide *A* cathodes for the required time during transfer, it is essential that the guide *A* potential remain negative with respect to the guide *B* potential for this period; i.e. the crossover point, Fig. 8.9, should occur at a time t_p after the start of the guide *A* pulse, where $t_p \geqslant$ the recommended pulse width. The effective guide *B* pulse amplitude is the potential difference between the guides after the crossover (lower curve on Fig. 8.9), and this must be taken into account when designing the circuit. Transfer to the main cathode is accomplished, as before, when the guide *B* potential becomes more positive than the main cathode potential. A simplified circuit for slow speeds up to 40 c/s using differentiated and integrated guide pulses is illustrated in Fig. 8.10(a). The

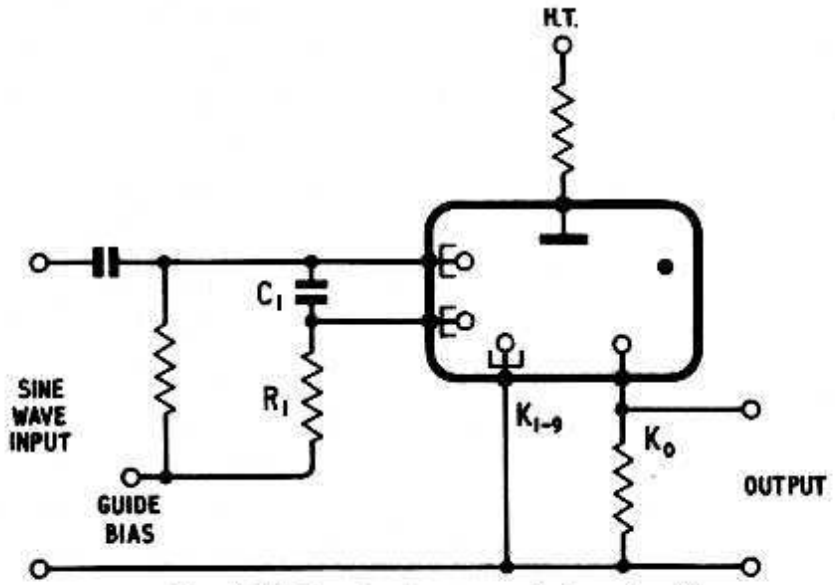


Fig. 8.11. Basic sine-wave drive circuit

amplifier is a Z700U cold-cathode trigger tube operating in a self-extinguishing circuit. The output pulse of the Z700U is fed through the coupling capacitor C_2 to the pulse shaping network. The potential divider R_1 and R_2 applies the required pulse to the guide *A* cathodes whilst the integrating network R_3 and C_3 applies a delayed pulse to the guide *B* cathodes. The relevant waveforms for this circuit are shown in Fig. 8.10(b). It will be seen that at no time can the guide *B* voltage become more negative than the potential at the trigger tube anode V_a, so that if V_a were applied directly to guide *A* the discharge would not transfer to the guide *B* cathode, hence the

need for the attenuator R_1R_2. A modification of this circuit can be used to increase the speed, but the long exponential recovery of the guide potential limits the counting rate with this type of input to about 500 c/s. However, this speed is adequate for coupling stepping tubes in decade scalers, where differentiated/integrated guide-pulse drive is most frequently used.

There are certain applications where a sine-wave drive for the first tube in a chain is useful. Providing the input frequency does not vary appreciably and the signal is applied continuously, the simple circuit illustrated in Fig. 8.11 can be employed. For reliable counting the capacitor-resistor network C_1R_1 should be chosen to give a phase advance of approximately 45° at the operating frequency.

8.3.2. POSITIVE GUIDE BIAS RANGE AND PULSE AMPLITUDE

The positive d.c. potential on the guides ensures that the discharge is stable on the rest or main cathodes, and also that at the end of the guide *B* pulse, the discharge transfers reliably to the next main cathode i.e. does not remain on guide *B* or transfer back to guide *A*. For a perfectly symmetrical tube this required bias can be relatively low. Indeed initially, after processing, a minimum bias of 10–15 V is sufficient to operate the Z504S, although this increases if the rest period on the main cathodes is less than 50μ sec or if the tube current is increased.[14] However, the tube does not remain so symmetrical, in terms of the breakdown and maintaining potential of each cathode-anode gap, on life, particularly if the discharge is resting on one cathode for a considerable length of time. To take care of the variations around the tube during life a higher bias potential is required, and most manufacturers quote a minimum bias potential of around 30 V for the inert-gas stepping tubes; for the Z504S, Barker [14] suggests a minimum guide bias of 25 V— with main-cathode rest periods greater than 1 msec—and a minimum of 35 V for rest periods of less than 1 msec. These values were arrived at after life tests of 25,000 hr with the tubes cycling and 9,000 hr standby with the glow at rest on one cathode.

The ultimate maximum guide bias potential will be reached when the guides act as anodes, causing spurious breakdown to an adjacent main cathode. However, in practice, an upper limit occurs at a much lower value, which is associated with adjacent cathode contamination, discussed in Section 8.2.2. It was found that the rate of contamination, measured as the change in transfer voltage, of the guides adjacent to the current-carrying cathode is dependent on the guide potential; the more positive the guide potential the

greater the contamination effect. This is probably due to the fact that as the voltage is increased the positive-ion current to the guides, which tends to counteract the contamination by re-sputtering, is reduced. It has also been found that high guide bias potentials tend to inhibit deionisation of the extinguished main gap, possibly by focusing the positive ions around the main cathode. As a result of a combination of these two effects the discharge may fail to transfer with short duration pulses after long standby periods (the discharge re-ignites on the previous cathode at the end of the guide *B* pulse). It is therefore recommended that the guide bias potential is kept as low as possible; for the Z504S Barker suggests a maximum value of 45 V.

In a similar way to the guide bias potential, the pulse amplitude is limited in its range. At the start of life the minimum pulse amplitude required to transfer the discharge is about 20 V negative with respect to the main-cathode potential. On life, when a wider spread in the breakdown and maintaining potential occurs, larger input pulses are necessary for reliable stepping. The value depends on the voltage spread which in turn is a function of the operating life conditions. As already stated, standby with the glow on one cathode for long periods is the most detrimental. From life experiments on the Z504S it was concluded that, with the absolute minimum recommended pulse duration of 60μ sec, the region of reliable operation commenced with negative guide-voltages of $\geqslant$ 45 V (i.e. total pulse is 45 V plus the guide bias potential). The reliability, however, increased as the negative guide-voltage increased. The maximum pulse amplitude is limited only by the necessity to restrict the voltage difference between any pair of adjacent cathodes to less than the primed breakdown potential, which for inert gas tubes is about 150 V.

8.3.3. THE OUTPUT SIGNAL

The output signal can be obtained by connecting a resistor in the required rest cathode lead. When the discharge transfers to the output cathode, the potential rises giving a positive potential step. When the discharge transfers from the output cathode, a negative step is obtained. There will obviously be a limit to which the output cathode potential can be made positive without causing spurious transfer. This maximum permitted potential is related to the guide bias potential as shown by the typical curves in Fig. 8.12. In the region *AB* the value of the output resistor may be increased until the cathode voltage becomes almost equal to the guide bias potential, and current sharing occurs between the output cathode

and adjacent guide cathodes. If, however, the current to the previous guide *B* cathode is too high it will prime the preceding guide *A* cathode, and on receipt of the next pulse it may ignite instead of the following guide *A* cathode, so that the discharge cannot be stepped from the output cathode. In region *BC* where the maximum positive output potential decreases as the positive

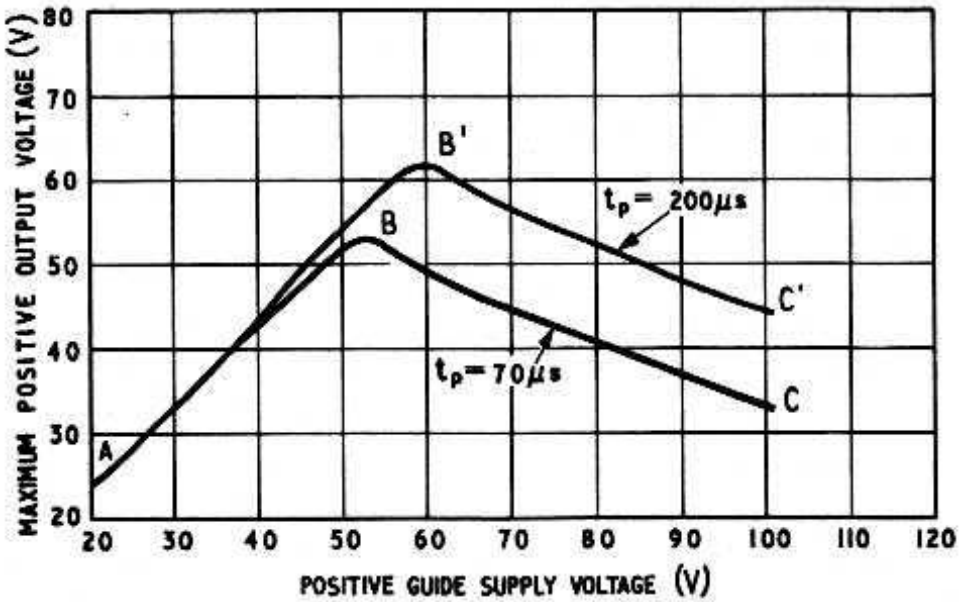


Fig. 8.12. Maximum positive output voltage as a function of the guide bias potential. Barker[14]

guide bias is increased, failure occurs due to the discharge transferring back to the previous main cathode at the end of the driving pulses, i.e. it will not step on to the output cathode. This is probably due to the high bias potential inhibiting deionisation as suggested in the last section. As a result the breakdown potential of the previous digit gap is less than the sum of the maintaining potential of the anode–output cathode gap plus the voltage developed across the output resistor. This premise is supported by the fact that increasing the pulse duration increases the maximum positive excursion of the output voltage (see Fig. 8.12). Normally the tube will operate in the region *AB*, and it is advisable for reliable stepping to limit the output cathode potential to at least 10 V below the guide bias potential, where current sharing does not occur.

When a larger output signal is required at the cathode, the cathode resistor can be returned to a source of negative bias, providing of course the negative potential does not cause spurious transfer. As with other potentials the maximum value of negative bias which can be tolerated depends on the variation of breakdown potential around the tube. It also depends on the duration of the guide *B* pulse, since it is whilst the discharge is on the following guide *B* cathode that the output cathode–anode gap deionises and the following main cathode–anode gap is primed. With minimum

pulse duration only a few negative volts can be tolerated on the output cathode, but with longer pulses -12 to -15 V can be employed. Use of a negative bias on the output cathode requires a corresponding increase in the guide pulse amplitude and/or reduction in the guide bias.

8.3.4. LIFE PERFORMANCE

The main problems in the life of stepping tubes are the variations in maintaining and breakdown potentials of individual cathodes which may occur during operation. These changes tend to reduce the characteristic operating ranges and may result in miscounts. However, providing these changes are allowed for in the circuit design, as described in the previous sections, the chances of a miscount can be reduced to a minimum. In general the changes are reversible, and when the discharge moves on to a 'contaminated' cathode rapid cleaning of the surface takes place restoring V_s and V_m to their initial values. Thus when the tube is operating in circuits where the discharge is stepping from one rest cathode to the next at intervals of, say, less than 3 hr, very little deterioration in the characteristics is observed. Under these conditions no miscounts occur and life is limited only by obscuration of the viewing window at some 15,000 hr. Where a visual indication is not required, a life in excess of 25,000 hr may be expected.

In applications which require long periods of standby life on one cathode, the operating ranges must be restricted to obtain maximum reliability. The optimum operating range depends somewhat on the circuit conditions and is usually given by the manufacturer. Table 8.1 gives the recommended operating ranges for the Z504S as detailed by Barker.[14] It is worth noting that a nominal anode current of 330 μA is given. At the start of life the tube operates over a wide current range of 100 to 1,300 μA, but if the current is too high ($> 500\,\mu$A) deterioration in the stability of the tube characteristics is more rapid. It is therefore advisable to keep the anode current as low as possible, and in this respect 250 μA is the minimum for reliable operation at all speeds. The nominal value of 330 μA, therefore, allows for up to 20% variation in the supply.

Deterioration in the operating range also occurs if the tubes are stored at high temperature, due to gas evolved from the electrodes and glass envelope. It is therefore recommended that tubes are not stored above an ambient temperature of 60°C if they are to maintain their wide operational range.

Table 8.1

Recommended Characteristic Operating Ranges of the Z504S

(a) All operating speeds where a standby period < 1 hour is expected
(b) For maximum counting speeds and where a standby period > 1 hour is expected
(c) For counting speeds ⩽ 3kc/s and where a standby period > 1 hour is expected

		(a)	*(b)*	*(c)*
Nominal anode current	(μA)	330	330	330
Minimum dwell time:				
guide A cathode	(μsec)	60	60	60
guide B cathode	(μsec)	60	80	120
main cathode	(μsec)	75	75	75
Negative guide voltage for transfer:				
minimum	(V)	−45	−45	−45
maximum	(V)	−140 minus the positive guide supply voltage	−70	−140 minus the positive guide supply voltage
Positive guide supply voltage:				
minimum				
main cathode dwell time > 1ms	(V)	25	—	25
main cathode dwell time ⩽ 1ms	(V)	35	35	35
maximum	(V)	50	45	50
Maximum negative bias voltage on any main cathode	(V)	−14*	−14*	−14
Maximum positive voltage excursion on any main cathode	(V)	Positive guide supply voltage −10V		

*where a 14V negative bias is used with guide B rest periods <120μsec the minimum recommended negative guide voltage for stepping should be increased to −60V.

N.B. The characteristics rapidly return to their original values if the discharge is rotated around the tube for a few seconds. It is recommended that this process should be carried out whenever possible after standby periods in excess of one hour.

8.4. Stepping Tube Applications

8.4.1. SCALER CIRCUITS

The majority of stepping tube applications where several tubes are coupled in cascade can be grouped under this heading. Included in this group are random pulse counting, frequency division, timing, batching and measurement.

Important in such applications is the coupling circuit. If the input to the first tube is from random pulses then the rest period on the output cathode will not be a constant, and consequently, although a rectangular output pulse is obtained, it will be of irregular duration. Thus it is not possible merely to invert and amplify the pulse to drive the next tube, but a coupling circuit must be used which will give a differentiated and integrated pulse from a step function. The trigger tube circuit of Fig. 8.10a and the thermionic valve circuit

of Fig. 8.8a can both be used for this purpose, by directly applying the output of one stage to the input terminal of the coupling circuit driving the next stage. For higher speeds more sophisticated circuits may be required. Suitable coupling can also be achieved by transistors; several circuits have been proposed in the literature.[12, 13, 15] A three transistor coupling circuit is illustrated in Fig. 8.14 which is explained later. Even when the input is at a constant frequency, pulse shaping is often desirable to obtain the optimum drive pulse for the succeeding tubes. Suitable coupling

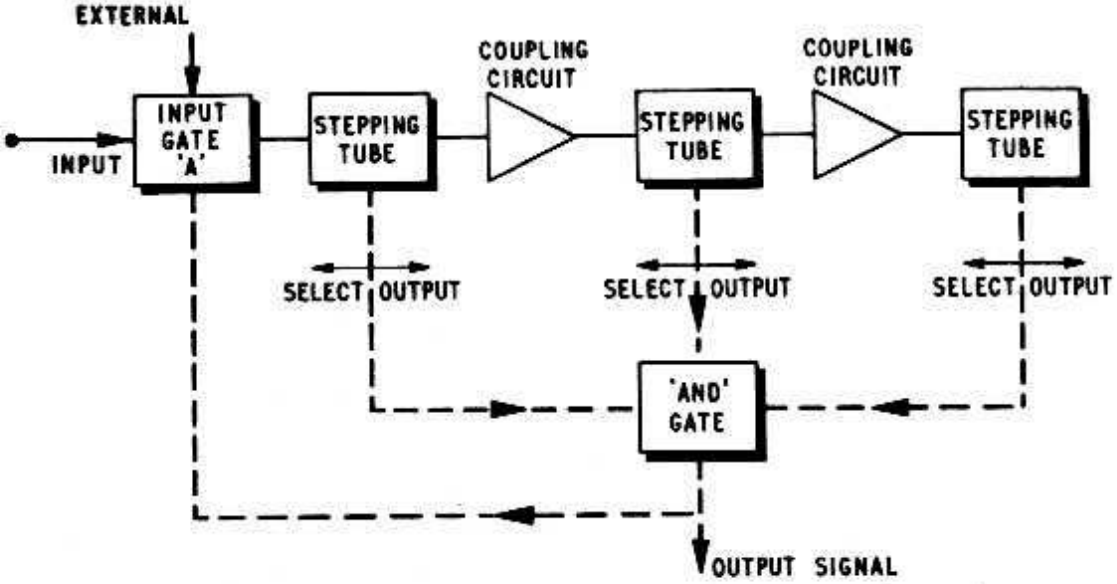


Fig. 8.13. Block schematic diagram of the basic stepping tube scaler with connections for batching and timing applications

circuits are usually recommended by the manufacturers, and can be found in their handbooks.

To illustrate the various scaler applications, consider the block schematic diagram of Fig. 8.13 which shows a chain of counting tubes with associated circuits (full lines). When the tenth or zero output cathode K_o of each stage is connected to drive the succeeding coupling circuit, the arrangement becomes a simple random pulse decade-counter. The number of input pulses are counted, the state of the count observed visually or electrically, and the counter reset to zero after the state of the count has been read.

A practical scaler for up to 4,000 pulse per second input using Z504S tubes with transistor coupling [12] is shown in Fig. 8.14. In the coupling circuit a pulse is applied to the following tube when the discharge transfers from the output cathode in the driving tube, i.e. the coupling circuit functions from a negative pulse. Thus K_9 and not K_o must be connected to the coupling circuit. Assuming that the discharge is stepping off K_9 in the first tube $V1$, then $Tr2$, which was off whilst the discharge rested on K_9, is switched on and a positive step is applied via C_4 to the base of $Tr3$, switching off

the cascade circuit: the two ACY17 transistors could be replaced by a single high voltage transistor if an inexpensive one were available. Whilst *Tr*3 and *Tr*4 are conducting, a positive bias is applied to the guide cathodes since the main rest cathodes are at −45 V. When they are switched off, the potential on guide *A* drops to the negative potential of 110 V. However, *Tr*3 is switched on again as C_4 recharges via R_{14}, and the potential at guide *A* recovers. Thus a negative pulse is applied to the guide *A* cathodes which is roughly rectangular with a duration determined by C_4R_{14}. The pulse is fed via an integrating circuit to the guide *B* cathode. The

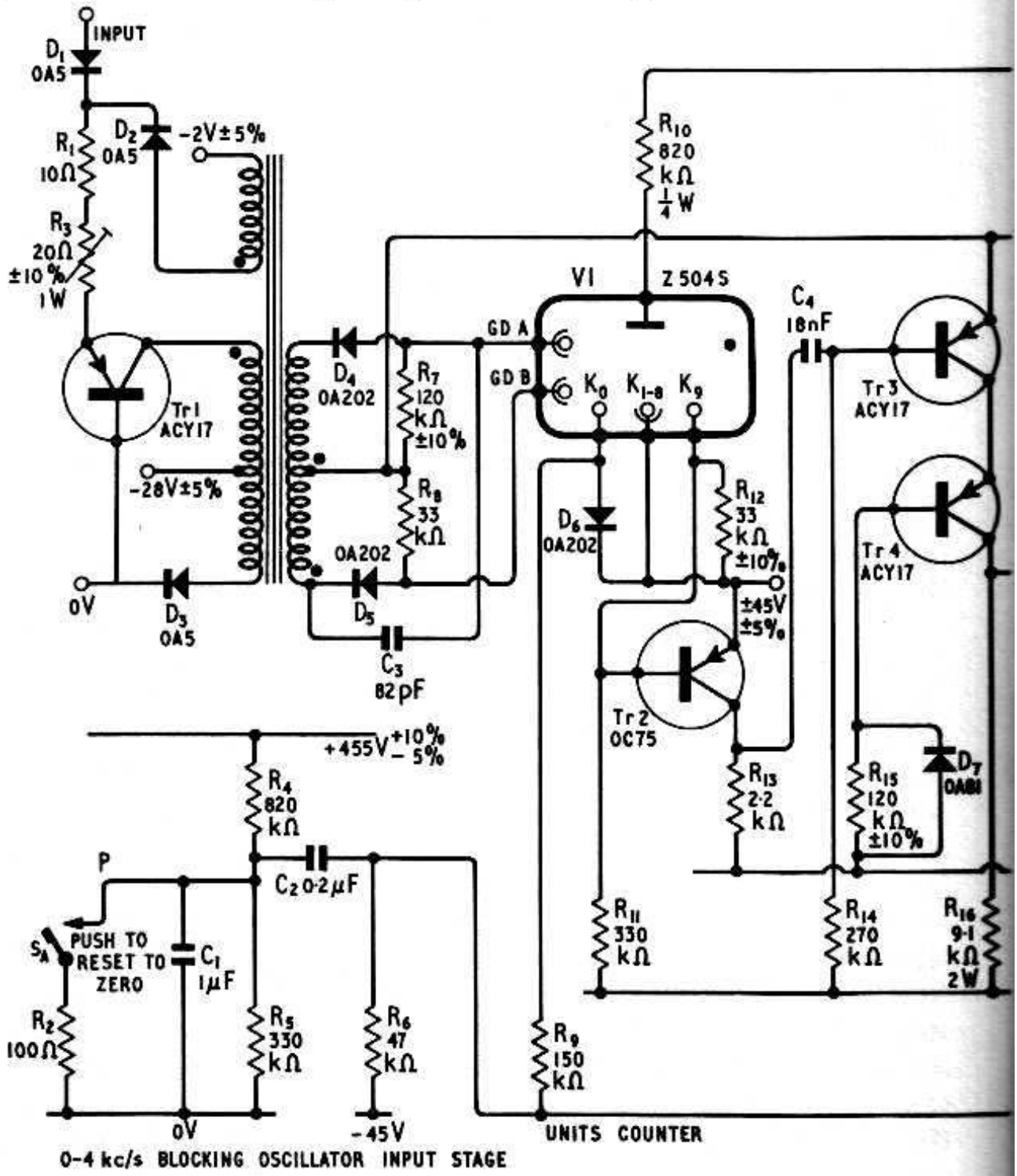
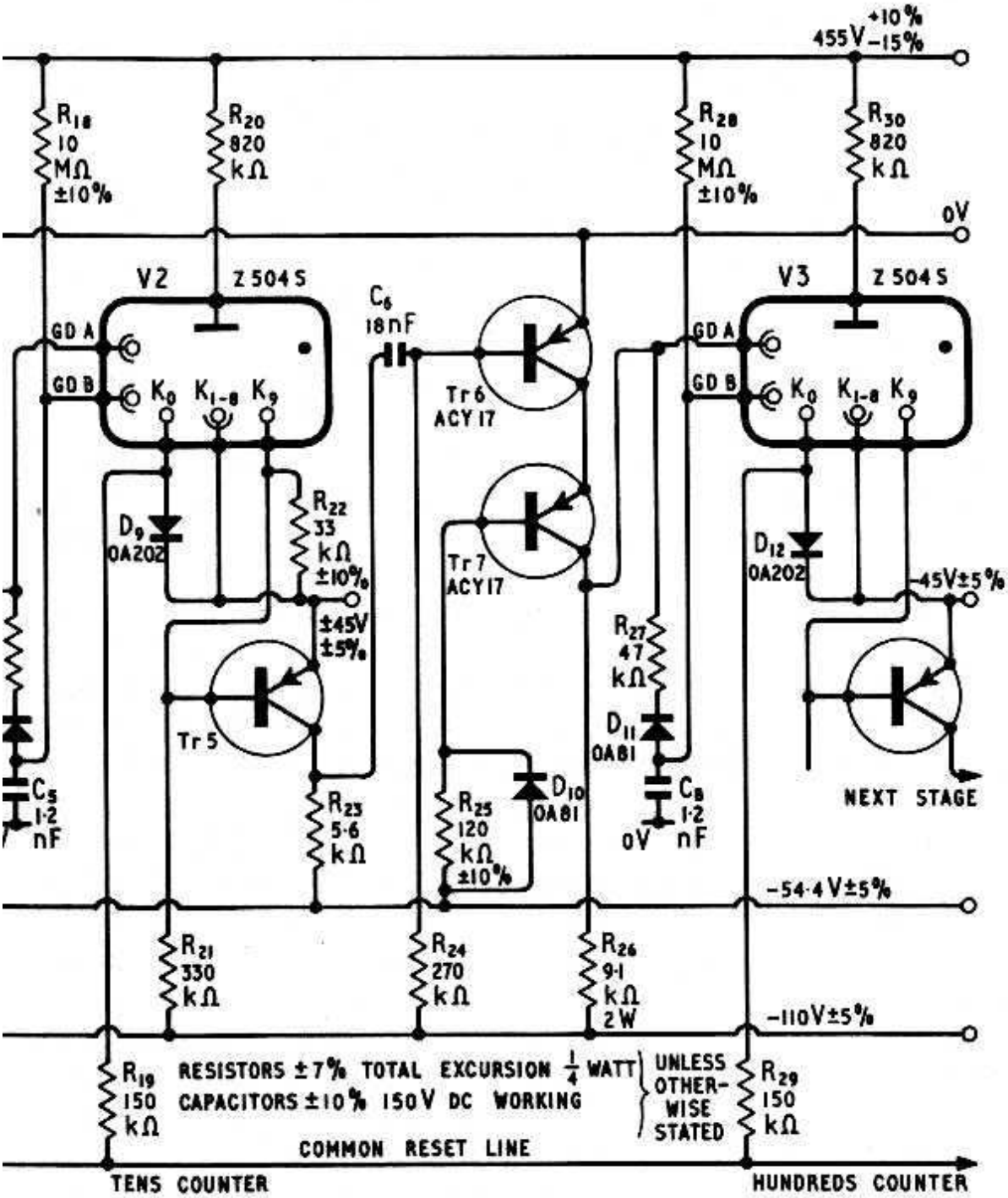


Fig. 8.14. Z504S scaler with 0–4,000 pulse per second input circuit and reset facilities

scaler is reset by applying a large negative pulse to the K_o cathodes.

Frequency division can also be performed with the basic scaler of Fig. 8.13. With a standard frequency source applied to the input terminal, and output obtained from appropriate cathode combinations, the input frequency can be divided by factors of two, five or ten. For example, by connecting every other cathode as illustrated in Fig. 8.15, half the input pulse frequency is obtained at the output. If tubes with twelve digit or main cathodes are employed, division by two, three, four, six or twelve is possible.

In a further application the circuit can be connected to act as



an accurate timer. This arrangement is indicated by the dotted lines of Fig. 8.13, where an output is obtained from a preselected cathode in each tube and applied to an 'AND' gate. Assume, for example, that the circuit operates as a decade scaler and that cathodes K_5, K_6, and K_7 of the units, tens, and hundreds stages respectively are brought out to the 'AND' gate. If now an external signal opens the

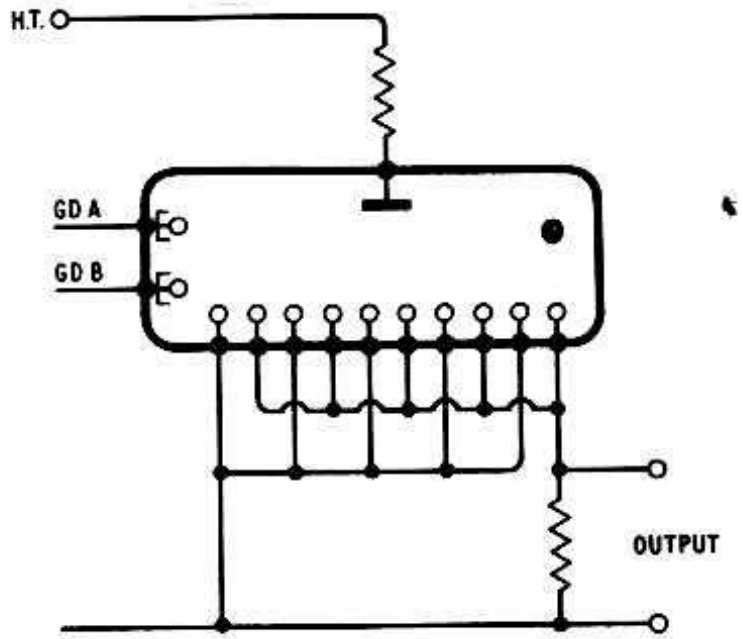


Fig. 8.15. Basic frequency division circuit

input gate 'A', and the circuit counts from zero at the rate of 1 kc/s, an output will be obtained from the 'AND' gate 0·765 sec after the input gate is opened. This pulse can be used to operate a further circuit.

The same scaler can also be employed to measure time intervals in an external circuit. In this case a signal corresponding to the start of the period to be measured is used to open the input gate, and a further pulse at the end of the period closes this gate again. During the period the gate is open, the scaler counts pulses from a standard frequency source. An application of this principle to an inexpensive digital voltmeter has been described by Cleghorn.[16] An amplified value of the analogue voltage is compared with a voltage 'staircase' having accurately defined steps. The number of steps from zero voltage to the point where the two voltages are equal is counted on a Dekatron scaler and fed to a digital output.

For batching, the process is similar to timing, the number of articles in the batch is preselected by connecting the desired output cathodes in the scaler to the 'AND' gate. The input signal, however, now consists of a train of random pulses, each pulse corresponding to one article in the batch. Thus when the preselected number of articles has been counted the 'AND' gate output signal can be used to terminate the process. Basically this technique can be used to

control any quantity which is expressible in pulse form. For example, the input could be obtained from a photo-conductor, giving a pulse as an article passes between it and a lamp. Batching counters based on this arrangement have been described in the literature.[17, 18]

8.4.2. LOGIC CIRCUITS

The stepping tube can also be used in a potentially wide field of application in industrial control equipment and small calculating machines, where pulse-train generation and pulse programming at relatively low speeds are required. For example, it has been employed as a pulse-train generator in a resistance welder controller.[19, 20] In this application the output cathodes of the tube are connected either direct to a common negative line or through an output resistor to this line. This arrangement is shown in Fig. 8.16. When the operator initiates a weld, the discharge makes one complete rotation around the tube and a train of output pulses (1 to 7) corresponding in number to the cathodes connected to the output resistor is generated across R_K. These pulses can be amplified and used to switch a pair of thyratrons connected 'back to back' in series with the welder transformer. Thus the number of welding cycles can be preselected by connecting the required number of stepping-tube cathodes to the output resistor.

A more complex pulse-train generator combining more than one

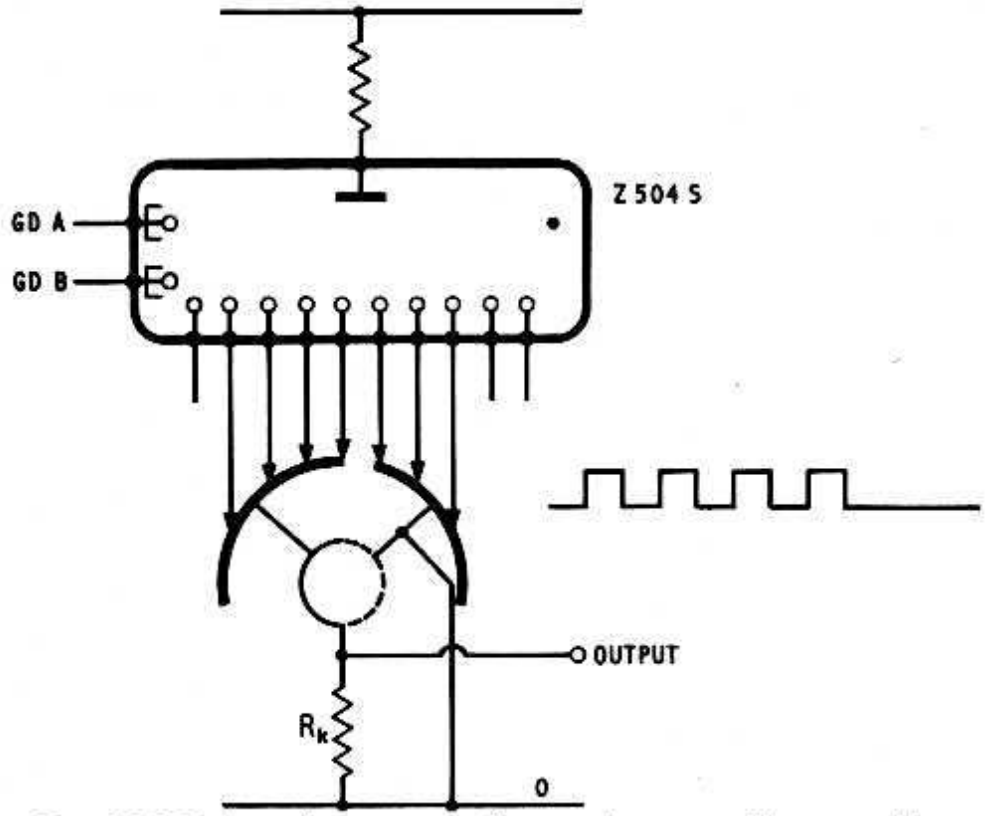


Fig. 8.16. Pulse train generator for a resistance welder controller

stepping tube has been described by Flood and Warman,[21] capable of supplying pulses at frequencies from 5–50 c/s either continuously or in trains of from 1 to 11 pulses, with pauses between trains adjustable between 66 msec and 4 sec.

Alternatively, the stepping tube can be used as a pulse programming device by connecting suitable combinations of cathodes to separate output terminals in the manner shown in Fig. 8.17. The

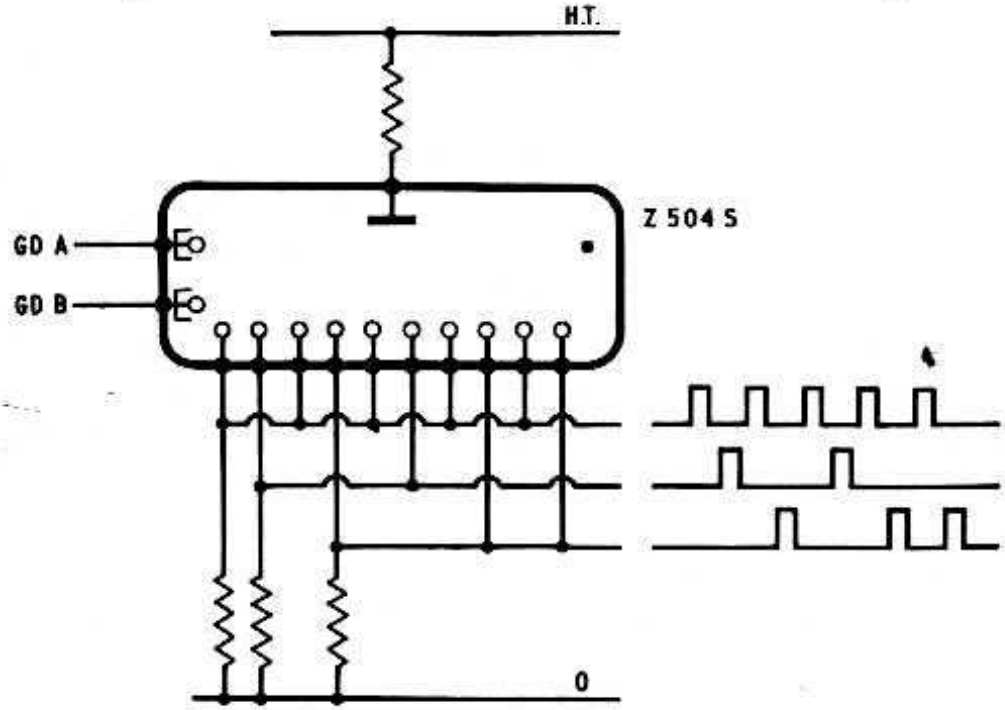


Fig. 8.17. Basic pulse programme circuit

stepping tube can be used in this way as a 'master clock' in a logical system. In general this system is limited to a single tube giving a maximum of ten or twelve output channels. However this can be extended by using two stepping tubes suitably inter-connected with diode gates, or by connecting two or more stepping tubes in parallel and arranging that they count the pulses in succession. Two methods of achieving the latter process have been described by Stearman.[22]

In the small business machine the stepping tube can be used both in the accumulator or register circuit, carrying out addition and subtraction, and also for information storage. For the accumulator it is convenient to use a reversible counter rather than subtract by adding the 10's complement. Although a single stepping tube can be readily reversed, when additional stages are added to extend the reversible counting to tens, hundreds, etc., difficulties arise in deriving the necessary carry signals. When any stage changes from 9 to 0 it must be arranged to add or 'carry' a digit to the next higher stage, whilst a 'borrow' must be subtracted from the next stage when the discharge moves from 0 to 9 (i.e. the following tube must decide on the counting direction in the previous stage). Branson[23] overcame this problem by using stepping tubes in which two of the

guide cathodes GDA_9 and GDB_9 were connected separately and used as 'routing' guides. However, it is relatively simple with selector tubes, where the digit cathodes can be separately connected, to arrange a gating system for determining the direction of the count and to carry the correct signal. Such a circuit has been described by Barber[24] and is illustrated in block diagram form in Fig. 8.18. A, B, and C are transistor gates which are normally open,

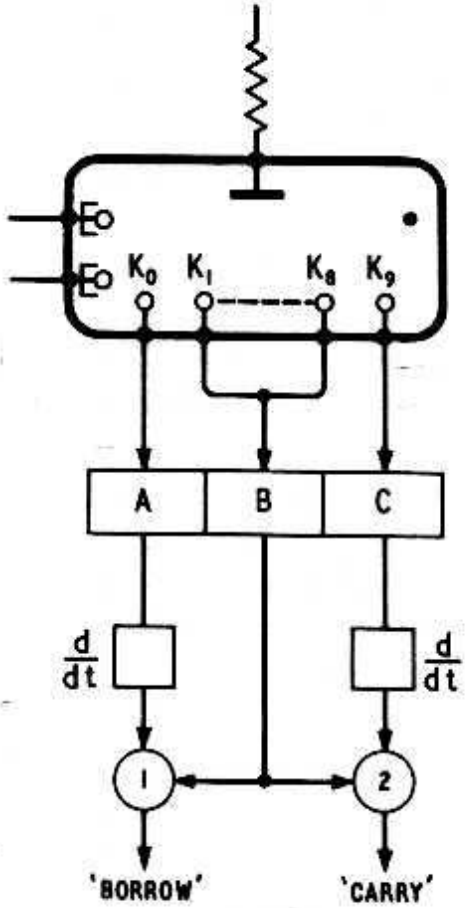


Fig. 8.18. Gating circuit for obtaining the 'borrow' or 'carry' pulse in a reversible scaler

but closed when the respective cathode of the stepping tube is carrying the discharge. Thus if the discharge is on K_9, A and B are open giving a positive d.c. output, whilst C is closed. When the discharge transfers from K_9 to K_0 a positive step pulse is derived at C, which is differentiated and fed to gate 2. The signal from B holds gate 2 open and thus the pulse is fed as a 'carry' on to the next stage. A similar pulse will be obtained at C when the discharge transfers from K_9 to K_8, but now the discharge on K_8 closes B, and therefore also gate 2, and the pulse is inhibited. In a similar way a 'borrow' pulse is fed to the next stage when the discharge transfers from K_0 to K_9. With a slightly more complicated circuit, Oxley[25] has shown that reversible counting can be obtained with stepping tubes having only one output cathode if the input signal to the scaler is used to gate the output of each stage. A practical accumulator for use with a stepping tube calculator has been described by Townsend and Camm.[26]

Recently Zilkha[27] has described an information store, using stepping tubes and in-line numerical readout. The novelty of the system is the 'dynamic' display which utilises a single set of cathode switches for operating any number of tubes (see Section 9.4.5.2). Only one output cathode is required from the stepping tube and the read out is non-destructive. Where it is necessary to store, process and display information in more than one store, a number of stepping tubes may be connected together, with this system, to form a matrix. Any tube in the matrix may be selected for operation on 'cross-point' principles by applying both guide *A* 'column' and guide *B* 'row' pulses to the selected tube.

Finally the stepping tube may be used to generate various waveforms. In particular a 'staircase' waveform can be generated by connecting successive cathodes of the stepping tube to progressively higher tappings on a common cathode load.[28] The voltage is of course limited to the maximum output which can be obtained from a single cathode, 30–40 V.

REFERENCES

1. The X-2076 developed by The Federal Telecommunications Lab. Inc., *U.S. Nat. Defence Contract report 17*, (1950)
2. HOUGH, G. H. and RIDLER, D. S., *Electron. Eng.*, **24**, 272, (1952)
3. KOEHLER, R. B. and RICHARDS, R. K., *Electronics*, **27**, 151, (Nov. 1954)
4. PECK, D. S., *Elec. Eng.*, **71**, 1136, (1952)
5. TOWNSEND, M. A., *Elec. Eng.*, **69**, 810, (1950)
6. BACON, R. C. and POLLARD, J. R., *Electron. Eng.*, **22**, 173, (1950)
7. ACTON, J. R., *Electron. Eng.*, **24**, 48, (1952)
8. APEL, K., *Elektron. Rundschau*, **14**, 405, (1960)
9. REANEY, D., *Electron. Eng.*, **34**, 412, (1962)
10. APEL, K., *Rad. Electron. Engr.*, **28**, 147, (1964)
11. KANDIAH, K., *Proc. Inst. Elec. Engrs.*, **101 (Part 2)**, 227, (1954)
12. CHAPPELL, G. C. and JEYNES, G. F., *Mullard, Tech. Commun.*, **7**, 310, (1964)
13. WARMAN, J. B. and BIBB, D. M., *Electron. Eng.*, **30**, 136, (1958)
14. BARKER, C. S., *Mullard Tech. Commun.*, **8**, 131, (1965)
15. CHAPLIN, G. B. B. and WILLIAMSON, R., *Proc. Inst. Elec. Engrs.*, **B105**, 231, (1958)
16. GLEGHORN, P., *Rad. Electron. Engr.*, **28**, 107, (1964)
17. CRAXTON, R. T., *Electron. Eng.*, **25**, 424, (1953)
18. TOOKE, P. E., *Electron. Eng.*, **26**, 160, (1954)
19. BRADY, T. W., *Electron. Eng.*, **28**, 70, (1956)
20. CROWTHER, G. O. and JEYNES, G. F., *Mullard Tech. Commun.*, **4**, 289, (1959)
21. FLOOD, J. E. and WARMAN, J. B., *Electron. Eng.*, **27**, 13, (1955)
22. STEARMAN, G. H., *Electron. Eng.*, **31**, 69, (1959)
23. BRANSON, L. C., *Electron. Eng.*, **27**, 266, (1955)
24. BARBER, D. L. A., *Electron. Eng.*, **31**, 42, (1959)
25. OXLEY, A. J., *Electron. Eng.*, **32**, 746, (1960)
26. TOWNSEND, R. and CAMM, K., *Electron. Eng.*, **29**, 58, (1957)
27. ZILKHA, S., *Mullard Tech. Commun.*, **8**, 162, (1965)
28. DUFFY, R. M. and GILBERT, C. P., *I.R.E. Trans. Electron. Computers*, **EC10**, 71, (1961)

CHAPTER 9

Indicators and Display Tubes

9.1. Introduction

Most of the functions of the glow discharge tubes which have been described in the previous chapters can be performed by modern solid state devices, albeit at higher cost and with greater complexity. However, the gas-discharge tube remains unchallenged as an efficient light source for illumination or display purposes.

The light radiation emanates from excitation of the gas molecules and radiative recombination of the ions and electrons. The positive column is the most efficient source of light in a discharge, and is that part of an arc discharge which is exploited for illumination in, for example, fluorescent lamps, neon signs, and flash tubes. The negative glow of a cold-cathode discharge, although less efficient, is also a fairly bright source of light, especially in neon where it is clearly visible in daylight even when the discharge is in the normal glow region. The negative glow has the added advantages of following the contours of the cathode so that the glow may be located precisely within the envelope, and also it can be produced in devices having voltage characteristics compatible with valve and transistor circuits. This makes it particularly suitable for register and display purposes, and a number of practical indicator tubes utilising the characteristics of the negative glow have been developed since the 1950's for use with computing and control circuits.

In general such tubes can be classified into three groups:

1. On-off indicators. Included in this group are mains indicator lamps, also diodes designed to operate from transistorised circuits.
2. Glow-position indicators or register tubes. These normally have ten glow positions, and are used in conjunction with decade counting circuits.

3. Numerical or character display tubes. In these tubes the glow is in the shape of a number or character and each character within the tube appears substantially at the same position to give 'in-line' display.

Tubes in all three of the above groups are commercially available, and examples are discussed in some detail in this chapter. However, other types of glow discharge devices have been described in the literature for display applications, and some of these are also mentioned to illustrate the versatility and potentialities of the glow discharge in this field.

When considering the visual output of display tubes it is the brightness or luminance of the glow which is important rather than the total light output. For example, two tubes, one with a diffuse glow and the other a concentrated glow, may give the same light output; but the one with the concentrated glow will appear brighter since the light is coming from a smaller area. Luminance is expressed as the light intensity per unit area,[1] and in English units is given in foot-Lamberts. The metric unit of luminance, which is becoming more popular, is the nit, defined as 1 candela per square metre; 1 nit = 0·292 foot–Lamberts. The luminance of the negative glow is directly proportional to the tube current, providing the glow completely covers the cathode. It also depends on the gas pressure and composition and, perhaps a little surprisingly, on the cathode material. For a normal glow in neon at a pressure of 40 torr the luminance is around 500 nits.

9.2. On-off Indicators

9.2.1. MAINS INDICATOR LAMPS FOR ELECTRONIC EQUIPMENT

Neon lamps have been employed for many years as a.c. mains voltage on-off indicators. They are designed for direct connection across the mains supply with a series limiting resistor, and therefore the geometry and gas filling must be chosen to give a breakdown potential V_s which will always be below the minimum peak voltage likely to be met in practice. There is no lower limit imposed on V_s by the circuit. However, since the energy dissipated in the series resistor depends on the difference between the supply voltage and V_m, to prevent excessive dissipation V_m should not be too low. For the same reason there is a practical limit to the current; it is usually restricted to a few milliamps. Within these limitations, however, considerable variation can be tolerated on the electrical characteristics of the mains indicator lamp, and there is no need

for the elaborate processing schedules used in the manufacture of stabilisers, trigger tubes, etc. Neon indicators, therefore, can be made very cheaply.

Early tubes tended to be large with wire electrodes in the form of spirals, or plate electrodes. To obtain complete coverage of the electrodes by the glow at reasonable currents, low pressures were employed with consequent diffuse glows and rather low luminance. With the development of more compact electronic equipment the need arose for smaller indicator lamps, and neon tubes became available of equivalent size to torch-light filament bulbs. The electrodes were normally in the form of two parallel rods which were viewed end on, and tubes of this type are still available today. Although the smaller electrodes allow for higher current densities, and therefore greater luminance, bulb blackening due to material sputtered from the electrodes is more pronounced in the smaller envelope, and for a reasonably long life the current is limited to approximately 0·5 mA. The main development over recent years has been towards producing tubes with greater luminance, 'high brightness' tubes, without reducing the life. These tubes have something like three times the brightness of the standard tubes at the same current. However, in general they will pass a higher current without shortening the life, so that in practice ten times the brightness can be obtained. A good example of a modern high brightness tube is illustrated in Fig. 9.1. The two electrodes form

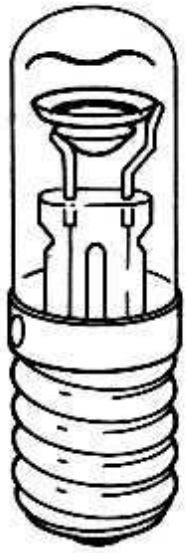


Fig. 9.1. 'High brightness' neon indicator lamp GL42M

sections of a cup facing the end viewing window of the tube. The electrodes are coated on the back with an insulating layer which ensures that the discharge only passes to the front surface. The cathode material and neon pressure chosen give a concentrated glow, and allow higher currents to be taken without excessive bulb blackening. The light from the glow is focused by the convex

lens window which is an integral part of the bulb. The improvement of the 'effective' brightness of the glow discharge by the lens is illustrated in Fig. 9.2, where the luminous flux has been plotted on a polar diagram for a Russian tube of similar structure.[2]

Most of the mains voltage indicators are neon filled giving the

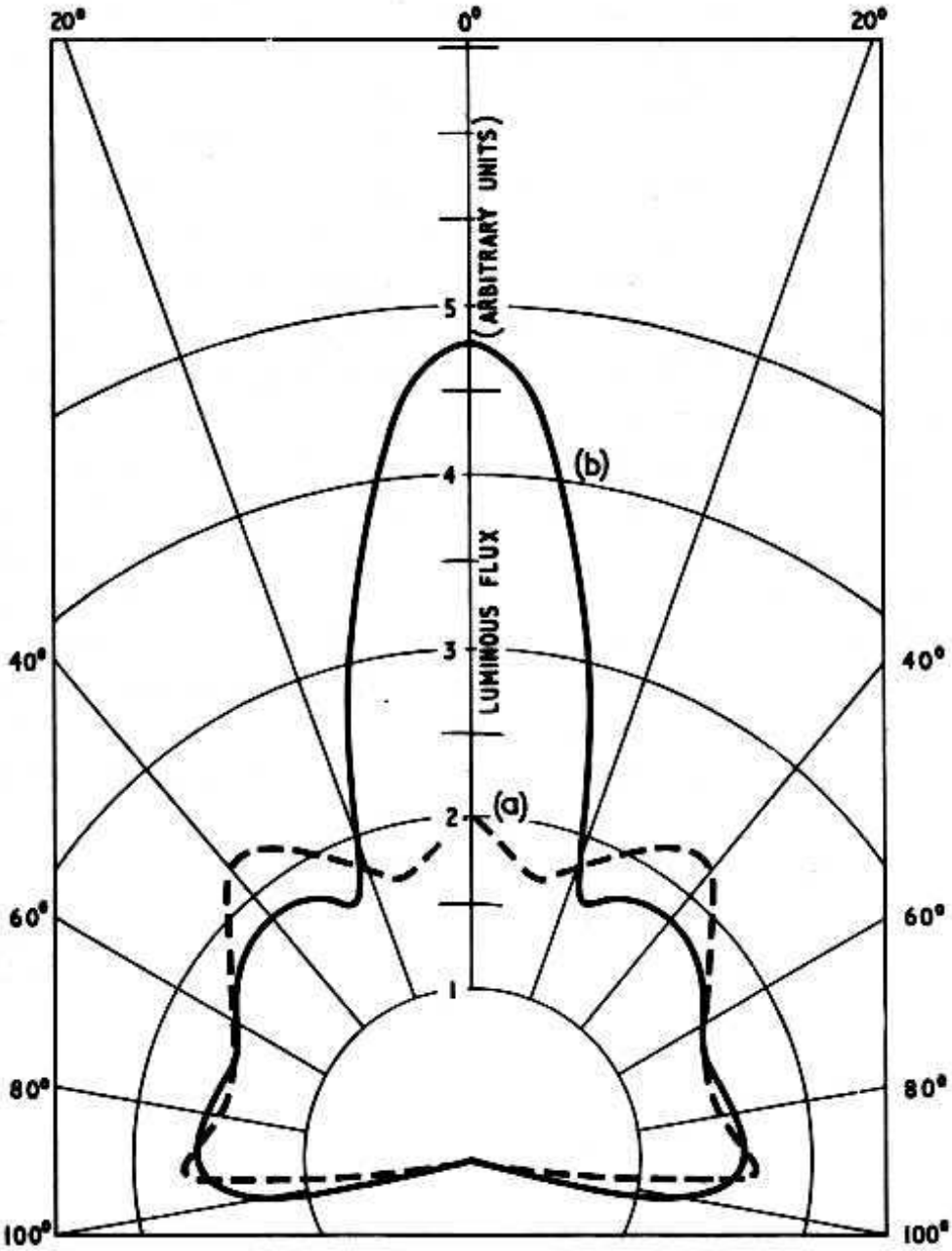


Fig. 9.2. Luminous flux distribution of tube type TN–*0–2: (a) with normal bulb, (b) with lens end to bulb. Gavanin* et al.[2]

characteristic red-orange glow. However, tubes are also available having a mercury vapour filling. The mercury discharge itself is a pale bluish colour, but the envelope is coated on the inside with a fluorescent powder which is activated by the ultra-violet radiation of the discharge to give a green indicator lamp.

9.2.2. 'ON-OFF' INDICATORS FOR TRANSISTOR CIRCUITS

One of the tendencies over the last decade has been to improve the electrical characteristics of the small neon indicator lamp by ageing schedules, so that it may be used as a cheap circuit element for logic and read-out applications (see Chapter 7). A more satisfactory approach, however, has been the design of small cheap glow discharge tubes specifically for these applications. The switching

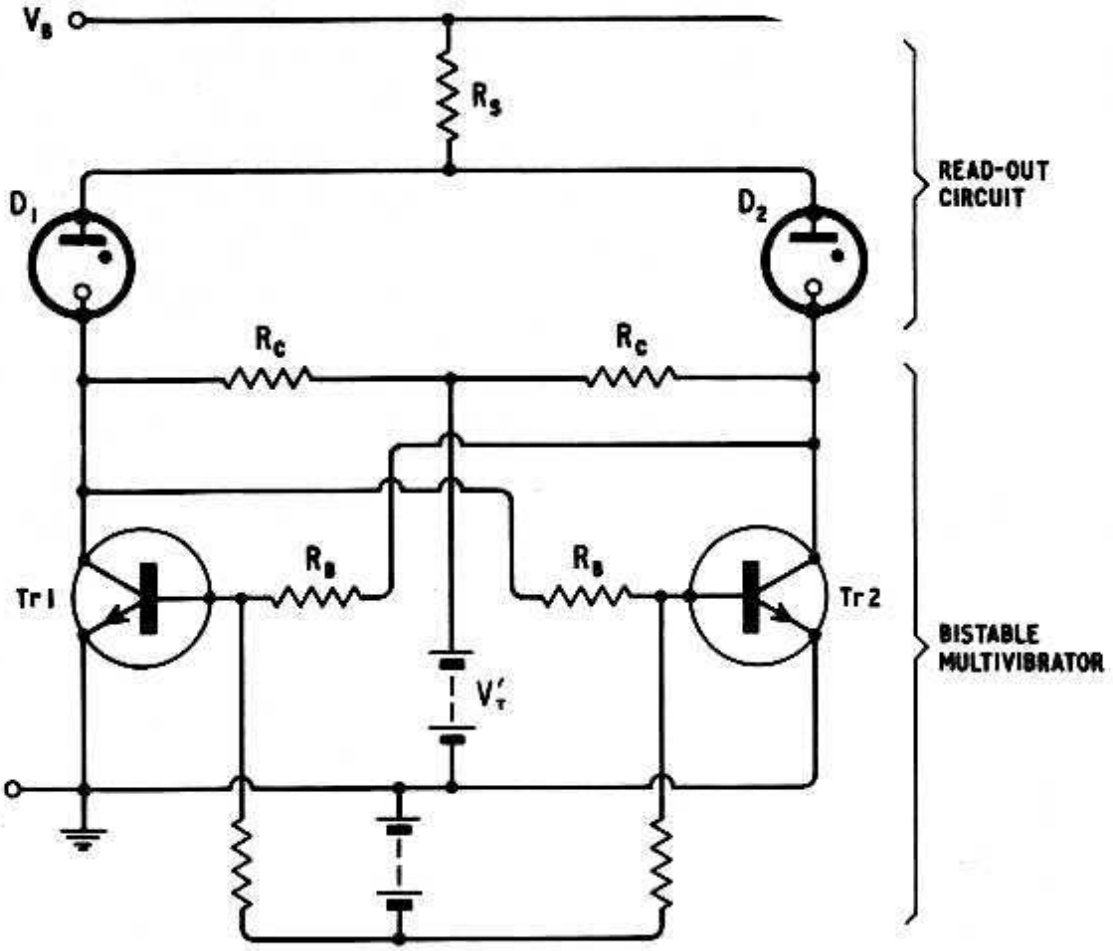


Fig. 9.3. Basic read-out circuit for a bistable multivibrator using either ZA1003 or ZA1004 diodes

diodes designed for logic circuit applications have already been described. In some of these tubes the light output is used for read-out purposes, but there are also tubes of similar design which have been specially developed to 'read-out' the state of transistor circuits and can be activated by a low voltage.

Although, by their nature, glow discharge tubes require a higher operating voltage than is normally available in transistor circuits, if they are fed from a simple auxiliary supply they can be designed to be satisfactorily controlled by the small signals obtained from transistors. Two such tube designs are available: one type is

switched on and off in the conventional manner by the input signal, whilst in the other type the signal transfers the glow from one cathode which is hidden to another which is visible, without actually extinguishing the glow.

In the former type of tube, the low control voltage depends on achieving a breakdown potential V_s, a maintaining potential V_m, and an extinguishing potential V_e which all lie close together and remain extremely constant during life. In addition, the difference between the characteristics of individual diodes must be small. Two tubes of this type are the ZA1003 and ZA1004 described in Section 7.2. Both are sub-miniature tubes having the central molybdenum rod cathode and coaxial cylindrical mesh anode of the ZA1000 series.[3] They are filled with a neon-argon Penning mixture to give the smallest differences between V_s, V_m and V_e, and the cathode is thoroughly sputtered to give stability of the characteristics over life. The ZA1003 can be reliably switched by a 12 V signal whilst the ZA1004 only requires a 6 V signal.

The basic read-out circuit using either the ZA1003 or ZA1004 is shown in Fig. 9.3. The read-out diodes D_1 and D_2 are controlled by two *npn* transistors *Tr*1 and *Tr*2, which form part of a bistable multivibrator. The transistors are fed from a common supply

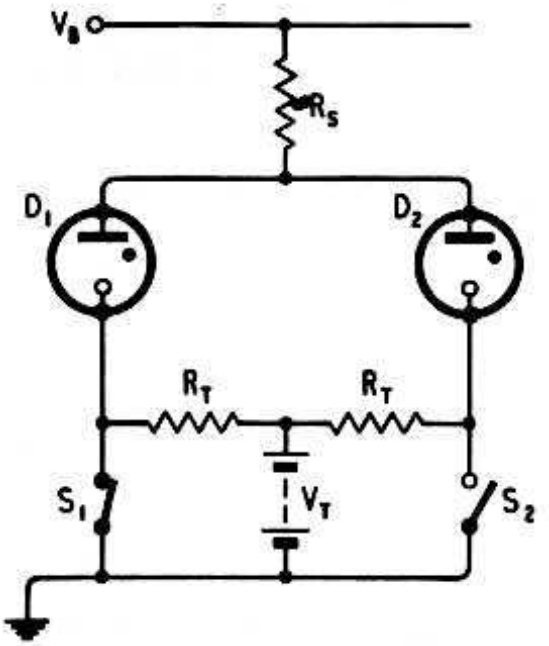


Fig. 9.4. Equivalent circuit of Fig. 9.3 in which the transistors are represented by switches

source V_T', and the diodes from the auxiliary supply V_B via a common resistor R_s. Correct operation of the circuit implies that during read-out the diode corresponding to the bottomed transistor is conducting whilst the other diode is non-conducting. The circuit must therefore be so designed that there is no possibility of the wrong diode conducting or of both tubes conducting. The diodes do not necessarily have to follow the changes of state of the multi-

vibrator so long as they register the final static state. Kerr, Poorter, and van Vlodrop[4] have studied this circuit and given design data for use with ZA1004 diodes.

For the sake of clarity Fig. 9.3 can be simplified to the equivalent circuit of Fig. 9.4 where the transistors $Tr1$ and $Tr2$ are represented by switches S_1 and S_2. In this circuit V_T represents the no-load

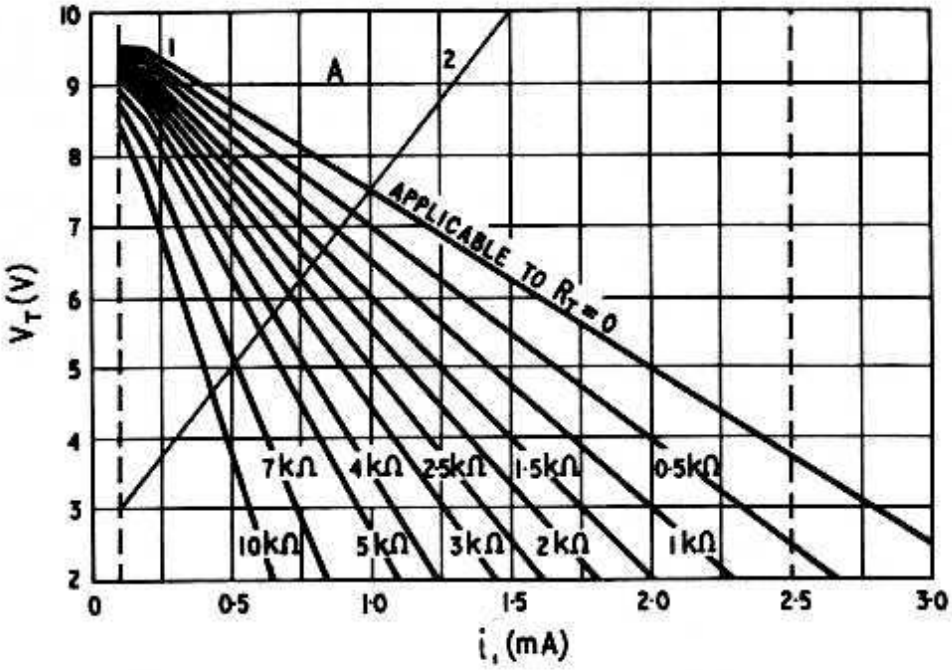


Fig. 9.5. Diagram of V_T, the supply voltage of the switching circuit in Fig. 9.4, against current, in which expression 9.1 for various values of R_T (curves 1) and expression 9.2 for the ZA1004 (curve 2) are plotted. Kerr, Poorter and van Vlodrop[4]

collector voltage measured on the cut-off transistor minus the saturation voltage of the bottomed transistor, and is therefore rather lower than V_T'. Also R_T represents the output resistance as measured at the collector of the cut-off transistor and, as such, is approximately equal to $R_C R_B/(R_C+R_B)$. Kerr *et al.* showed that for correct indication the following conditions had to be satisfied.

$$V_{m(\min)} + i_1 R_T + V_{T(\min)} \geqslant V_{s(\max)} \tag{9.1}$$

and

$$V_{m(\max)} - V_{T(\min)} \leqslant V_{e(\min)} \tag{9.2}$$

where i_1 is the current through the conducting diode D_1. Since V_m is a function of the tube current both the above expressions can be represented on a V_T against i diagram. This is illustrated in Fig. 9.5 where curve 1 is derived from expression (9.1), for various values of R_T, and curve 2 is derived from expression (9.2), using published data for the ZA1004. The area above the curves for a particular value of R_T, i.e. region A, represents the field within which reliable operation is ensured. From this diagram the permissible current

range for any given V_T can be determined, and the value of V_B and R_s deduced, allowing for circuit tolerances.

For direct current supply to the diodes the minimum required signal voltage V_T for the ZA1004 is approximately 5 V. However, Kerr *et al.* pointed out that the minimum signal depended on the nature of V_B, and that if an interrupted d.c. supply, such as full-wave rectified a.c., is employed then the extinguishing voltage does not have to be considered, and the minimum signal could be reduced to 4·5 V. Further, if a supply consisting of a combination of d.c. and a.c. of suitable pulsed waveform were employed, the diodes could be switched with only 3 V.

The indicator in which the glow is transferred from one stable position to another without it being extinguished, the TG121A, has been described by Omi, Fukukawa, and Nakajo.[5] The construction of the tube is shown schematically in Fig. 9.6(a). It consists of a

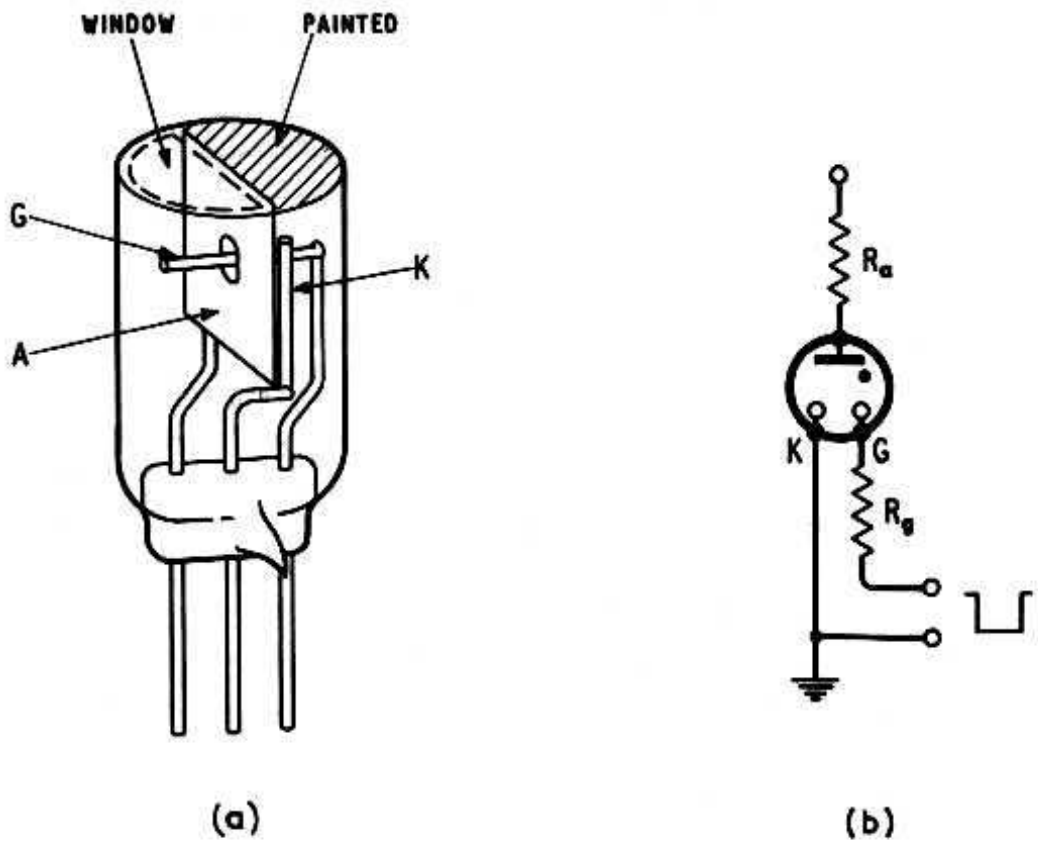


Fig. 9.6. Glow-transfer indicator tube TG121A, *(a) schematic view of the construction, (b) basic drive circuit. Omi, Fukukawa and Nakajo*[5]

nickel anode *A*, which substantially divides the tube into two compartments, and two molybdenum cathodes—an indicating cathode *G*, and a sustaining cathode *K*. The anode has a small hole at its centre through which the indicating cathode projects so as to be observed through the window at the top of the bulb. The rest of the bulb is painted black. On the other side of the anode, the

sustaining cathode is mounted close to the indicating cathode so that one cathode lies within the glow of the other. As a result the anode current is divided between the two cathodes, according to their sizes, shapes, and the potential difference between them. The tube is filled with a neon-xenon mixture, and the total size is about 8 mm diameter and 18 mm long.

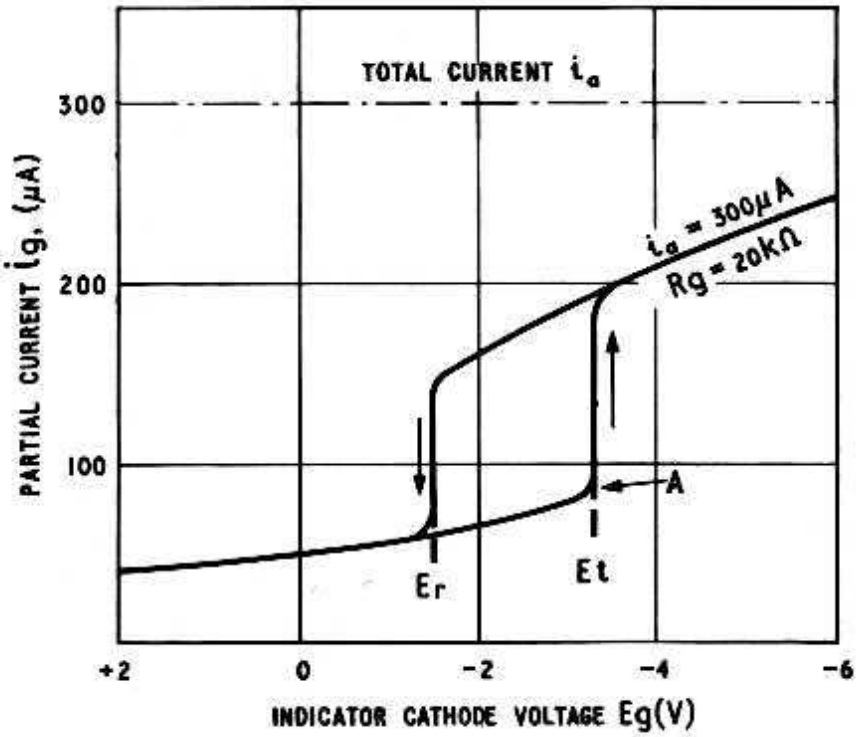


Fig. 9.7. i-V characteristic of the glow-transfer indicator tube

The circuit arrangement for the tube is shown in Fig. 9.6 (b). In the 'off' state the two cathodes are connected to earth, but the resistor in the indicating cathode lead, R_g, restricts its current to a small value, and most of the current passes to the sustaining cathode. When a negative potential is applied to the indicating cathode, the current to it increases, and if the potential is large enough the glow will entirely cover the indicating cathode and be visible through the window. Since the current to each cathode follows the V-i characteristic of a normal-glow discharge only a small voltage is required to produce a large current change (cf. Chapter 3) and thus the increase in current to the indicating cathode and the decrease in that to the sustaining cathode only requires a few volts difference. When the glow expands along the indicating cathode wire to reach the anode hole, the closeness of the anode to the cathode causes higher electron and ion losses from the discharge at the hole and a barrier is presented to further expansion of the glow. As a result the increase of current with bias voltage is not completely continuous, but shows an abrupt

change in current as the glow expands or contracts through the hole. This is illustrated in Fig. 9.7 where the i-V characteristics for the indicating cathode–anode gap is plotted. It will be seen from this figure that there is a hysteresis effect, and that the voltage required to switch the indicator on is higher than that required to turn it off. E_t represents the minimum signal required to switch the tube. However, since it is the potential V_c at the cathode which is important for the transference of the glow, and this is modified by the voltage drop across the cathode resistor, i.e. $V_c = E_t - i_g R_g$, the value of E_t will depend on the tube current and also on the value of R_g.

Typical ratings for the glow transference tube described are as follows:

Anode breakdown potential		150 V(max)
Maintaining potential at $i_a = 250\ \mu A$ to $400\ \mu A$		$112\ V \pm 3\ V$
Indicator transfer voltage		
$i_a = 250\ \mu A$	$R_g = 20\ K\Omega$	-5 V(max)
$i_a = 400\ \mu A$	$R_g = 20\ K\Omega$	-4 V(max)

The tube can be used with a transistor multivibrator in which case the cathode resistor is not required, and the indicator can be switched with only 3 V. The circuit is shown in Fig. 9.8.

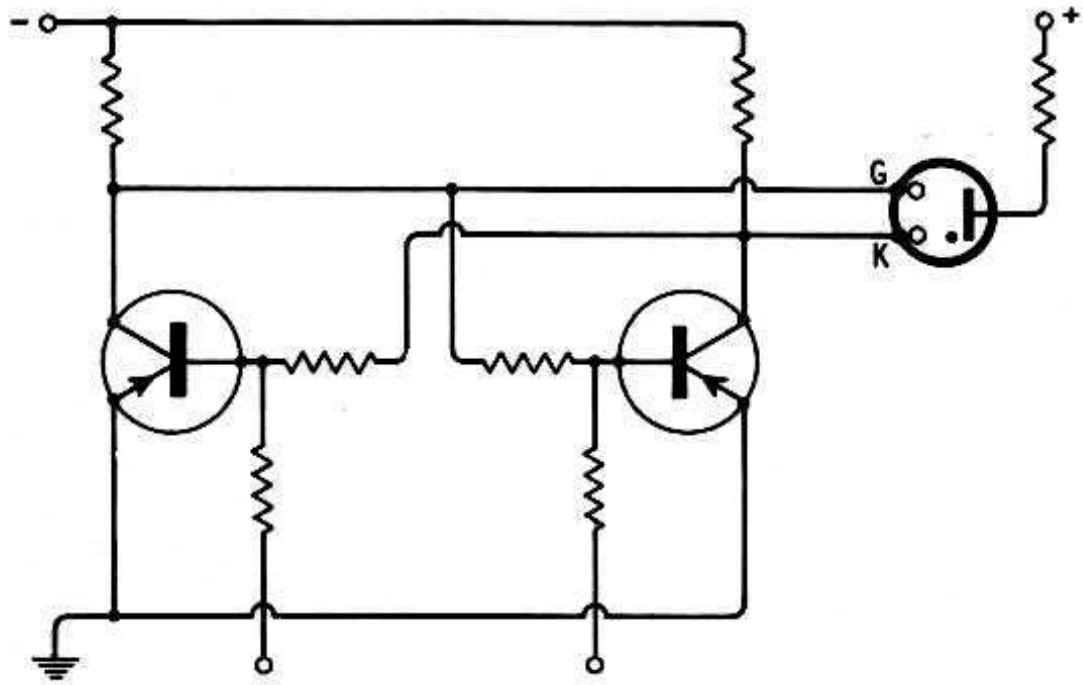


Fig. 9.8. Read-out circuit for a bistable multivibrator using the TG*121*A

The tube has been life tested both in standby and in switching operation for 10,000 hr without appreciable change in the characteristics.

The main objection to the tube has been the rather low brightness

of the glow. Recently an improved version has been described by Fukukawa[6] which can be run at a higher current (800 μA). A further improvement in the brightness is also obtained by the constructional design which is shown in Fig. 9.9. The indicating cathode is disk-shaped and illuminates the full diameter of the tube when viewed through the end dome of the envelope.

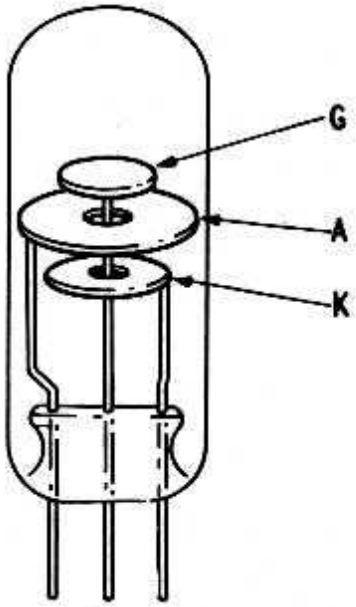


Fig. 9.9 Construction of the improved glow-transfer indicator. Fukukawa[6]

An alternative method of transferring a glow from one cathode to another by a low signal voltage is employed in the clock face positional indicator tube described by Botden[7] and summarised in the next section. The principle used in this tube, in which an intermittent discharge is transferred from one cathode to another by a small d.c. signal applied to an appropriate auxiliary anode, could equally well be applied to an on-off indicator. A cathode having two sections to which the glow can go would then be required, one being screened from view.

9.3. Glow Position Indicator Tubes

In scaling applications where the counting rate is too high for cold-cathode tubes, the first stage can be a thermionic or solid-state decade counter, whilst later stages are provided by stepping tubes. In the interests of uniformity, an indicator tube is then required for this first stage which will give a read-out of similar appearance to that of the stepping tubes. The glow position indicator was developed for this purpose, being constructionally similar to a selector type stepping tube but without the guide cathodes. It has ten rod cathodes symmetrically spaced around the same pitch

circle as used for the corresponding stepping tube, and is mounted in a similar bulb. The anode, however, is usually larger to reduce the anode-cathode spacing, and thus lower the breakdown potential.

In normal operation the anode is connected via a suitable current limiting resistor to the h.t. whilst the cathodes are at a common potential, e.g. earth. Initially when the tube is switched on, the voltage across the tube will be above the breakdown potential, and a discharge may be initiated to any one of the ten cathodes. The potential across the tube will then fall to the maintaining potential and no other anode–cathode gap will break down. To transfer the discharge from the conducting cathode to a selected non-conducting cathode a negative bias is applied to the latter of sufficient amplitude to cause breakdown; the required bias will be equal to $V_s - V_m$. The discharge to the biased cathode will extinguish the previous anode–cathode discharge, and thus the tube will register by glowing on the cathode to which the signal is applied.

Since transfer characteristics and switching speed are unimportant in the indicator tube, a different gas filling to that of the corresponding stepping tube can be employed. In fact the only requirement imposed on the gas filling in terms of compatibility with stepping tubes is that the glow should be of similar colour and brightness. The important consideration in the choice of gas is the desirability of switching the glow to the appropriate cathode with the minimum signal i.e. the lowest value of $V_s - V_m$. The Penning mixture of neon plus a small admixture of argon (0·1–1 %) gives the lowest value of $V_s - V_m$ for inert gases, particularly if the pressure is chosen so that the breakdown potential is at the Paschen minimum (cf. Section 6.3). It is, therefore, the most suitable gas filling, having the added advantage of being the most satisfactory gas for stable characteristics over life. The latter is important, since the minimum input signal is equal to the maximum breakdown potential likely to be encountered in any tube, initially and during life, minus the minimum value of the maintaining potential, also during life.

The electrode material is similarly chosen to give stable values of V_s and V_m during life, bearing in mind that the glow must be visible and therefore the cathodes cannot be extensively sputtered during processing. As was seen in Section 8.2.2, nickel is the most satisfactory material for cathodes in tubes where adjacent cathode contamination is a problem, and it is therefore normally used in the glow position indicator tubes. Values of breakdown and maintaining potentials for the Z503M which is typical of this design of glow position indicator are $V_{s(max)} = 129$ V and $V_{m(min)} = 105$ V. The tube can therefore be operated by the output from

relatively cheap 30 V transistors. Further, the recommended anode current is only 60 μA so that the tube consumes very little power.

A glow position indicator of a rather different construction which can be switched by as little as 5 V has been described by Botden[7] and is shown schematically in Fig. 9.10. The tube had a

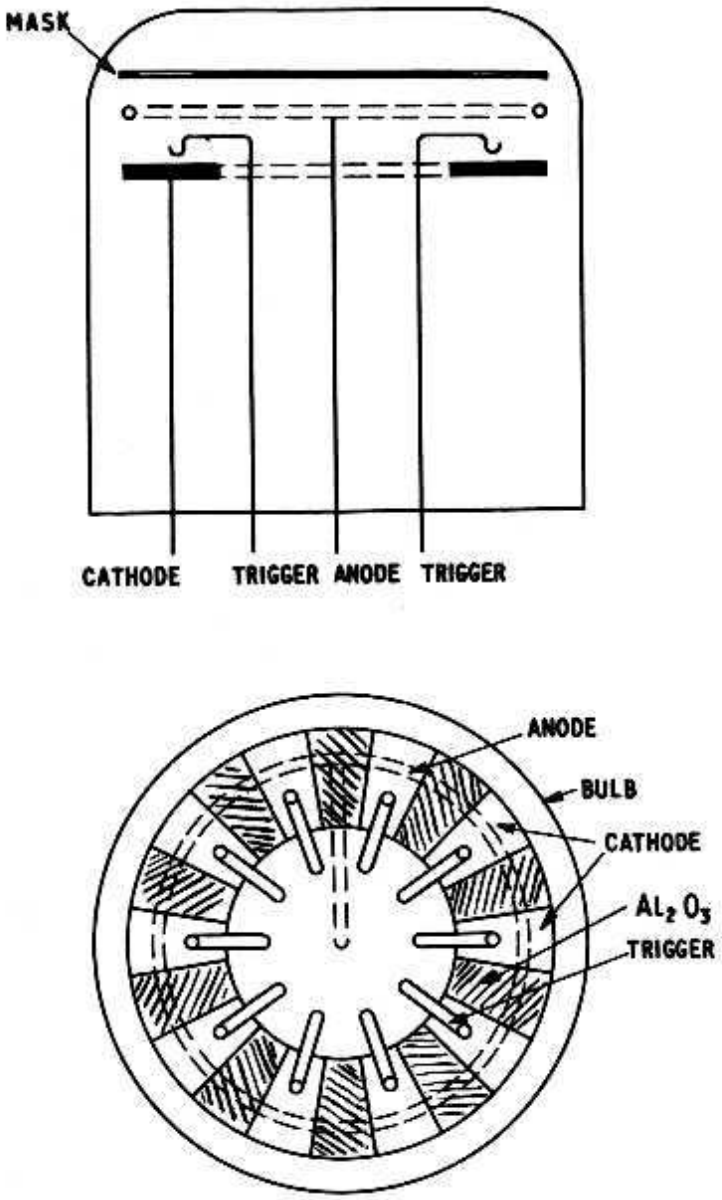


Fig. 9.10. Schematic diagram of the glow position indicator which can be switched by a low voltage (~ 5 V) signal. Botden[7]

circular wire anode of nickel mounted 5 mm from a molybdenum cathode in the form of a flat annular disk. The surface of the cathode was divided into ten parts by coated strips, which restricted the glow discharge to the uncoated areas; a thin Al_2O_3 coating was employed. Close to each uncoated cathode area was a wire trigger electrode, each of which was accurately mounted at the same distance from the cathode plane (1 mm or less). A mask plate

in which the numbers 0–9 were perforated was placed above the anode ring, so that the glow illuminated the appropriate stencilled number.

The tube was filled with the Penning mixture 99·9% neon + 0·1% argon at a pressure of 150 torr, which according to Fig. 2.10 gives the minimum breakdown potential for a 1 mm gap. The Paschen curve for this particular mixture is very 'flat' and a variation, therefore, in the trigger-cathode spacing will make very little difference to the breakdown potential; a variation of ±20% in the spacing for a 1 mm gap will change V_s by less than 1 V for plane-parallel geometry and similar results were obtained using a plane cathode with a wire anode (see Fig. 2.10). The cathode can be fairly thoroughly cleaned by sputtering since the mask acts as a screen to prevent the sputtered material being deposited on to the viewing window. Thus stable and closely similar values of breakdown and maintaining potential can be obtained in each anode–cathode gap, so that a small signal on the trigger is sufficient to ensure breakdown to any selected cathode area.

The operation of the tube can be explained with reference to the circuit diagram of Fig. 9.11. The anode is at earth potential, and a half-wave rectified voltage of peak amplitude greater than the anode–cathode breakdown potential is applied to the cathode

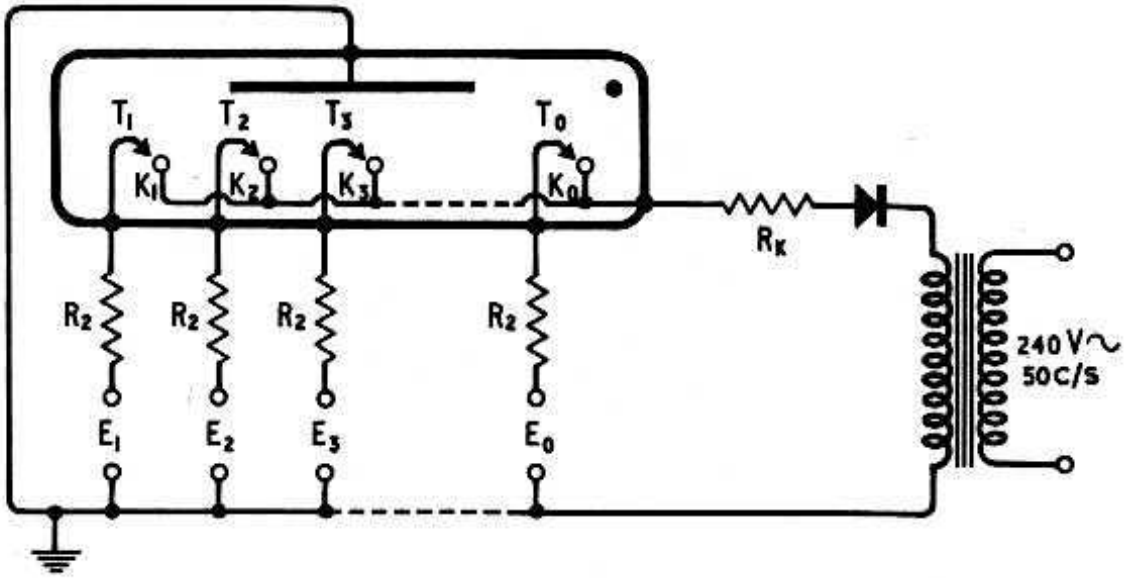


Fig. 9.11. Basic drive circuit for the low-voltage glow-positional indicator ZM*1050*

(negative half-cycles). The resistance R_K limits the current so that the glow is normally only invested on one portion of the cathode at any one time. The triggers T_1, T_2, . . . T_9, T_o, are connected to the anode via resistors R_2 (of 0·1–1 MΩ) and the respective voltage source E_1, E_2, . . . E_9, E_o. The voltage source represents the input signal which in practice would be derived from a transistor in, say,

a ring counter. As the voltage increases between anode and cathode on the negative half-cycle, it simultaneously increases between the triggers and cathodes. Applying a bias to the selected trigger of a few volts ensures that that particular trigger-cathode gap will ignite before any other, and the discharge current will prime the associated anode–cathode gap. If the priming current is sufficient, immediate breakdown of this anode–cathode gap will occur. The potential across the tube is then reduced to the maintaining potential, preventing breakdown to any other portion of the cathode. At the end of the half-cycle the discharge is extinguished. The glow will rest on the selected cathode area during each subsequent half-cycle, providing the corresponding trigger bias is maintained.

From the experimental tube Botden found that the necessary bias potential varied initially, from tube to tube and from trigger to trigger, from a minimum of less than 1 V to a maximum of 4 V. Although changes occurred during life, the required bias did not increase above 5 V over a period of 6 months. The trigger current necessary to lower the breakdown potential of the anode–cathode gap to that of the trigger-cathode gap was about 10 μA. In the commercial version, ZM1050, a signal of 5 V and 50 μA is recommended. The anode–cathode current of 3 mA provides adequate brightness for the numerals to be clearly visible in daylight.

The tube is conveniently supplied via a transformer from the a.c. mains at 50 c/s, but it can be operated at higher frequencies. If the frequency rises above 500 c/s, however, a larger triggering bias is required due to the residual ionisation. The maximum frequency limited by the deionisation time is found to be 3 kc/s.

9.4. Character Display Tubes

9.4.1. GENERAL CHARACTERISTICS

Probably the most successful glow discharge indicator or display tube has been that developed for in-line read-out, in which the character or number is 'picked out' by the negative glow of a cold cathode discharge.

In-line digital read-out has become increasingly popular in electronic measuring instruments, business machines, and industrial remote indicating applications. It provides an unambiguous display which is free from the reading errors, due to confusion of scale or parallax, associated with meters and positional indicators. The success of the glow-discharge display tube over other devices for

in-line read-out lies in the intrinsic reliability and long life of the cold-cathode discharge, the low cost, and the brightness of the display which is clearly visible in daylight with relatively low input powers (< 0·5 Watts per digit). It also has the advantage of compatibility with both thermionic valve and solid-state circuits.

The presentation of the characters in such tubes may be achieved either by a matrix of separately connected cathode segments mounted in one plane, a suitable combination of which form the required character,[8] or the cathodes themselves are shaped to form the character and stacked one behind the other to give a range of characters in a single tube.[9] Since, in the latter design, the cathodes must be insulated from each other, there is a limit to the number that can be stacked in one tube due to the depth of the stack. If this is too great, the spacing between the plane of the first and last character in the stack would be unacceptable when several tubes are used to display a line of characters. The matrix display tube is less restricted in the number of characters which can be displayed in one tube, and of course all the characters appear in the same plane. However, the character shape for some of the numbers and letters tends to be rather poor, and further, the selection of the required combination requires a complex coding circuit. For these reasons, the stacked-cathode tube is normally used for the display of numerals and of associated signs such as + or −, and the matrix tube is confined to applications where letters of the alphabet are also required.

9.4.2. THE NUMERICAL INDICATOR TUBE

The term numerical indicator tube is commonly applied to the stacked-cathode type of display tube, and will be adhered to in this book. The most obvious use for these tubes is as read-out devices for decimal counting systems where the tubes will contain the digits 0–9. A diagram of the basic construction is shown in Fig. 9.12. Each cathode in the form of a number has a tag at the top and bottom with locating holes. The cathodes are threaded on to two insulated rods via these holes with ceramic or glass spacers in between, so that each cathode is electrically isolated; electrical connection to the individual cathodes is taken out separately through the valve base. The assembly of ten cathodes is surrounded by a common anode structure, the front and part of the sides of which are in the form of a mesh with a good optical transmission. Because the anode completely surrounds the cathodes, the spacings and thus the breakdown potentials between anode and each

cathode are similar. Also the anode mesh helps to prevent sputtered material depositing on to the bulb in front of the numbers (see Section 4.4.4). Because the visible glow is considerably wider than the metal numeral on which it occurs, the selected number is clearly visible and not appreciably obscured by any other cathode which may be in front of it.

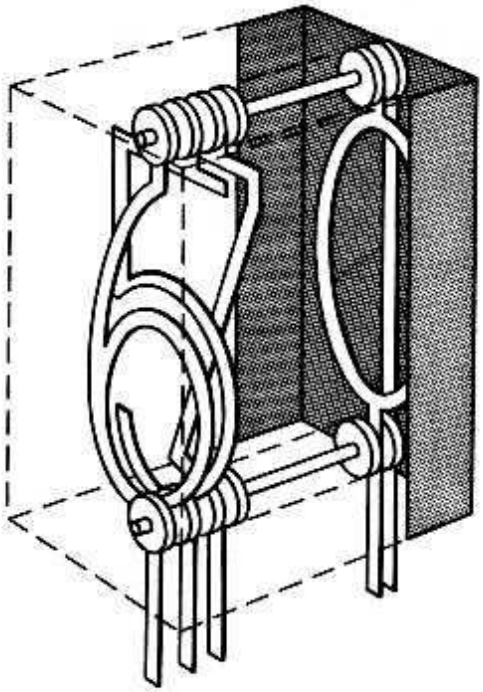


Fig. 9.12. Schematic diagram of the electrode construction of the stacked-cathode numerical indicator

The numbers can be mounted suitably within the envelope for viewing either through the end of the bulb or through the side. The size of the numbers can be varied within wide limits to suit the application, and tubes are available with numerals ranging from 8 mm high suitable for desk instruments to 50 mm high for viewing from a distance.

To ascertain the design requirements for the numerical indicator the basic operation conditions of the tube must be considered. The simplest method of operating the tube is shown in Fig. 9.13 (a). Here the anode is connected via a limiting resistor R_a to a positive potential V_B, which is above the breakdown of the tube, and the cathode is selected by switching it to earth. The non-conducting cathodes are left disconnected, i.e. electrically floating. This is not a very practical circuit, and a more useful arrangement is shown in Fig. 9.13 (b), where a bias is applied to all the cathodes and the required cathode is selected by pulsing it negative. Here the chief problem is to ascertain the voltage difference required between the selected cathode and the rest of the cathodes for satisfactory indication. In any tube having an array of cathodes with a common anode, the current is shared between the cathodes, the percentage passing to each cathode depending on its position and relative

potential. In the case of the numerical indicator tube, if the negative potential applied to the selected cathode is insufficient, the current to the other figures will produce a background glow and reduce the clarity of the indication. Thus a minimum potential difference can be defined where this background 'haze' becomes objectionable. The other electrical parameter of importance is the tube current. This must be large enough for the glow to cover uniformly the cathode surface, without the possibility of it extending on to the connecting wires.

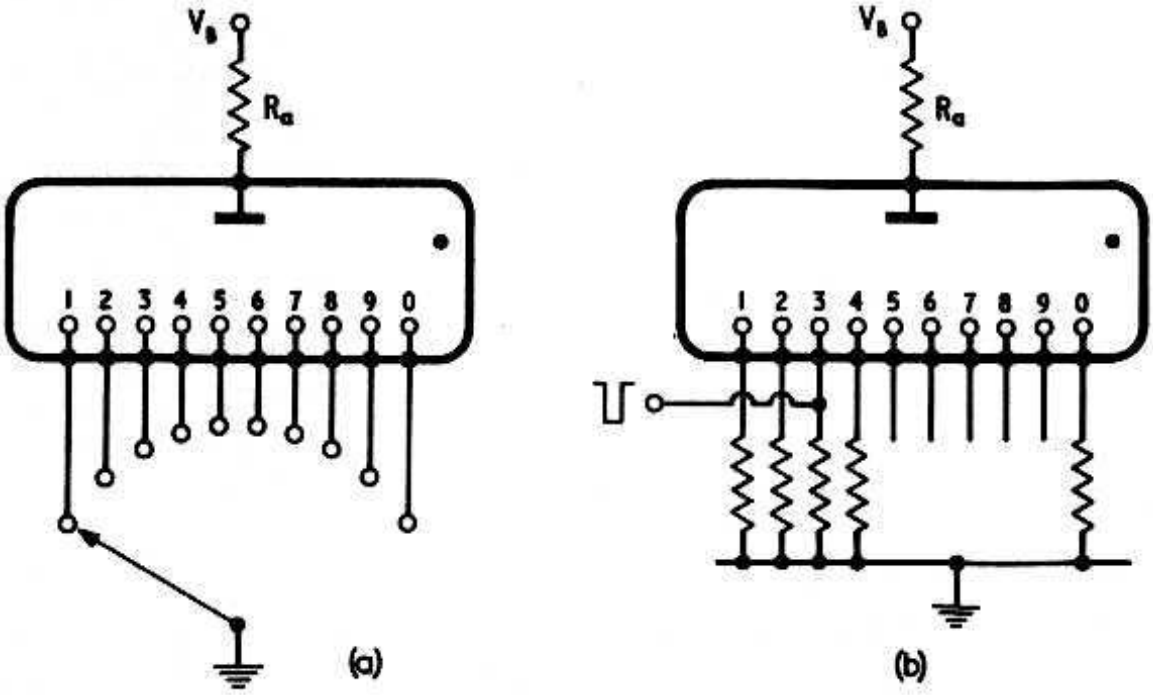


Fig. 9.13. Basic drive circuits for the numerical indicator

9.4.3. THE DESIGN AND CONSTRUCTION OF THE NUMERICAL INDICATOR TUBES

As with all glow discharge tubes, the electrical characteristics and life performance of a particular design depends on the gas filling, the material and geometry of the electrodes, and on the manufacturing procedure.

The gas filling is chosen not only to give the optimum electrical performance, but also from the subjective point of view of giving the best presentation. The glowing numbers should be clearly visible in a normally lit room, and have a well-defined outline. Further, to ensure that the number is not obscured by the cathodes in front of it, the glow should be 4 or 5 times the width of the cathode strip.

From the point of view of visual brightness the neon glow is

vastly superior to that of any other inert gas, and practically any gas mixture containing a reasonably large percentage of neon will give a satisfactory display, providing the correct pressure is chosen. However, there is a practical limit to the choice of gas set by the tube current; since the tube is usually driven by thermionic valves, trigger tubes or semiconductors, currents in excess of a few milliamps would be unacceptable. Therefore the current density in the normal glow must be such that the glow will cover the whole cathode at a few milliamps; for a standard tube with 15 mm high numerals a covering current of 1·5 mA gives a current density of roughly 5 mA cm^{-2}. The way in which the current density in the normal glow region depends on the gas composition and pressure is illustrated in Fig. 9.14, where it has been plotted against the

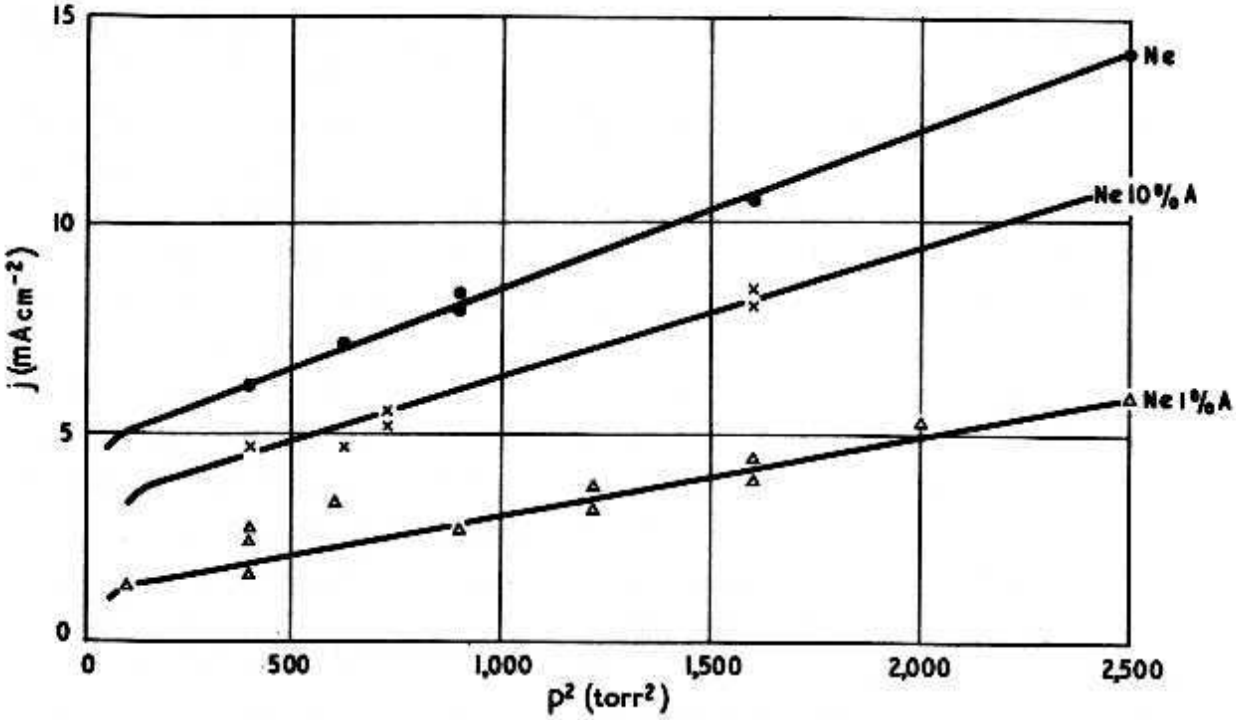


Fig. 9.14. Current density as a function of the pressure squared for neon and neon-argon mixtures in experimental tubes having 0·25 mm diameter molybdenum rod cathodes

square of the pressure for neon and two neon-argon mixtures. Although the values in Fig. 9.14 are not directly applicable to the numerical indicator tube since they were obtained from experiments with molybdenum rod cathodes, nevertheless they do give a rough indication that if a current density of the order of 5 mA cm^{-2} is required, then for pure neon, the pressure would have to be low, of the order of 10 torr. On the other hand for the Penning mixture with 1 % argon the pressure could be as high as 40–50 torr. Since the glow is very diffuse at low pressures, the outline would be too 'fuzzy' in pure neon for an acceptable presentation; generally, the

outline does not become sharp enough until the pressure is raised to about 30 torr. Thus, unless higher currents are used, 99% neon + 1% argon is the most suitable gas mixture for the numerical tube. This gas mixture has the added advantages of low breakdown and maintaining potentials which will allow for lower supply voltages, and also the required voltage differences between the selected cathodes and the other cathodes are relatively small, 30 V–60 V depending on the circuit (see Section 9.4.4).

The higher pressure is also desirable from the point of view of life. The end of life of a tube of this type is reached when the glow either no longer completely covers the required character or spreads on to some other part of the tube. Two effects can bring about this state of affairs, contamination or sputtering. The experience gained in the processing of stabilisers and trigger tubes, which was discussed in some detail in Chapters 6 and 7, can be used to minimise contamination effects. The effect of sputtering, however, cannot be so easily dismissed, and since the sputtering rate increases markedly with decreasing pressure there is advantage in keeping the pressure as high as possible in spite of the difficulty of cleaning cathodes at such pressures. Not only does the sputtering form deposited layers on the insulators, causing short circuits or extension of glow on to the deposits, but also, since the other cathodes are in close proximity, it causes adjacent cathode contamination (see Section 8.2.2). So serious was the sputtering problem in the early tubes that lives of more than 1,000 hr could not be expected if the glow remained stationary on one digit.

Investigation of the sputtering difficulties in numerical indicator tubes showed that if a small amount of mercury was introduced then a considerable improvement in the life performance could be achieved. The improvement was due to a reduction of the sputtering rate in the presence of the mercury and is illustrated in Fig. 9.15, where the weight of material sputtered from various metal cathodes in a neon-argon glow discharge was measured under controlled conditions. In the test, the sputtering rate of the nickel-iron alloy was so reduced by the introduction of the mercury that even after 1,500 hr no detectable change in weight could be measured. The only disadvantage of the mercury is the blue haze formed round the negative glow, but this can be filtered out by mounting a suitable filter in front of the tube or by coating the bulb with a red transparent layer. Care must be taken not to introduce too much mercury, as it can condense on to the bulb in the form of small globules obscuring the display.

Since adjacent cathode contamination is as much a problem in the numerical indicator tube as in the stepping tube, nickel would

appear an obvious choice for the cathode material, and indeed it is used by some manufacturers. However, as the cathodes are rather fine and easily distorted during assembly there is advantage in using a more rigid material. Nickel-iron alloys or stainless steel, which have this advantage, have proved to be suitable alternative materials for the cathodes giving similar initial characteristics and behaviour on life. The numerals are either etched or stamped out from sheet material, their dimensions being chosen partly for their performance in the tube, and partly from the more subjective

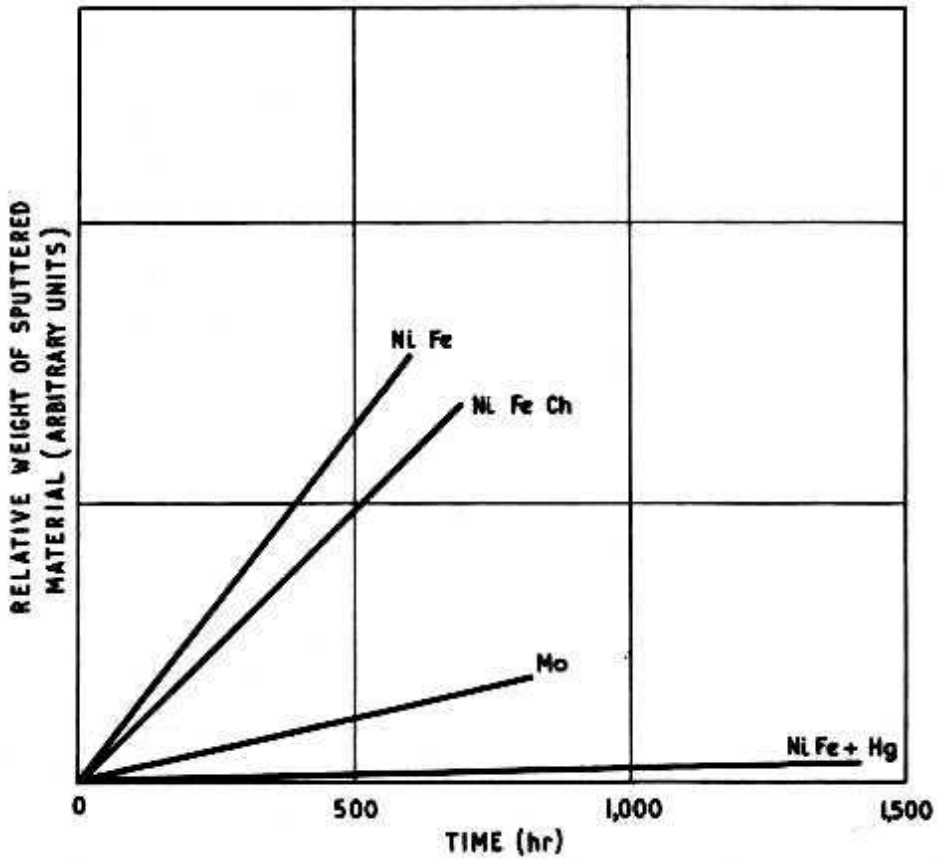


Fig. 9.15. Weight of sputtered material as a function of time in a 99% neon – 1% argon discharge with various cathode materials

criterion of the aesthetic appearance. Thus for the wide-viewing tubes, where the numbers can be mounted close together, a height to width ratio H : W of 2 : 1 is normally chosen, whilst for the end-viewing tubes, where the numbers will be further apart, a more square format, with H : W of 1:1·6, is preferred.

Clearly the cathodes should be stacked as closely as possible to minimise the change in apparent position of the number when switching from one to another. However, since the cathodes are not particularly substantial, allowance must be made for the possibility of slight distortion both on assembly and during operation. The latter can occur because of heating of the cathode; even with only 0·2 W the cathode temperature can be a few hundred

degrees centigrade. For the standard tube the minimum spacer width is about 0·6 mm, allowing a stacking depth of about 7 mm. The spacers are normally made of glass or ceramic and chamfered to prevent electrical leakage from any deposited metal layer. Even though the stacking depth is small, it is possible for the numbers at the centre of the stack to have rather higher values of breakdown and maintaining potential than the numbers at each end. It is common practice, therefore, to insert an extra electrode (often an '8' numeral) halfway down the stack electrically connected to the anode potential.

The design of the anode box is not critical, providing it is fairly evenly spaced around the cathode stack, but it can be optimised in relation to the visible display. Thus the solid back portion is normally coated or suitably processed to give a dark matt surface to reduce reflections, and the front mesh portion has been designed to give the maximum light transmission compatible with the required rigidity. Plate 9.1 shows some examples of numerical indicator tubes available in the U.K.

9.4.4. OPERATION AND PERFORMANCE OF THE NUMERICAL INDICATOR TUBE

At low currents, where the discharge is in the normal glow region, the maintaining potential is constant and the area of the negative glow increases in size as the current is increased to maintain approximately the same current density. This continues until the whole cathode is covered with the glow, when the discharge becomes abnormal. In this abnormal glow region the maintaining potential increases with the increasing current density (for a more detailed description see Chapter 3). In the case of a numerical indicator tube, therefore, the lower limit of current for operating the tube is just above that necessary for the glow to cover the cathode completely. The tube, however, should not be operated at too high a current because of the increase in the cathode sputtering rate, which as previously noted is closely tied to the life expectancy of the tube. Other considerations which limit the maximum current are that the glow should be prevented from spreading on to the cathode lead-in wires, and that the total power dissipated should be kept low enough to avoid overheating the bulb with the consequent release of contamination. The current limits are quoted by the manufacturer for each tube type, and it is important to observe strictly their recommendations. This implies choosing suitable values of anode supply voltage V_B and anode resistor R_a.

If the circuit diagram of Fig. 9.13 (a) is considered the current is given by

$$i = \frac{V_B - V_m}{R_a}$$

where V_m is the maintaining potential of the tube. The maintaining potential, however, is a function of the current over the full glow

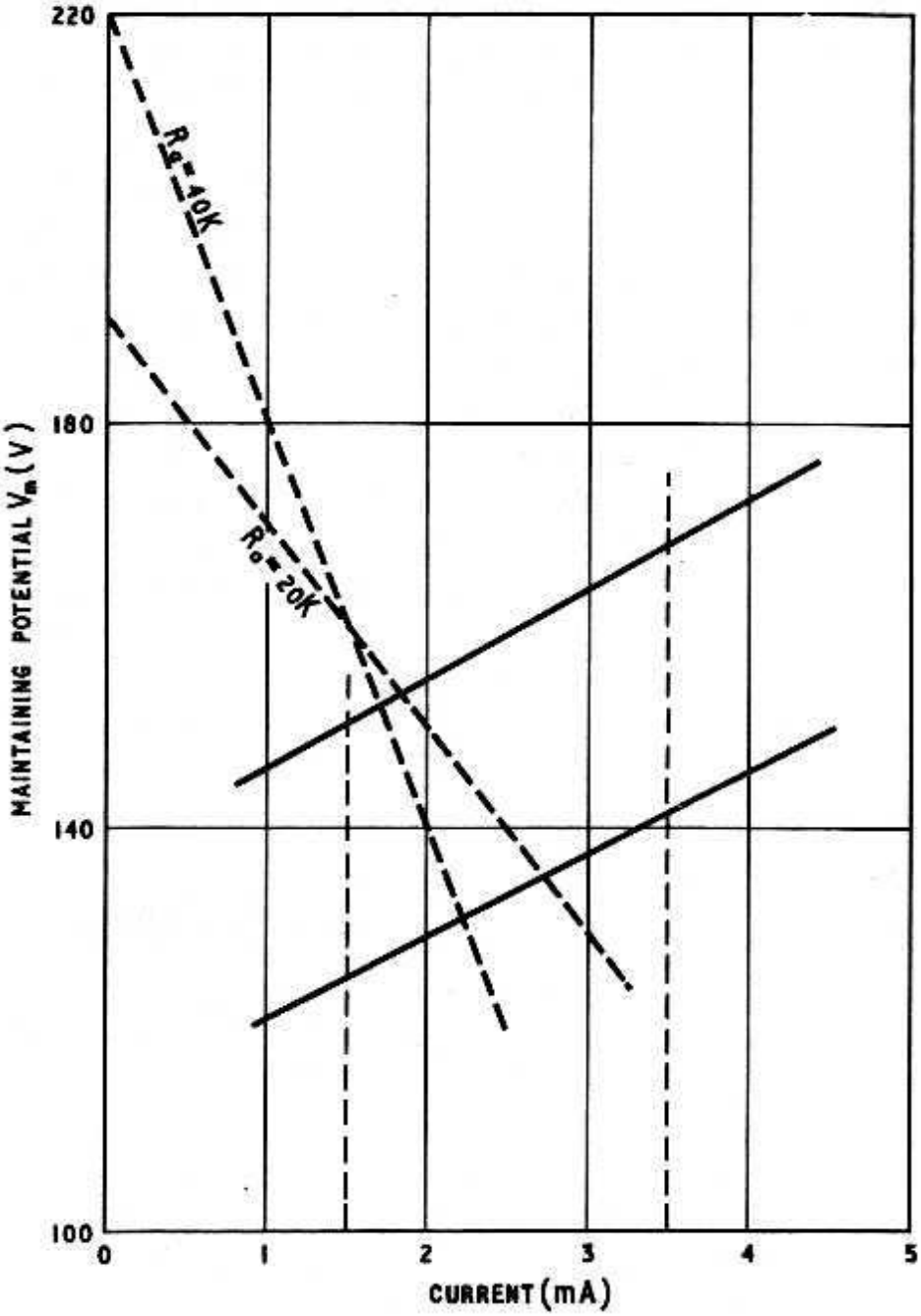


Fig. 9.16. Maintaining potential/current diagram, illustrating the graphical method of determining the anode supply voltage and series resistor for the ZM1080

coverage region, and also will vary from tube to tube and during life, so that V_B and R_a should be chosen to contain the current within the limits for all possible values of V_m. One method of

arriving at suitable values is graphically from the V-i characteristics, an example of which is shown in Fig. 9.16 for the ZM1080 tube. The two full lines represent the upper and lower manufacturing limits of maintaining potential, and the current limits are given by the vertical dotted lines. Thus the working area of the tube is the area within the four lines. A load line can be constructed on Fig. 9.16 corresponding to R_a, which clearly must intersect the maintaining voltage limit lines between the current limit lines. A series of such lines can be drawn to determine the permissible tolerances on V_B and R_a; two lines are shown in Fig. 9.16. V_B must of course be greater than the maximum breakdown potential of the tube type, which for the ZM1080 is 170 V. However, since it is better to run the tube near the minimum current (2 mA is recommended for the ZM1080) a higher value of V_B is advantageous allowing for wider tolerances on R_a. Alternatively, the design procedure for stabilisers (Section 6.2.1) can be employed to determine V_B and R_a, if the load current is taken as zero.

The other parameter important to the circuit design is the value of the potential difference which must be applied between the selected cathode and the unwanted or 'off' cathodes in the normal type circuit of Fig. 9.13(b) to give an acceptable indication. In Fig. 9.17 the current-sharing of the 'off' cathodes, which dictates

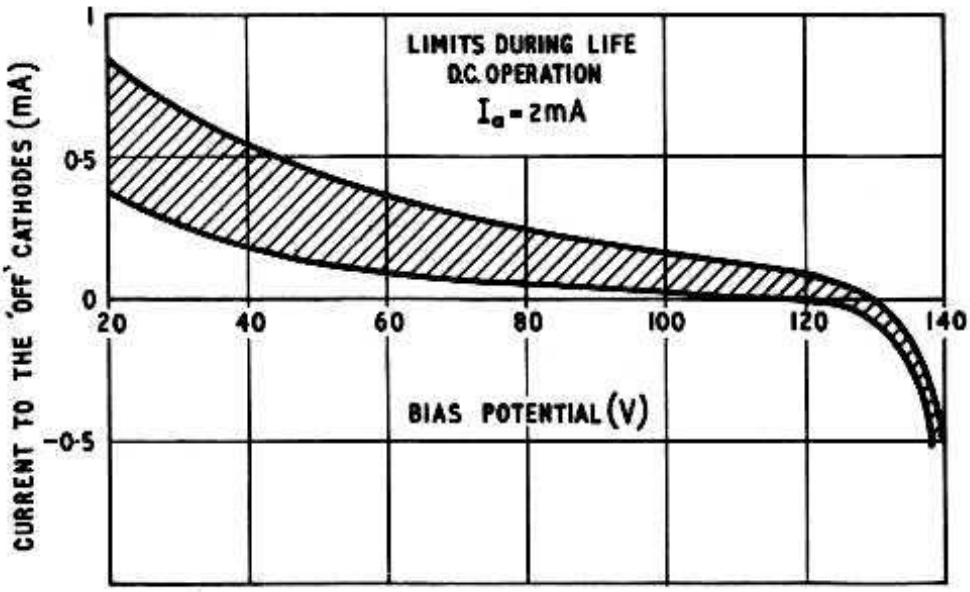


Fig. 9.17. Total current to the 'off' cathodes in the numerical indicator ZM1020 as a function of the bias potential. Total tube current = 2 mA d.c.

the clarity of the indication, is illustrated by a plot of the total current to the 'off' cathodes as a function of the said potential difference (expressed here as the positive bias applied to the 'off' cathodes, the selected cathode being at earth) for the ZM1020. The spread of the curves takes account of the variations that occur

when different numerals are selected and for variations on life. A satisfactory display may be obtained with an 'off' cathodes' current of up to 20% of the total current, i.e. with a bias of ~ 60 V according to the example of Fig. 9.17. A reduction of background 'haze' and hence an improvement in contrast may be obtained by increasing the bias, say up to 80 V.

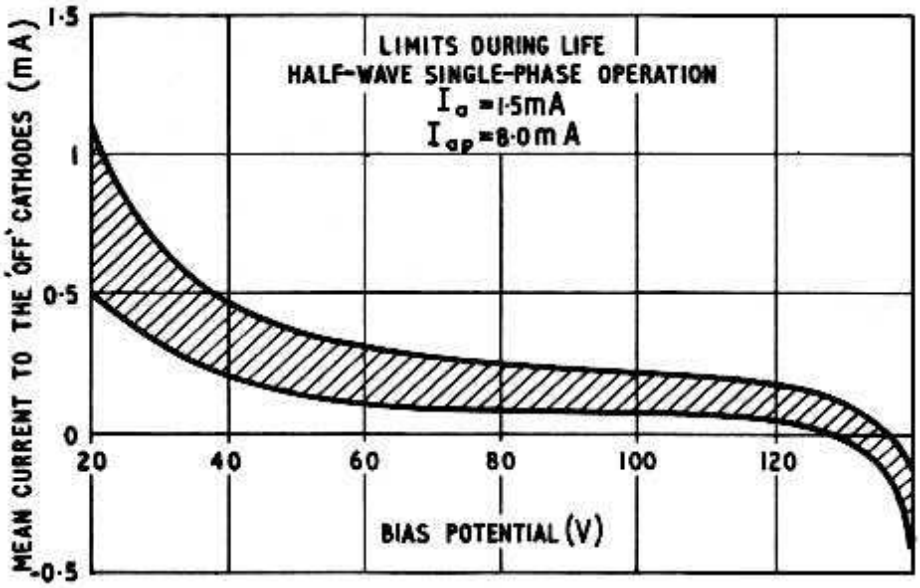


Fig. 9.18. Total current to the 'off' cathodes in the numerical indicator ZM1020 as a function of the bias potential. Total tube current = 1·5 mA half-wave rectified, peak current = 8 mA

In a number of circuit forms, particularly those associated with semiconductors, this relatively large bias potential is an embarrassment. An improvement, however, can be obtained if, instead of using a steady d.c. supply to the tube, a pulsed supply is used giving a higher peak current but still maintaining the average at about 2 mA. This increases the apparent contrast of the indication for lower values of bias potential. The brightness is unaffected since this is proportional to the mean current. A particularly convenient method of obtaining the pulse voltage is to use 50 c/s a.c., full or half-wave rectified, but without any smoothing. The current-sharing characteristic for the same tube type as used for Fig. 9.17 but for a.c. is shown in Fig. 9.18. With this circuit a bias of 40 V gives a satisfactory indication. Although the sputtering rate can be expected to be higher at the large peak currents, the cathodes are always covered even if slightly contaminated, and no deterioration in life performance is experienced. Design procedure for rectified a.c. operation has been outlined by Neale.[10]

As with any multi-cathode tube, the useful life might be expected to depend on the length of time that any one cathode is used, and therefore life could be expressed in terms of hours per cathode.

However, this is not so if the cathodes are being sequentially changed, because one of the major factors determining the life is the contamination of adjacent unused cathodes, and the small amount formed when the numbers are continually being switched is rapidly removed when current flows to the cathode in question. In the case of the numerical indicator tube, lives of over 30,000 hr can be expected if the numbers are sequentially changed, but this is reduced to 5,000 hr if the discharge is invested continuously on only one number. A point that is not always appreciated is the fact that end-of-life failure is rarely catastrophic as in the case of displays using incandescent lamps, but is more in the nature of deterioration of the display. Even when this deterioration becomes apparent, it may be possible to continue using the equipment in which the tube is installed until it is convenient to remove for servicing perhaps some 1,000 hr of operation later.

9.4.5. DRIVE CIRCUITS FOR THE NUMERICAL INDICATOR TUBES

The numerical indicator tube has a wide number of applications, such as digital frequency meters and voltmeters, floor indicators in lifts, small computers and desk calculators. The nature of the equipment in which they are installed—the type of components used and the functioning speed—dictate the manner in which the tubes will be operated. In the following sections the various methods of driving the tube are described to assist the design engineer in his specific application.

The simplest method of driving the tube is by mechanical switches such as uniselectors and relays, and although operation is limited to very slow speeds the method can be used with advantage in some process control applications. Rather higher speeds can be obtained by driving the tube from cold-cathode tubes or photoconductors, but for much higher speeds, thermionic valves, or, better still, transistors are mainly used. As the numerical indicator tube is normally a scale-of-ten read-out device it is most conveniently used in conjunction with decade counters. The examples of driving circuits that follow, therefore, are mainly based on suitable inter-connection between decade counting circuits and the indicator tube.

9.4.5.1. *Trigger-Tube Ring Counter Drive*

The numerical indicator tube can be used in conjunction with a trigger-tube ring counter of the type described in Section 7.10.2,

by directly connecting each cathode of the numerical tube to the appropriate anode of the trigger tubes in the counter. A circuit to run at a few hundred c/s has been described by Higgins,[11] in which the directional mechanism in the counter is provided by capacitor coupling the trigger-tube anodes. Higher speeds can be obtained from the more conventional ring counter in which the following trigger tube in the ring is pre-biased from the cathode load of the conducting trigger tube.[12] A suitable circuit based on this method is shown in Fig. 9.19 for driving the ZM1020 from a

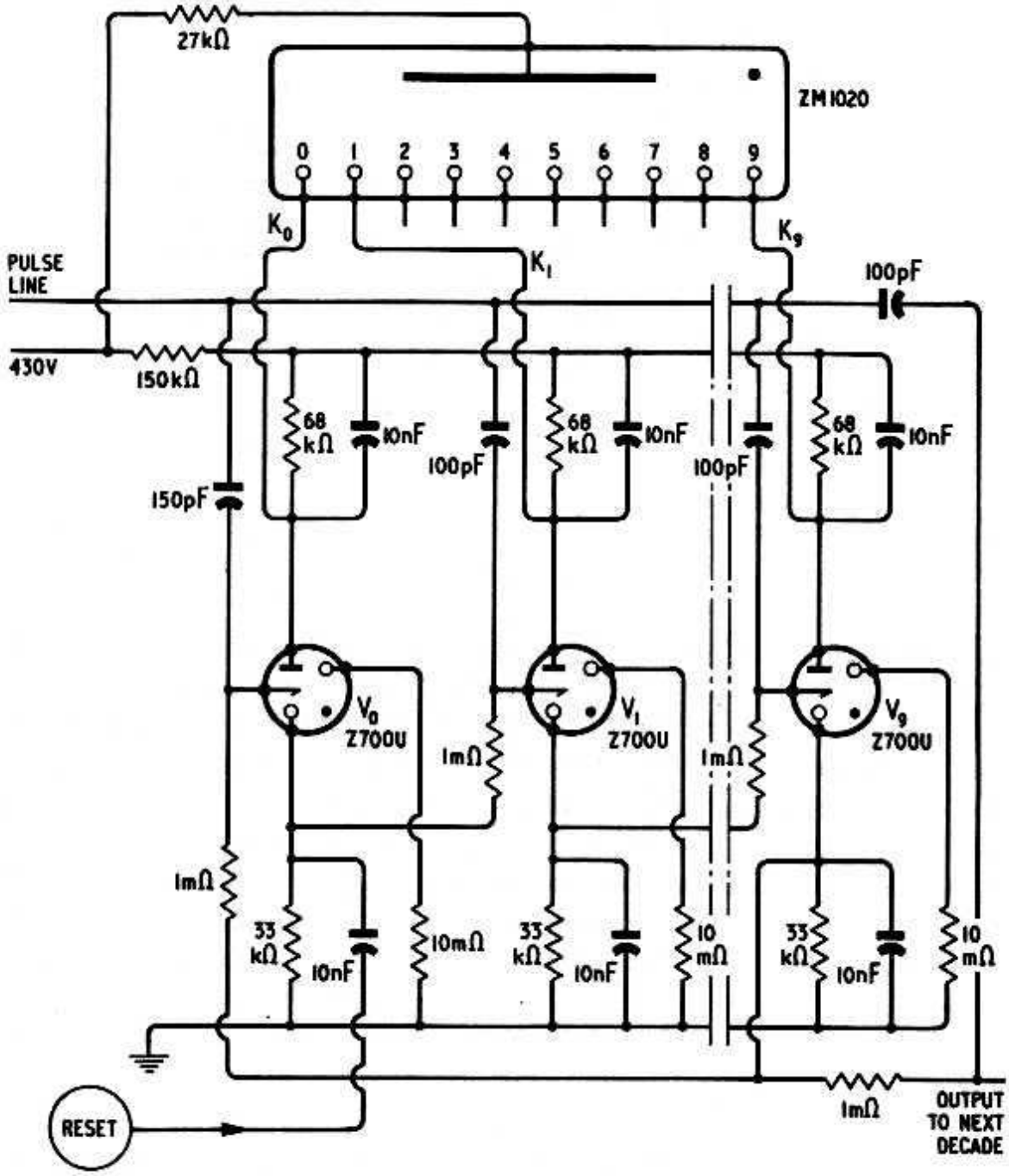


Fig. 9.19. Trigger-tube ring counter directly coupled to the numerical indicator tube

Z700U ring counter.[13] Each anode of the trigger tubes has a 68KΩ load resistor connected to it in series with the common 150KΩ load. The potential drop across this 68KΩ, which occurs when the trigger tube conducts, provides sufficient negative bias to the connected cathode of the indicator tube to ensure that the correct numeral

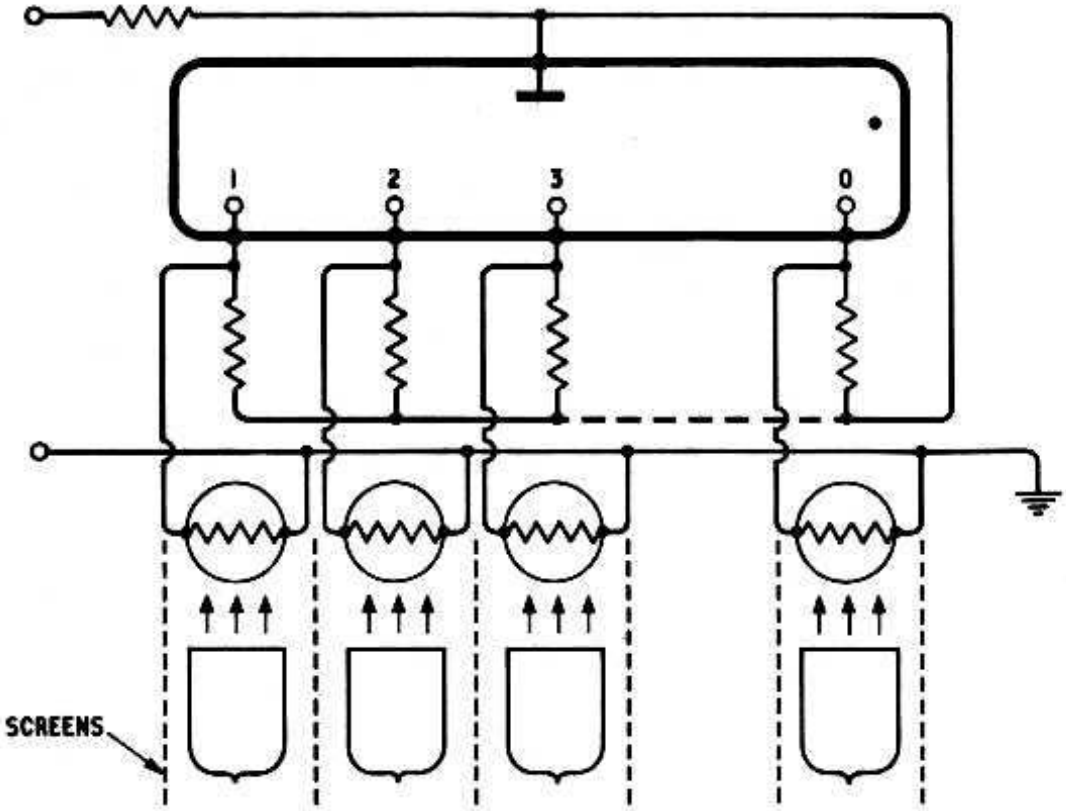


Fig. 9.20. Trigger tube ring counter coupled by photo-conductive cells to the numerical indicator tube

and only that numeral is displayed. The shunt capacitor allows the speed of the counting chain to be less dependent on the ionisation and deionisation time of the ZM1020. Nevertheless, the addition of the numerical indicator tube does restrict the counting rate to about 1 kc/s. If the full counting rate of the ring counter is to be utilised then the counting function must be separated from the read-out mechanism. This can be achieved by making use of the light output from the trigger tubes to activate cadmium sulphide cells which are connected at the cathodes of the numerical indicator tubes;[14] the circuit is illustrated in Fig. 9.20.

9.4.5.2. *Stepping Tube Drive*

Driving the numerical indicator tube from a stepping tube (selector) is obviously an attractive proposition. Unfortunately the output from a standard stepping tube is insufficient to drive directly the indicator tube, and some form of amplification is required for each

digit. The amplifying circuits, however, can be relatively simple and economic in components.

A circuit using thermionic triodes is shown in Fig. 9.21. The triodes are normally biased to cut-off, and the positive input from the stepping tube switches on the appropriate triode so that the numerical tube indicates the glow position of the stepping tube. The circuit is mainly of academic interest, since cold-cathode tubes and transistors offer a more economic solution in terms of space, power and price. The triodes, however, give a large output which is capable of driving the bigger numerical indicator tubes or several tubes in parallel.

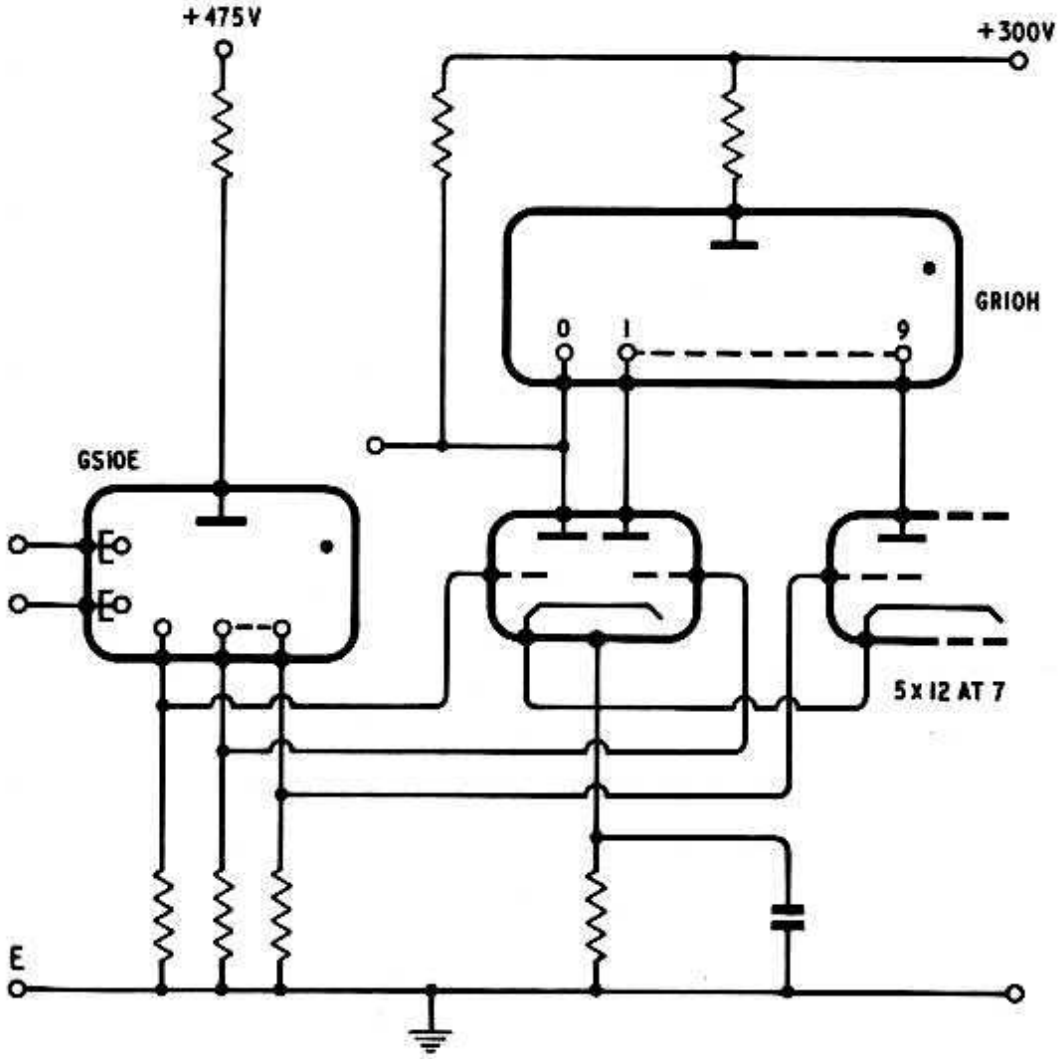


Fig. 9.21. Thermionic valve drive of the numerical indicator tube from a selector stepping tube

The triodes can be replaced by trigger tubes if the supply to the trigger tubes is pulsed so that they are extinguished at regular intervals. A suitable supply is obtained from unsmoothed rectified a.c., which can also be applied to the numerical indicator tube connected in series. A circuit showing this arrangement is given in

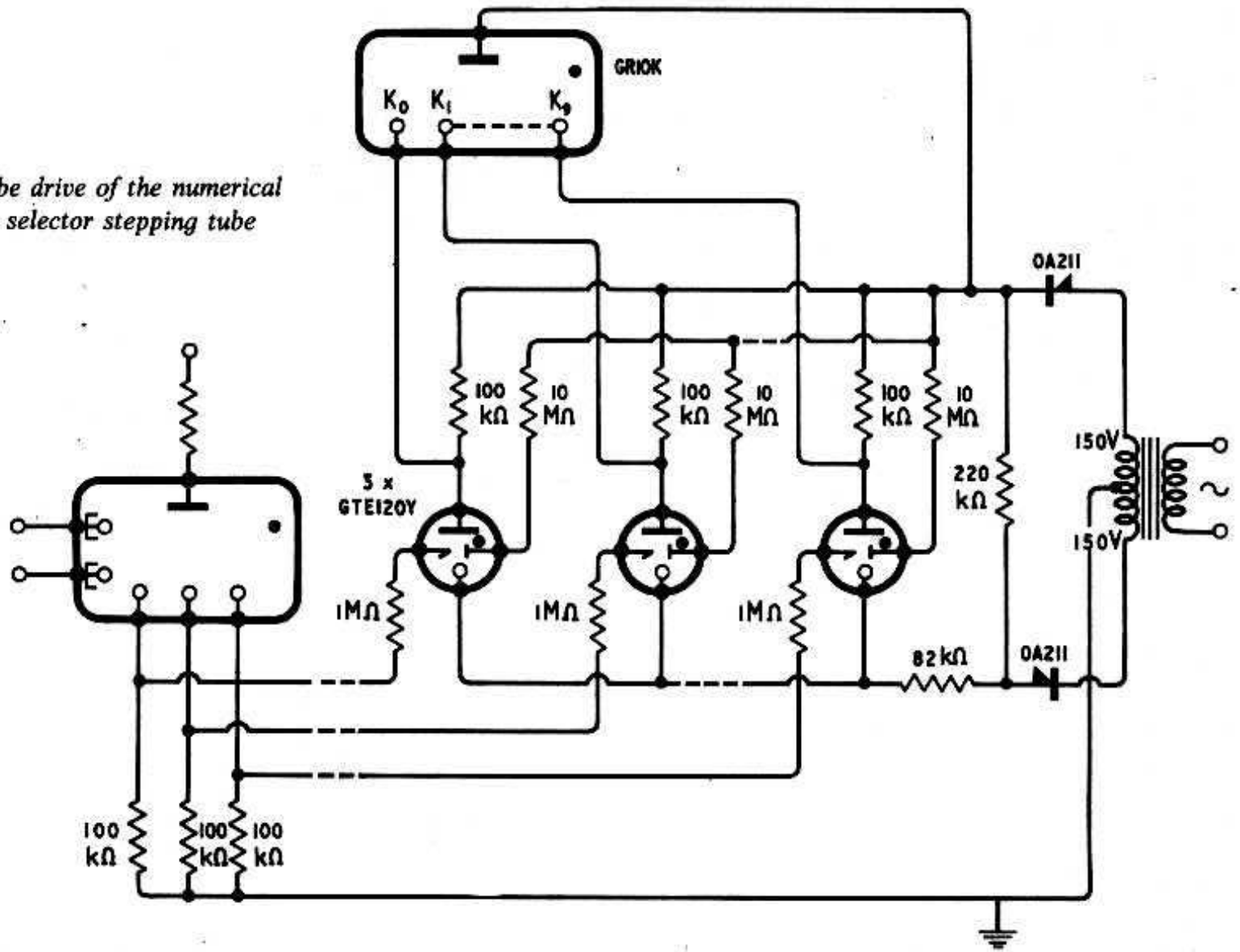


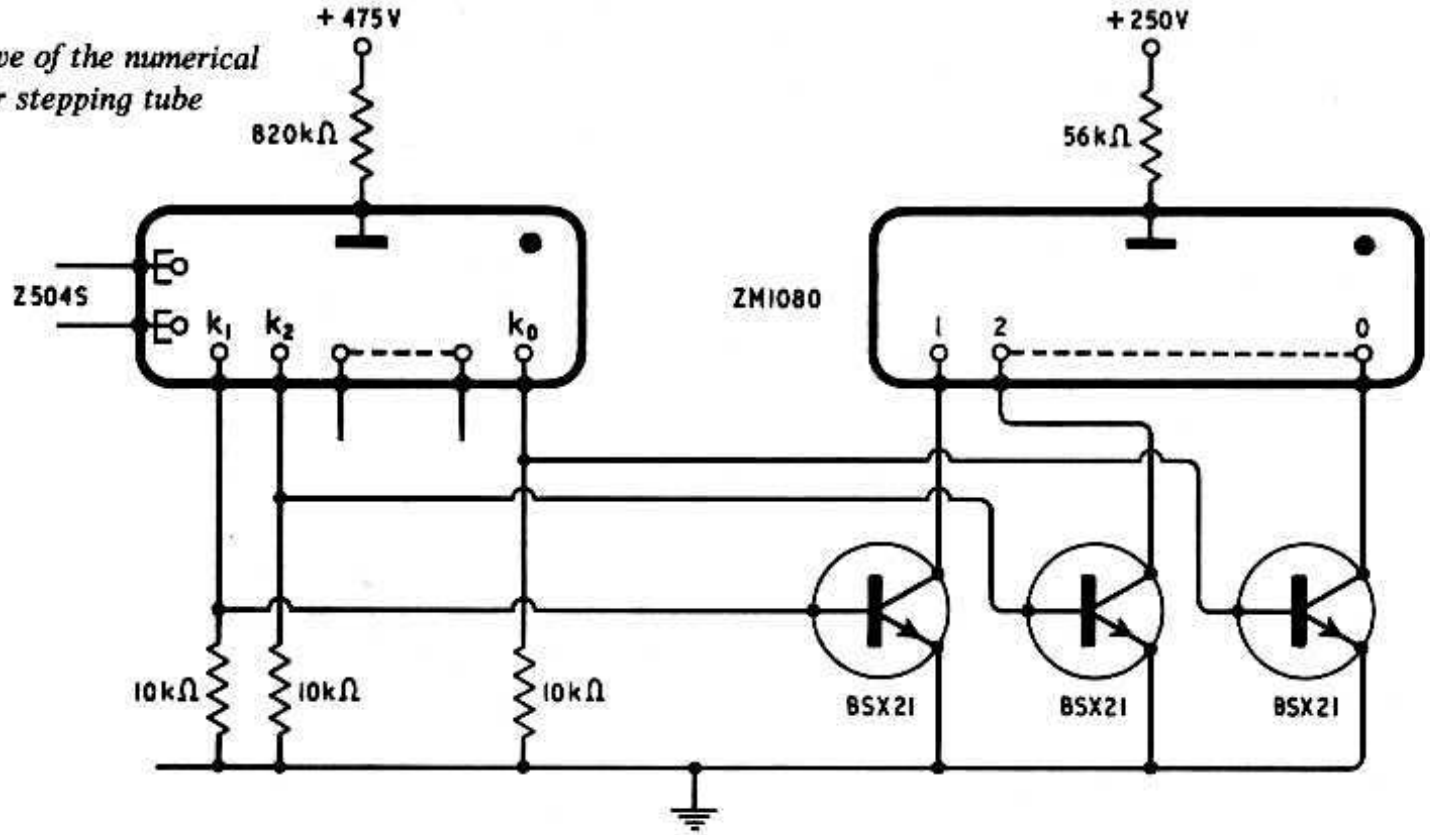
Fig. 9.22. Trigger tube drive of the numerical indicator tube from a selector stepping tube

Fig. 9.22. When the glow is resting on a particular cathode of the stepping tube the positive bias developed across the cathode resistor is applied to the trigger of the associated trigger tube. Thus, when the voltage increases across the anode–cathode gap of the trigger tubes on the conducting half-cycle, and the cathodes go negative with respect to the triggers, the biased trigger tube will ignite first. As soon as this tube conducts, the cathodes of all the trigger tubes will rise because of the 82 KΩ and ignition of any further trigger tube is prevented. The voltage drop across the anode resistor (100 KΩ) when the rectified voltage further increases will strike the corresponding cathode in the numerical indicator. At the end of the half-cycle both the trigger tube and numerical indicator will be extinguished, and the process is repeated at the beginning of each new cycle of the a.c. supply. Thus, in the indicator tube, the number will be displayed which corresponds to the position of the glow in the stepping tube.

Transistor coupling circuits using high voltage breakdown *npn* transistors now offer a simpler solution than trigger tube circuits. The early high voltage transistors were too expensive to be used for the purpose, but methods were developed to improve the breakdown characteristics of the cheap lower voltage transistors to give acceptable displays. In particular, Somlyody[15] demonstrated that the normal collector/emitter breakdown of such a transistor could be increased from 25 V to hold off 40–60 V by applying a reverse bias to the base-emitter junction. However, cheap mesa transistors are now available having breakdown potentials of the order of 120 V. Further, the leakage current in the 'off' condition is very low and produces negligible background haze in the numerical indicator tube. A suitable circuit connecting the Z504S to the ZM1080 is given in Fig. 9.23. Providing the current to the 'off' cathodes in the Z504S is small, and the temperature of the transistors does not exceed the recommended average, the 10 KΩ resistor can normally be dispensed with.

Either approach, transistor or trigger tube, provides an inexpensive method of coupling numerical output to the stepping tube where a few numerical indicators are required. In applications where a large number of numerical tubes are needed the cost of the transistors can be important. A dynamic driving technique which reduces the overall cost of components can be used in these applications. The principles of the system described by Jeynes[16] are illustrated by the block diagram of Fig. 9.24. The cathodes of all the indicator tubes are arranged so that all like numbers are connected to the same switch (i.e. all the 0's to switch S_0 all the 1's to S_1, etc.). An additional switch, driven from the output of the

Fig. 9.23. Npn *transistor drive of the numerical indicator tube from a selector stepping tube*



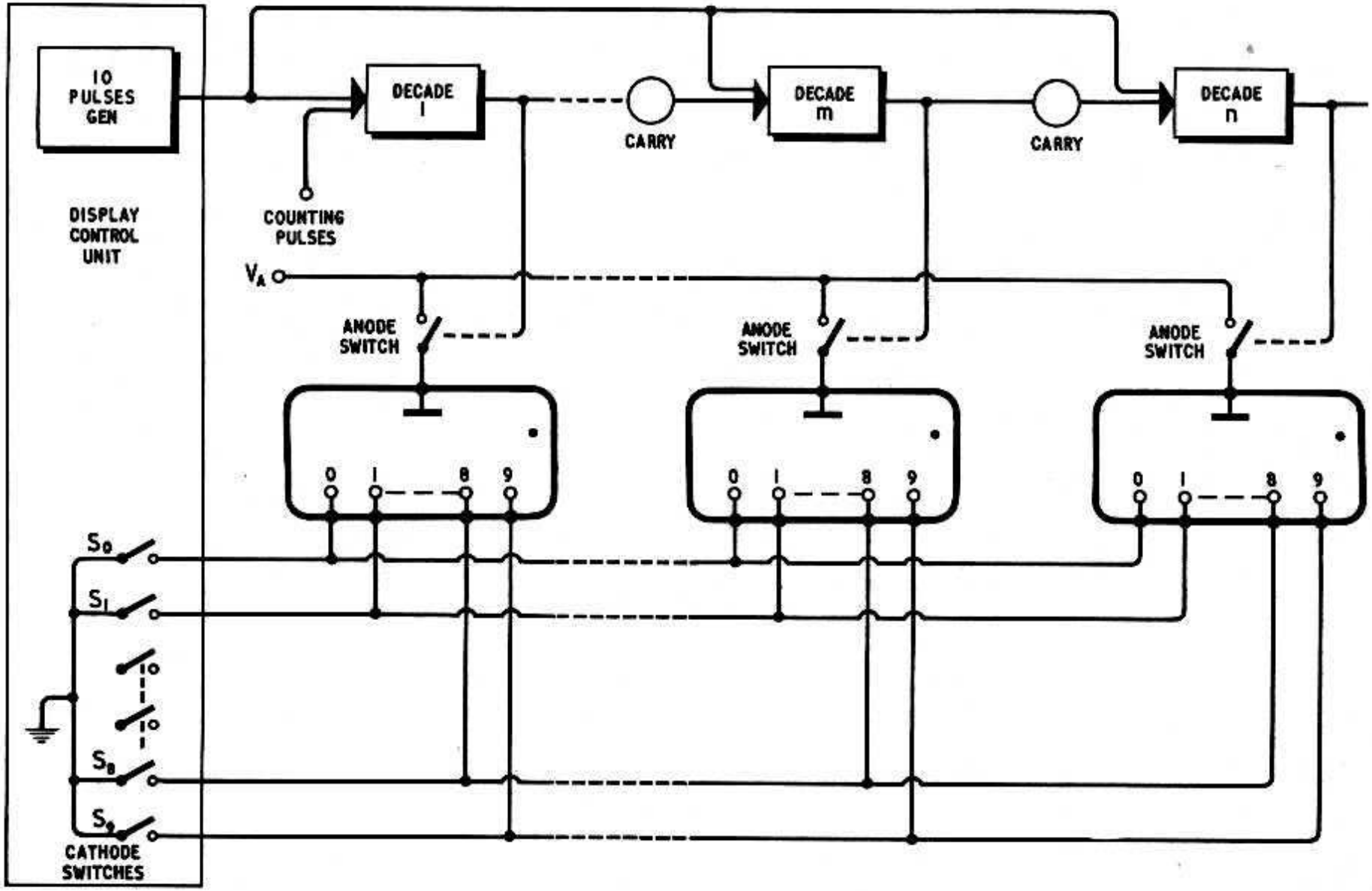


Fig. 9.24. Block diagram of the dynamic drive system of coupling numerical indicator tubes to a stepping tube scaler

counter to be displayed, is inserted in series with the anode of each tube. During counting, the cathode switches S_0 to S_9 are open, and the carry pulses in the counting chain are generated as usual, i.e. when the discharge steps off K_9 a pulse is delivered to the next decade. However, when the count is finished and read-out is required, the carry pulse is inhibited, and each tube in the counting chain is fed synchronously by a train of 10 inspection pulses. For the duration of the inspection pulses the carry pulse is used to operate the anode switch of the numerical indicator. The ten inspection pulses are also used to close the cathode switches one at a time, the switches closing in complementary order i.e. S_9 on the first pulse, S_8 on the second and so on. When the switches for both anode and a particular cathode are closed for any indicator tube, then the number will be displayed which will correspond to the number stored in the driving stepping tube of the counter. For example, if at the end of the count cathode K_7 of the first decade tube was glowing, then the discharge will transfer off K_9 on the third inspection pulse which is when S_7 is closed. Since the carry pulse closes the anode switch the number 7 will be displayed on the first numerical indicator tube. By repeating the inspection pulses at sufficiently high frequency an apparently static display in digital form is observed; the pulse current is chosen to give sufficient mean brightness. The dynamic display requires only $10+n$ switches for n decades which may be either transistors or trigger tubes, compared with $10 \times n$ for more conventional systems.

9.4.5.3. *Direct Drive by a Stepping Tube with Auxiliary Anodes*

The auxiliary-anode stepping tube GCA10G, mentioned in Section 8.1.2, was specifically designed to drive a numerical indicator tube without the need for intermediate amplification.[17] The basic circuit arrangement for a single decade is shown in Fig. 9.25. In the stepping tube a voltage difference of about 40 V can exist between the conducting auxiliary anode and the adjacent ones before the latter take appreciable current. This voltage difference provides the necessary bias to the numerical indicator, and the values of V_B and the resistors R_1 and R_2 are chosen, therefore, to give this 40 V. The potential of the auxiliary anode in the conducting gap is fixed at $V_m + i_K R_K$, where V_m is the maintaining potential of the gap and i_K the cathode current. Thus for the correct bias:

$$\left(V_B - V_m - i_K R_K\right)\left(\frac{R_1}{R_1+R_2}\right) = 40\ \text{V}$$

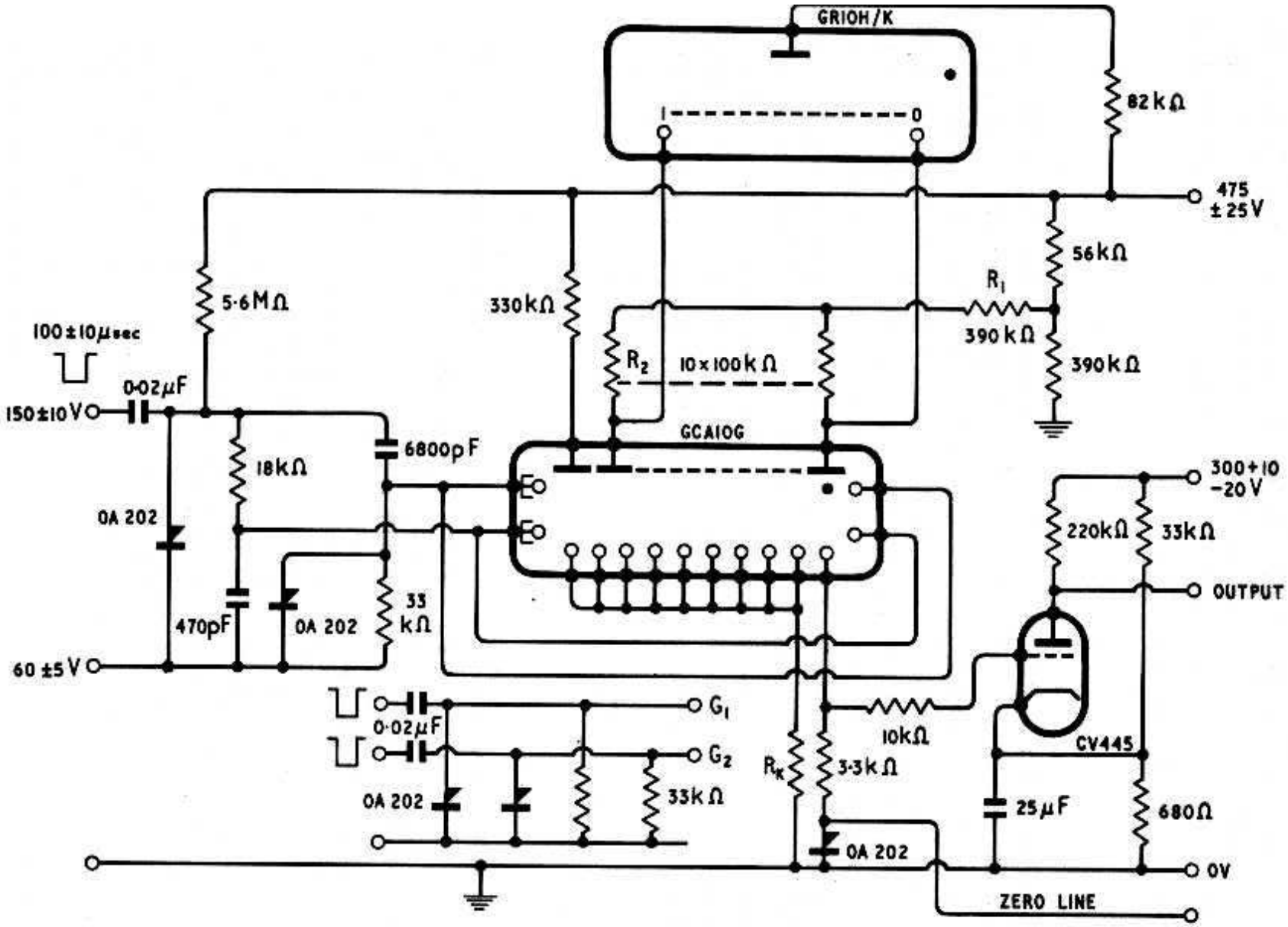


Fig. 9.25. Direct drive of the numerical indicator tube by an auxiliary-anode stepping tube

Since the output voltage of the auxiliary anodes is constant at $V_m + i_K R_K$, the current through the numerical indicator tube must be controlled by its anode resistor. The main disadvantage of the auxiliary-anode stepping tube is the close-toleranced resistors and well-regulated supplies which are required to keep the tube within its operational ratings.

9.4.5.4. *Trochotron Drive*

The circuits so far described are limited to maximum counting speeds of a few kc/s. Speeds of up to 2 Mc/s can be achieved by using beam-switching counting tubes of the trochotron type.[18, 19] The trochotron is a high-speed vacuum counting tube, and consists of a centrally located thermionic cathode surrounded by thirty other electrodes formed into ten similar groups of three electrodes. Each group comprises a target (or output) electrode, a spade (or beam locking) electrode and a grid (or switching) electrode. The whole tube is mounted within a cylindrical magnet which gives the electrons a trochoidal path. An electron beam can be directed from the cathode to any one of ten stable positions, and is switched

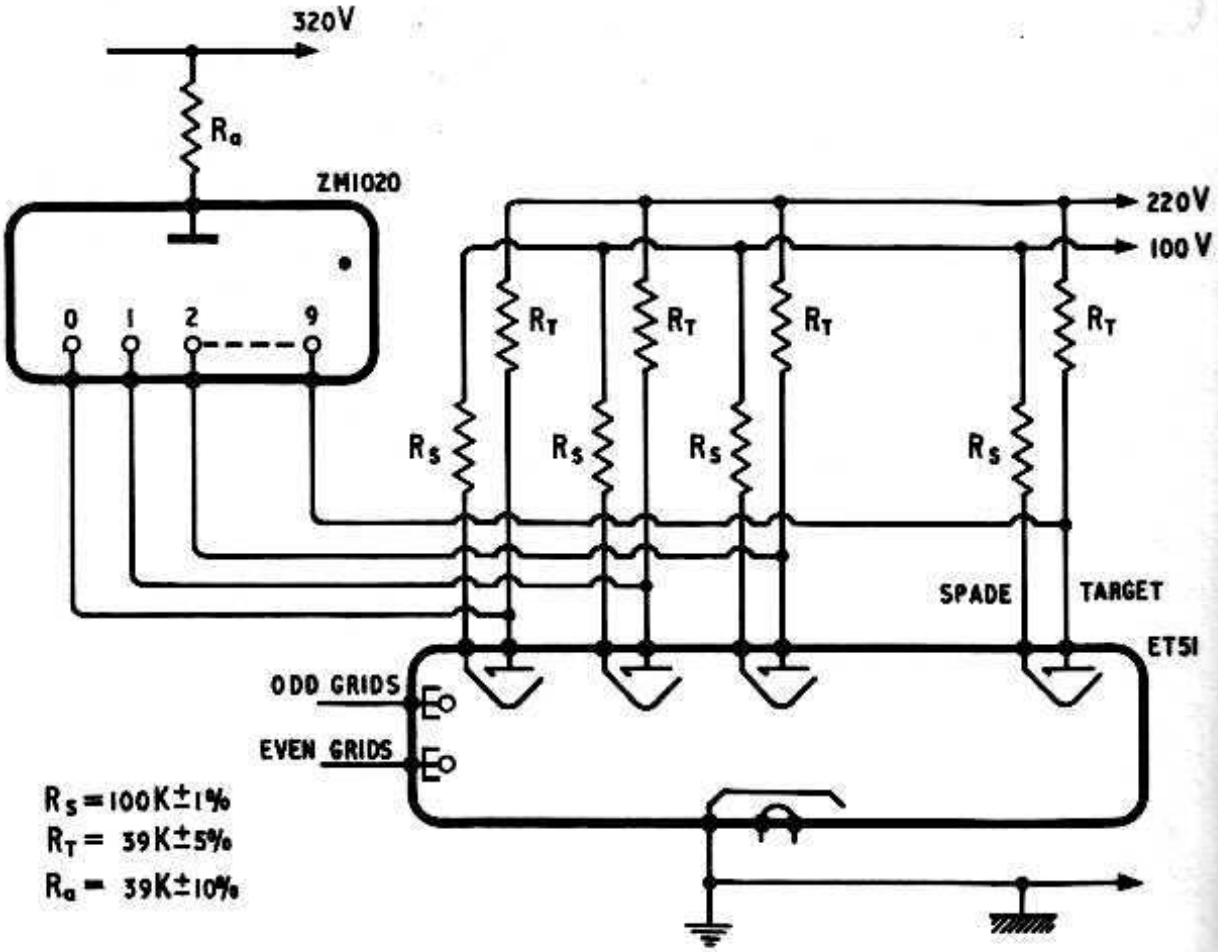


Fig. 9.26. Basic circuit for driving the numerical indicator tube directly from a trochotron

sequentially round the tube by a suitable signal on the grid electrodes, the direction being governed by the magnetic field. All the ten targets and ten spade electrodes are brought out individually through the valve base. Alternate grid electrodes are connected together so as to form a set of 'odd number' grids and a set of 'even number' grids; the drive pulses are fed alternately to odd and even grids to obtain the required sequential stepping. The output of the trochotron is essentially a constant current, the magnitude being controlled by the spade voltage, but with a maximum of 18 mA. It is therefore particularly suited for directly driving the numerical indicator tube; the basic circuit is shown in Fig. 9.26. Because of the constant current output, the resistor in the anode lead of the indicator tube is not strictly necessary. It is, however, desirable since it reduces the dependence of the indicator current on the total current to the target, important in view of the need for a target resistor. The target resistors are provided to give the required voltage difference between conducting and non-conducting cathodes, and also to provide a current path at stepping speeds above 300 kc/s where the indicator tube will fail to ionise during the counting cycle (it will of course operate when read-out is required). In the design of such a circuit the published data of the two tubes impose stringent limits on the choice of circuit parameters, and it is desirable to employ catching diodes for fixing the conducting target voltage if extreme values of the tube data are to be taken into account.

9.4.5.5. *Drive from a Binary Counting System*

Numerical indicator tubes are not necessarily confined to driving devices or circuits operating on a scale of ten. They can also be operated from circuits based on binary-coded-decimal counting (BCD) systems provided suitable decoding circuits can be constructed. This is not a complex task, and circuits have been developed using diode logic 'AND' gates in conjunction with transistor drive for this purpose.[11, 20] A conventional diode matrix for a decoder using the BCD code 2–4–2–1 is shown in Fig. 9.27. The input to the matrix is from binary circuits (A, B, C, D) switching between 0 and +6 V, and the matrix comprises of 28 diodes. When the cathodes of all the diodes connected to one driving transistor are switched to +6 V the number is 'on' but if any one cathode of the diodes is connected to zero the number is switched 'off'. By using the resistor chain between the 12 V and −6 V lines cheap 0A200 diodes are adequate for slow speed counting and BAY38 diodes for high speeds, there being no danger of the transistors

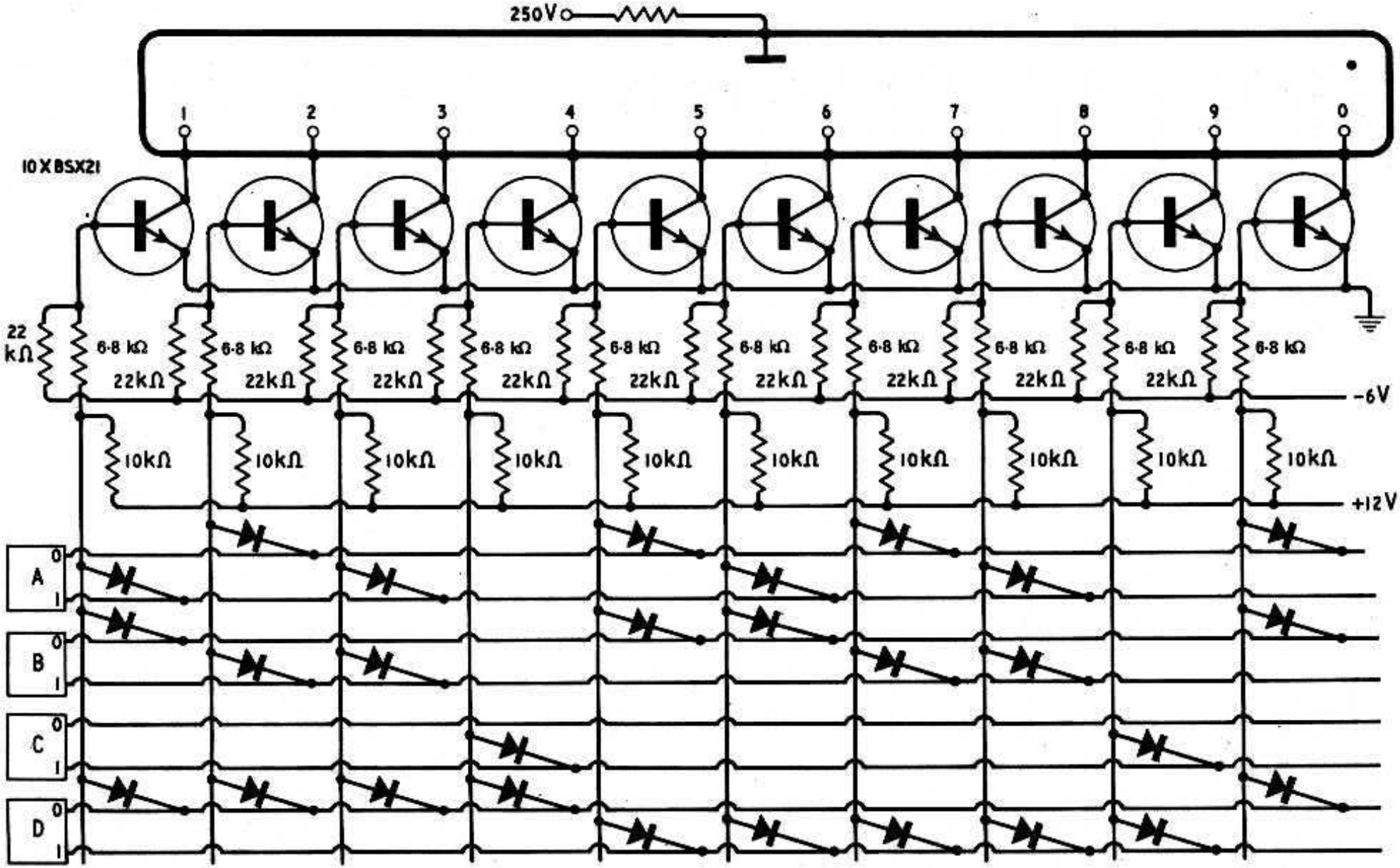


Fig. 9.27. Decimal indication from a binary scaler using a diode matrix BCD *converter, and* npn *transistor coupling to the numerical indicator tube*

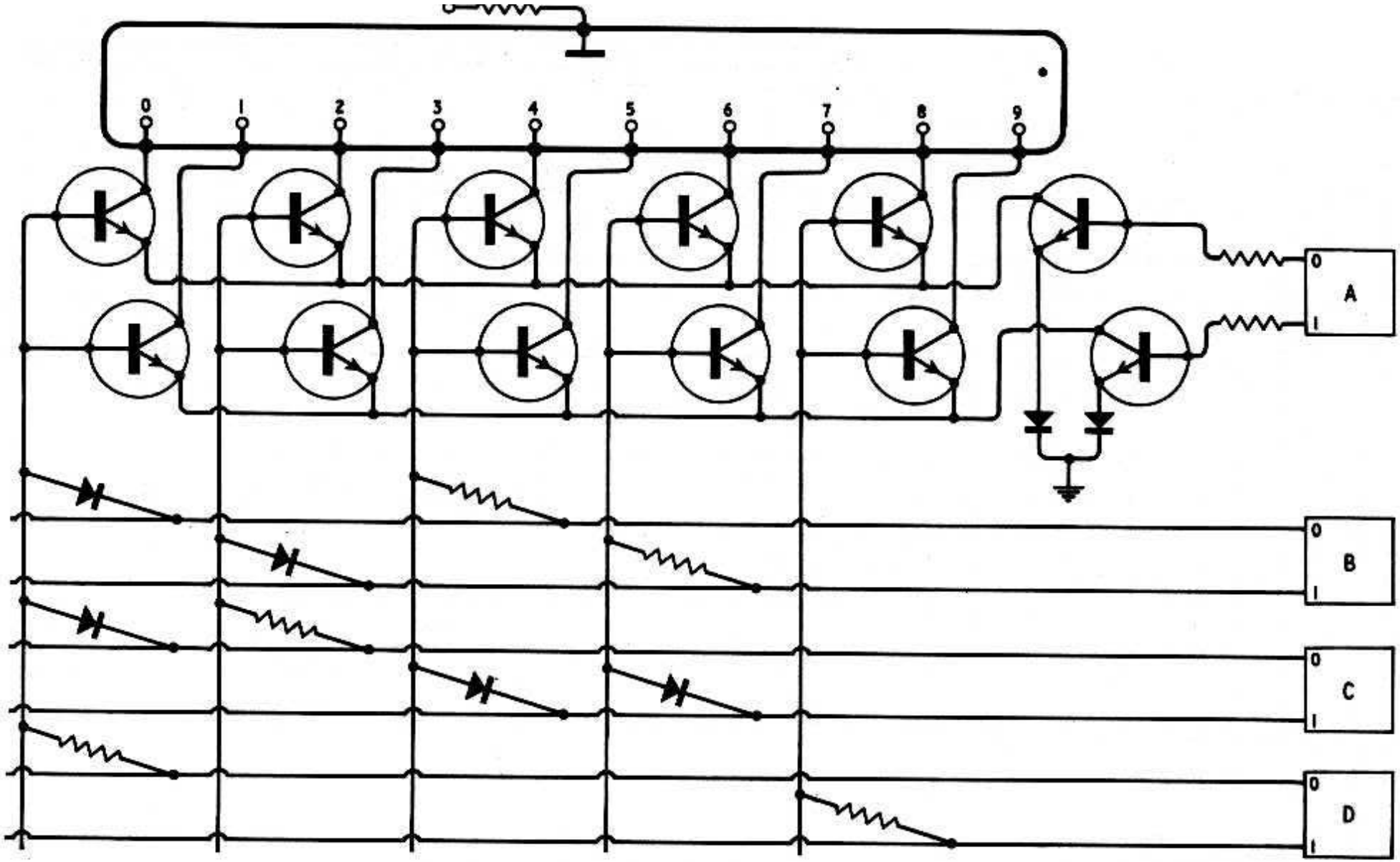
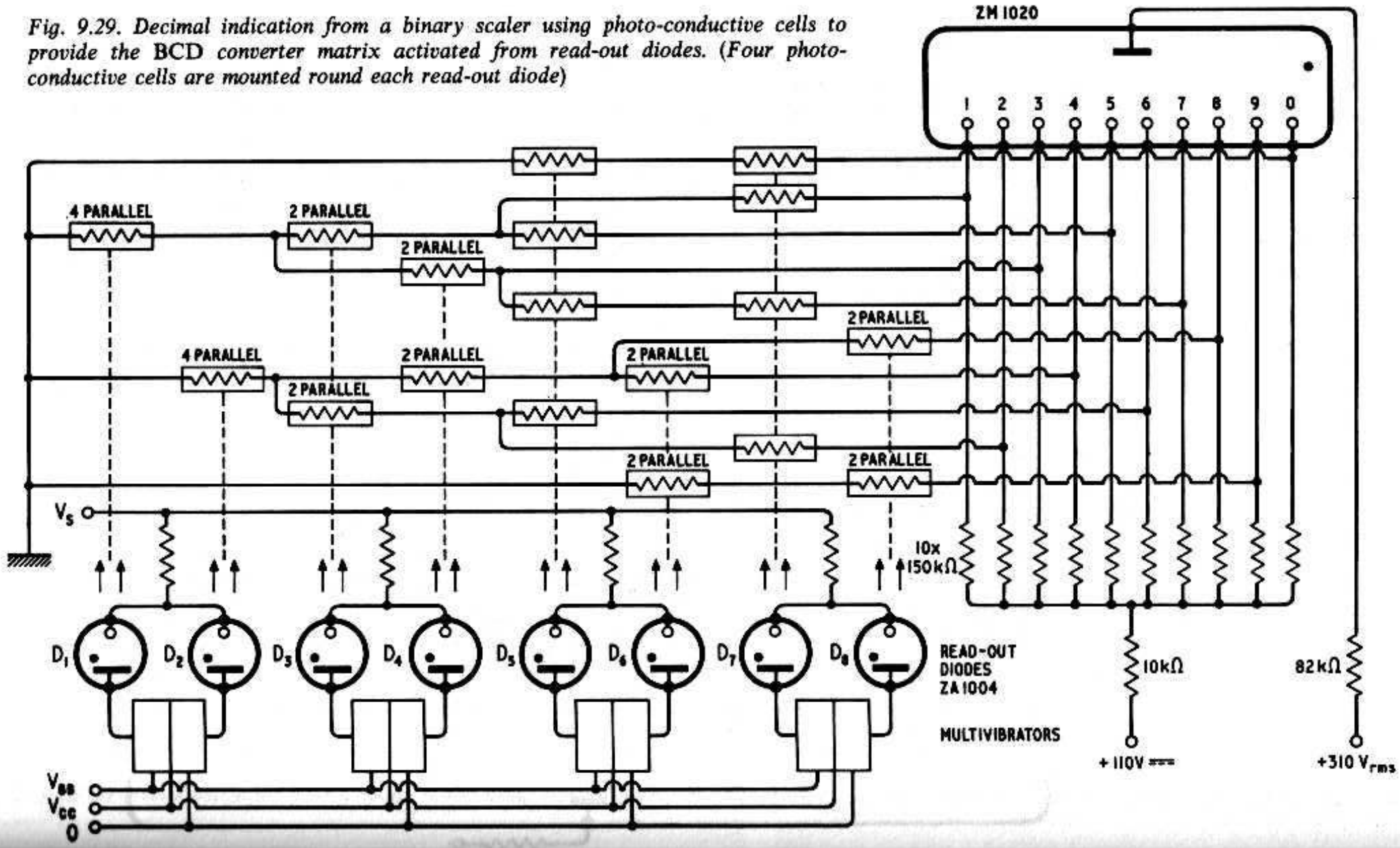


Fig. 9.28. Decimal indication from a binary scaler modified by emitter switching to reduce the number of components in the BCD *converter matrix*

Fig. 9.29. Decimal indication from a binary scaler using photo-conductive cells to provide the BCD *converter matrix activated from read-out diodes. (Four photo-conductive cells are mounted round each read-out diode)*



being switched on by the leakage current. Further, no precautions need be taken to control the temperature of the transistors.

It is seen from Fig. 9.27 that the binary circuit *D* switches 'on' either the driving transistors 0–4 or the transistors 5–9, and that in the matrix connected to *A*, *B* and *C*, 0 is identical to 5, 1 is identical to 6 etc., making in all five identical pairs. Adby[20] has pointed out that if the *D* circuit is used to switch the emitters of the driving transistors, then the bases of identical pairs can be connected together, reducing the number of diodes by 19. Further the reduction of the number of diodes connected to the base of any one input transistor from 5 to 2 makes it possible to replace some of the diodes by resistors. A circuit based on these ideas is given in Fig. 9.28, for an 8–4–2–1 binary code. It will be noted that since there is an extra transistor in the emitter lead of the driving transistor, there is no possibility of the transistor being switched on by the leakage current, and the negative line is no longer required.

An alternative approach is to provide the binary switches with read-out diodes using the circuit of Fig. 9.3 and to decode the binary number to decimal indication by interconnection of photoconductive cells mounted around the read-out diodes. A suitable arrangement described by Poorter *et al.*[21] is shown in Fig. 9.29.

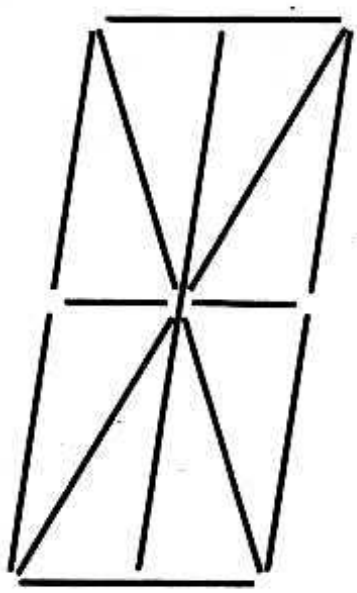
9.4.6. ALPHA-NUMERIC DISPLAY TUBES

In the last few years requirements have grown for displaying letters of the alphabet as well as numerals in such applications as airline arrival and departure boards, stock quotation information, military situation displays, etc. Previously such information displays had been limited to C.R.T. tubes or dot matrix displays, and although they have advantages in particular applications, high cost, complicated drive circuitry, and in some types poor readability, have limited their wider use.

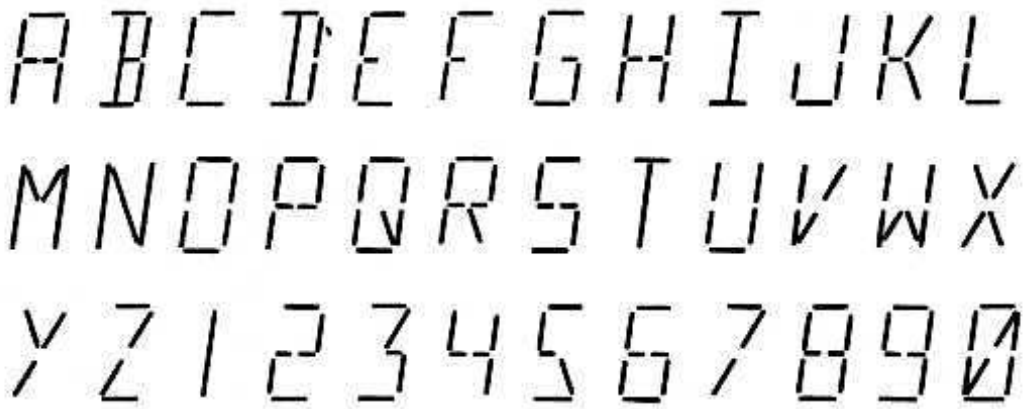
The segmented alpha-numeric display tube was developed specifically for these types of application. Within a single tube all the numbers and letters of the alphabet can be selected for display in one plane, a task which would be impossible in the stacked-cathode tube. The cathode of the tube is constructed of a number of segments mounted on an insulating support, each segment being separately connected electrically. The required letter or number is displayed by applying the necessary potential between a common anode and the appropriate combination of cathode segments. A common arrangement, which has 13 segments, is shown in Fig. 9.30 (a); inclining the matrix gives slightly better presentation.

The shape of the characters which can be formed from this arrangement is illustrated in Fig. 9.30 (b).

The manufacturing techniques used in the construction and processing of the segmented cathode tube are very similar to those used for the stacked cathode tube. Typically the cathode segments are stainless-steel strips mounted edge-on to enhance the brightness and provide uniform character height with good joining up of the glow at the corners. The anode is in the form of a mesh with high optical transmission placed in front of the cathode matrix. A similar gas filling is also used, namely 99% neon + 1% argon with a very small addition of mercury. Maloney and Glaser[8] claim that not only does the mercury give the longer life, but also it increases the dynamic impedance of the tube and permits common anode-resistor operation without causing unusual differences in brightness when characters employing different numbers of segments



(a)



(b)

Fig. 9.30. (a) 13-segment cathode array for an alpha-numeric tube
(b) construction of the letters and numbers using the array

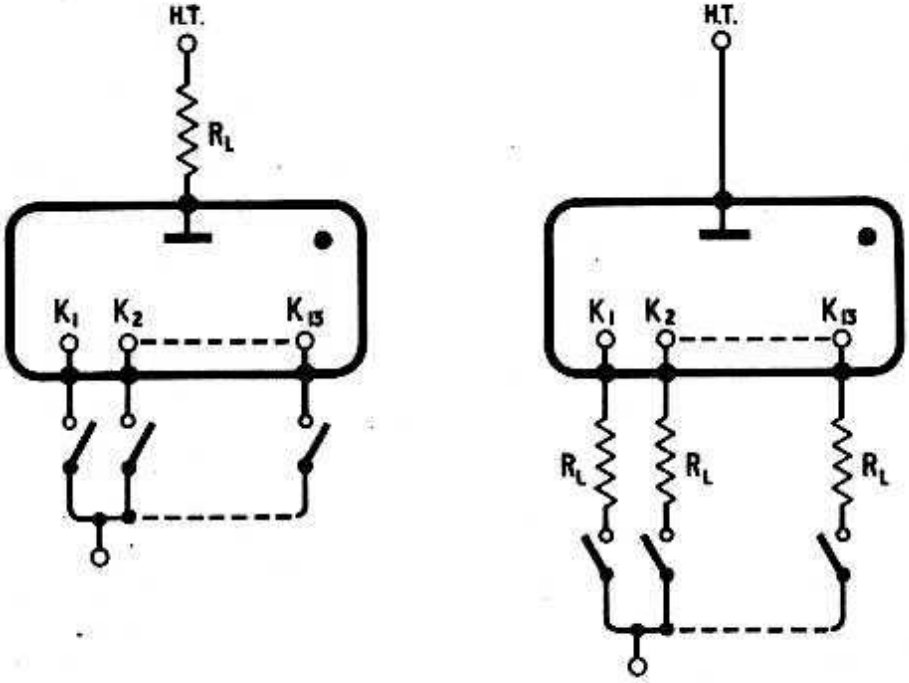


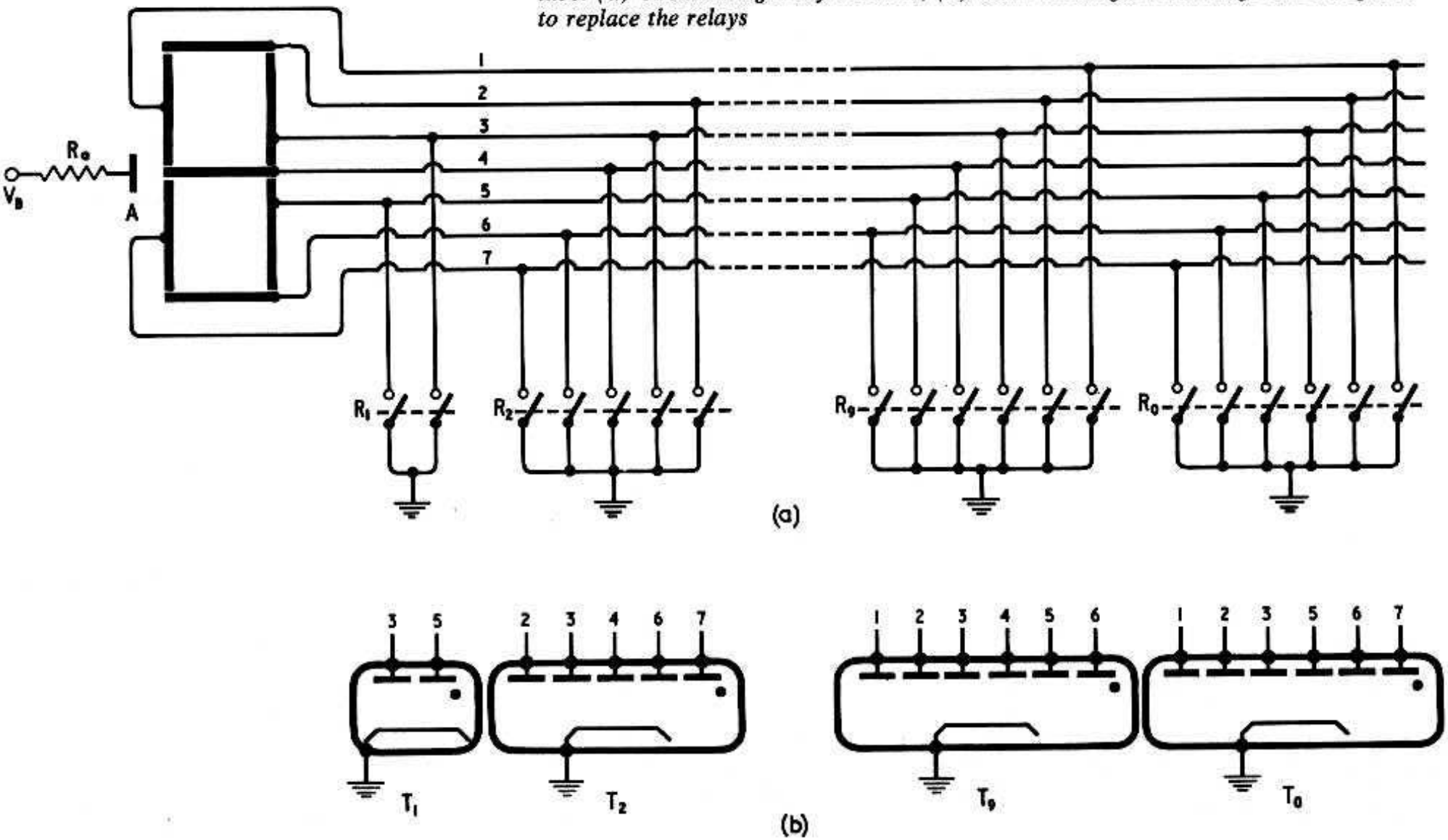
Fig. 9.31. Basic drive circuits for the segmented cathode alpha-numeric display tube

are displayed side by side. The alternative method of operating the tube is to connect the anode directly to the positive supply and place the current limiting resistance in the cathode leads. The two methods are illustrated in Fig. 9.31. The total current to the tube varies from about 3 mA to 12 mA, depending on the size of the tube. Both 'side viewing' and 'end viewing' tubes are available.

9.4.7. DRIVE FOR THE ALPHA-NUMERIC TUBE

Since to display any one character requires on average the operation of five cathode switches, the drive circuit is necessarily much more complex than for the stacked-cathode tube. Essentially, for displaying numbers and letters 36 inputs, one for each character, are required which must be coded to select 'n' out of 13 cathode switches. In practice, the information would be supplied from a 6-bit binary counter giving a coded signal rather than a separate input to 36 lines. Thus to provide character display in the tube one requires a 6-bit memory register, a 36-line decoder and an 'n' out of 13-line encoder. To provide such a system for each display tube would involve an excessive number of components and would be prohibitive in cost. A more reasonable approach is to provide a common circuit to convert the input information into the switch code for the required character, and to gate this information to the correct display tube. This entails a storage circuit connected to each display tube, but the expense involved need not be very high. The 6-bit register, decoder and encoder are best constructed using

Fig. 9.32. Basic drive circuit for displaying the numbers 0–9 with a 7-segment cathode tube. (a) circuit using relay contacts, (b) multi-anode glow-discharge tubes required to replace the relays



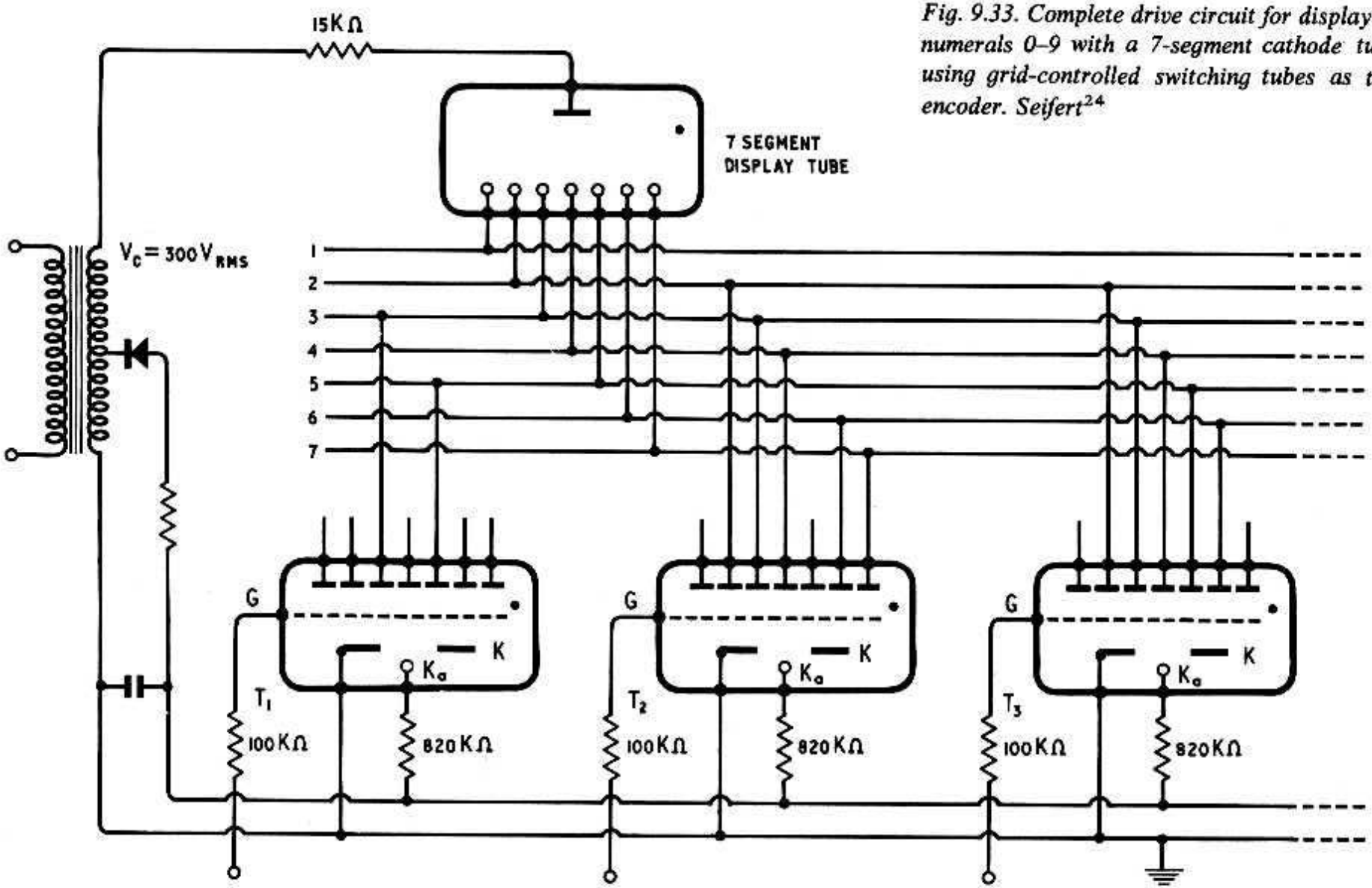


Fig. 9.33. Complete drive circuit for displaying numerals 0–9 with a 7-segment cathode tube using grid-controlled switching tubes as the encoder. Seifert[24]

transistor logic-matrix techniques or possibly using integrated circuits. The problem of designing suitable logic circuits is outside the scope of this book, there being no standard circuits for driving the segmented-cathode tube. The storage circuit should be compatible with the logic circuit, and most suitable is an npn/pnp latch, or 2 SCR's. For lower speed circuits, a latched relay or a trigger tube can be used, but they have certain disadvantages, in particular they require rather higher drive signals.

There is also the possibility of using specially constructed decoder/encoders such as have been described by Mash[22] and Greenberg[23] for use with electroluminescent panels. These, however, would be complex for displaying 36 characters. Recently, Seifert[24] has described a method of coding a segmented-cathode display tube for a limited number of characters using a cold cathode glow discharge tube, which could be incorporated within the envelope of the display tube. Since the function of cold cathode tubes may be compared with electro-mechanical relays, the method can be illustrated by first considering the equivalent circuit using relays. Fig. 9.32 (a) shows how a seven-segmented cathode matrix can be used to display the numbers 0 to 9 by connecting appropriate segments to the contacts (up to 7) of ten relays. The selected number is displayed by activating the corresponding relay; closing R_1 will display the number 1, closing R_2 displays 2, etc. The relays can be replaced by cold-cathode switching tubes having a common cathode and a number of anodes as illustrated in Fig. 9.32 (b), the character now being displayed by ignition of the correct switching tube. Although the connection of two glow discharge tubes in series presents certain problems particularly when both tubes contain several discharge gaps, a practical arrangement is possible when grid controlled switching tubes[25] are used. The circuit used by Seifert is shown in Fig. 9.33. In the switching-tube an auxiliary discharge is maintained between the auxiliary cathode K_a and the main cathode K (acting as anode) by the smoothed rectified a.c. The main cathode has a hole opposite the auxiliary cathode, through which electrons from the auxiliary discharge reach the cathode-grid space. However, the grid G acts in a similar manner to that of a thyratron, and if it is negatively biased the electrons will not penetrate into the anode–grid space and the main anode–cathode gap will not be primed. If, however, a positive signal is applied to the grid, electrons will prime the anode–cathode gap and breakdown will occur in the switching tube and in the display tube in series during the positive half cycle. The tubes can be switched by a 6 V signal on the grid. Seifert describes a construction whereby the ten switching tubes can be incorporated within a single

envelope. The construction has a common cylindrical auxiliary cathode with the ten electrode-arrays, i.e. cathode, grid and anodes, arranged in sectors radially around it. Each sector is enclosed within insulating ceramic walls. The tube can also be used in conjunction with electroluminescent panels.

9.5. Other Display Devices

9.5.1. LARGE AREA DISPLAY PANELS

The possibility of producing large area display panels using a matrix of glow discharges is actively being examined in a number of development laboratories, particularly in the United States. An example of a simple design of panel is shown diagrammatically in Fig. 9.34.[26] The display consists of a large number of individual gas discharge cells arranged in a row/column matrix. The panel is fabricated from three glass plates, the centre plate

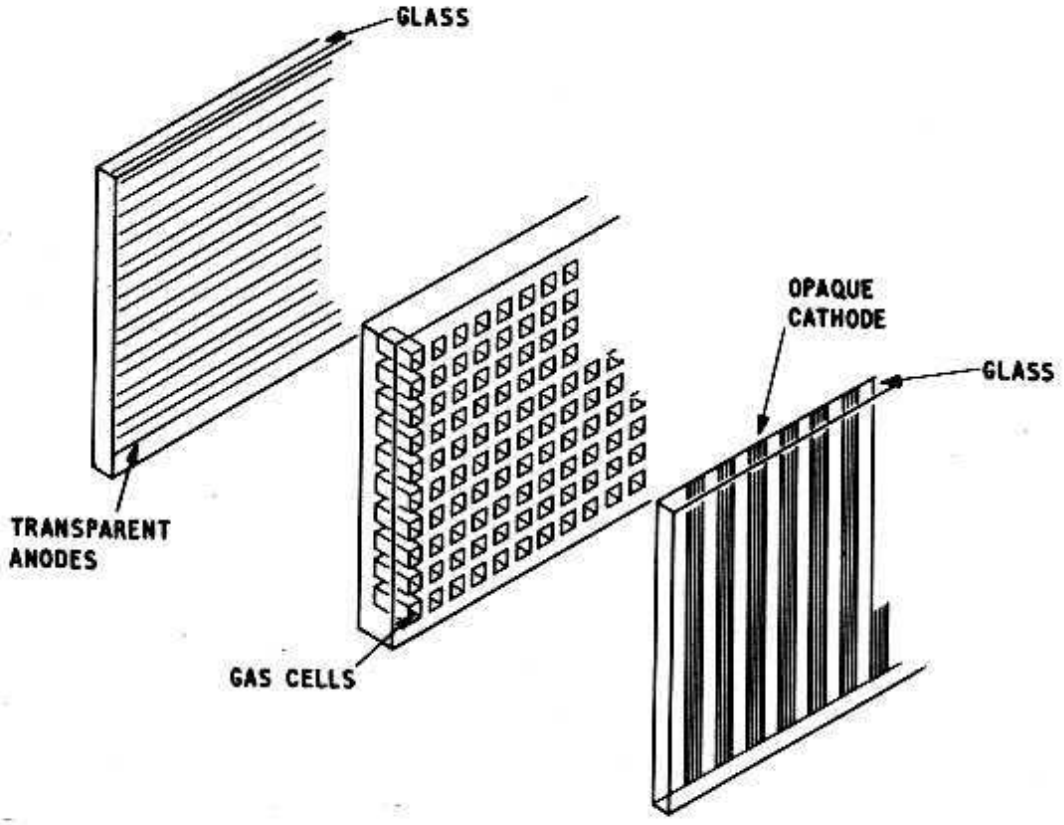


Fig. 9.34. Schematic diagram of the basic components for a glow-discharge matrix panel

forming the cells, and the outside plates sealing off the discharge areas and providing the electrodes. The holes in the central plate are photographically etched in Photoform glass which is subsequently heat treated to an opaque state. The electrodes are coated on to the end plates in strips and mounted with the anodes and cathodes at right angles to each other, lining up with the holes in

the centre plate. In this particular construction the anodes are transparent, probably tin oxide layers, but they could be fine metal strips providing the glow were not optically obscured. 4×4-in panels based on this design have been constructed $\frac{1}{4}$ in thick with a resolution of up to 20 lines per inch. The device is filled with an inert gas. The brightness of the display is around 2,500 nit, making the display visible in direct sunlight. The breakdown potential of each cell is about 285 V whilst the maintaining potential is about 200 V. By applying suitable potentials to the electrode pattern a spot of light can be made to move across the face of the panel. Thus the panel may be employed as a plotting board or for displaying tracks. For single line tracks where 're-entrants' do not occur, the trace pattern may be retained simply by applying a d.c. potential between the cathodes and anodes, which lies between the breakdown and maintaining potentials. For more complex patterns, for example the display of several tracks or of characters, a more complex switching system is required which must operate on a dynamic basis.

The panel has advantages over cathode ray tube displays in the flatness of the device, $\frac{1}{4}$ in compared with 20 in for a similar size CRT display, and also in its capability of digital operation. The size of the panel has an upper limit due to the use of glass for the structure. Unfortunately no other suitable transparent material which is stable in the presence of a glow discharge has yet been found. The applications of the device may also be limited by the writing speed which is relatively slow due to the ionisation and deionisation times of the discharge.

9.5.2. CURRENT INDICATING TUBES

In the normal glow region the area of the negative glow is proportional to the cathode current, and attempts have been made to exploit this parameter to produce a cold-cathode tube for indicating current. Usually long wire or rod cathodes are employed and the length of the glowing column is taken as a measure of current. The early tubes of this type were wide tolerance devices mainly used as tuning indicators in radio sets. Because of the difficulties of manufacturing reliable tubes with satisfactory life they have been superseded by other types of tuning indicators, and the only tube which has survived is the 4662 used in conjunction with a Penning gauge to indicate vacuum pressures in the 10^{-2}–10^{-4} torr region.

Recently the feasibility of applying modern processing techniques to improve the performance of this type of tube has been inves-

tigated by the author and Schagen[27]. The aim was to produce a tube with stable characteristics, where the glow length was linearly related to the tube current, and in which the negative glow was bright enough to be clearly visible in daylight. To obtain the linear function of glow-length to current the anode had to be mounted parallel to the cathode along its whole length, and in the initial design the electrodes consisted of molybdenum rods mounted 2–3 mm apart. The display brightness of this tube was not high enough for many applications and in a later design which allowed for greater visibility of the glow the anode consisted either of two rods placed 1 mm apart or of a metal strip with a longitudinal slit 1 mm wide with the edges bent round to give it depth. A schematic diagram of the tube, together with a photograph is given in Fig. 9.35 and Plate 9.2 respectively. Tubes up to 25 cm long and 1 cm diameter were constructed. Because of the random path taken by the sputtered atoms from the cathode at the filling pressure of 20 torr, few particles from the cathode reach the glass envelopes in front of

Fig. 9.35. Current indicator tube where the glow length is a linear function of the current (see also Plate 9.2)

the anode slit, even when the cathode is sputtered to the extent of producing an opaque layer on the rest of the tube wall (see Section 4.4.4). Thus the cathode can be thoroughly cleaned by the sputtering technique and yet the negative glow is clearly visible through the window provided by the anode slit.

To use the tube as a linear current-indicator, it is essential that the glow should commence at the base of the cathode, and that only the edge of the glow away from the base should move as the current is varied. This was achieved by welding a small spike of zirconium to the bottom end of the molybdenum rod cathode. Zirconium has a higher second Townsend coefficient γ than molybdenum, and therefore gives lower breakdown and maintaining potentials when used as a cathode. The spike ensured that

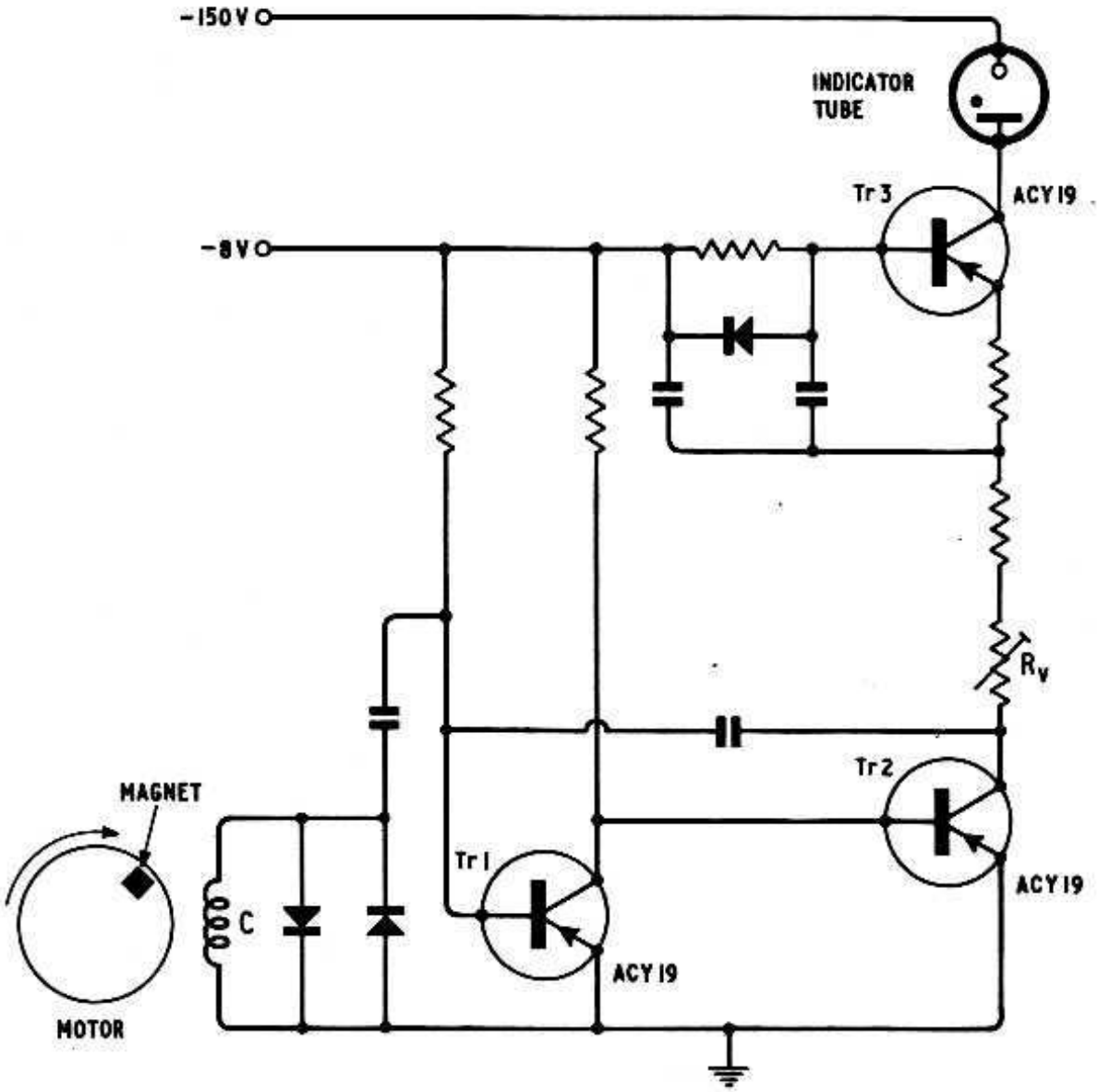


Fig. 9.36. Circuit for indicating motor speed by the glow length in the current indicator tube

the discharge always ignited in the same region, and effectively anchored one end of the glow.

To obtain the brightest glow pure neon was used at a pressure of approximately 20 torr; the measured luminance was about 350 nit. The low pressure gave a less defined edge to the glow than one would have liked, but raising the pressure gave too high a

current density and reduced the stability of the characteristics. In spite of the poor edge to the glow, the glow length could be read with an accuracy of better than 5%. The linearity of the glow length with current was good and could be reproduced from tube to tube.

The breakdown potential of the tube was ~ 146 V and the maintaining potential ~ 115 V, and thus the tube could be controlled by a relatively cheap transistor in series with it. A circuit showing its possible use as a motor speed indicator is given in Fig. 9.36. Pulses from the pick-up coil C are fed to the monostable multivibrator formed by the transistors $Tr1$ and $Tr2$, so that the speed of the motor controls the frequency of current pulses of given magnitude fed to the current control transistor $Tr3$. The condensers and resistor connecting the emitter to the base of this transistor convert the pulse current into a steady current so that the indicator tube does not flicker. The calibration is set by the variable resistor R_V.

The main problem with the tube was life, particularly if the glow remained stationary on a fixed portion of the cathode for any length of time i.e. more than 1,000 hr. Although the general glow length/current characteristic does not change, small jumps occurred on life which appeared to be associated with small areas of contamination. Although these jumps may be less than 10 mm even after thousands of hours life, they precluded the use of the tube in many applications.

9.5.3. GLOW MODULATOR TUBES

Although the exploitation of glow/current characteristics in the normal glow has not resulted in a completely satisfactory tube, the application of the glow/current characteristic to a useful device has been more successful in the abnormal glow region, where the brightness is a function of the current. The device is the glow modulator tube, a source of intense light of small area used in facsimile equipment. A photocell in the transmitter scans the original picture and the variations in current produced are used to modulate a carrier wave. At the receiver end the carrier is demodulated and the current variations fed to the glow modulator tube. This converts the current variations into brightness variations which are projected on to photographic material. For this application, the tube must have a high brightness, preferably in the blue part of the spectrum, which does not fall off during life. The curve of light output against current must be continuous and without steps or discontinuities (it does

not necessarily have to be linear), and the tube must be capable of being modulated at frequencies up to 100 kc/s. A tube with a hollow cathode which satisfies the initial conditions has been described by Rees[28] and is shown schematically in Fig. 9.37. The device consists of a hollow cathode *C* formed by drilling a central hole in the moybdenum cathode rod, which is enclosed in a ceramic insulating sleeve I to prevent ionisation from occurring outside the

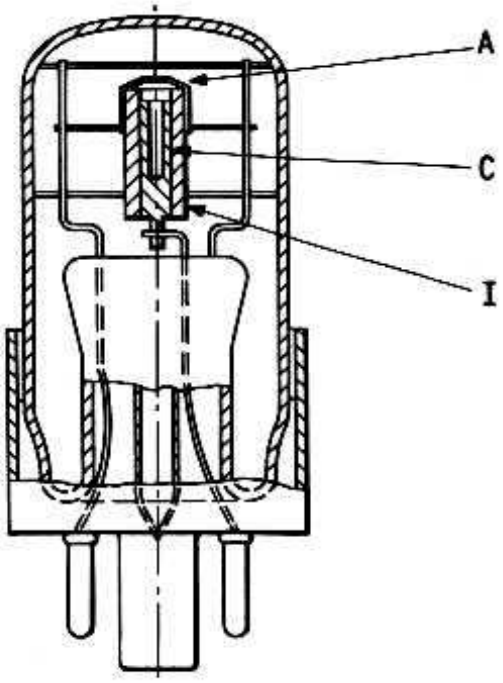


Fig. 9.37. Schematic diagram of the glow modulator tube. Rees[28]

crater. The anode A is mounted above the crater and a hole in the anode permits the beam of light from the hollow cathode to pass through. Two micas serve to position the assembly within the envelope.

Rees showed that the light output from such a tube depended not only on the gas composition and pressure but also on the hole diameter and depth. Neon-argon was found to be a suitable gas, providing the argon content was greater than 10%. The optimum gas pressure was related to the cathode hole diameter, increasing with decreasing hole size. With a cathode current of 30 mA the luminance was of the order of $1{\cdot}7 \times 10^5$ nit. With this current, however, the life was limited to a few hundred hours. The limit of life was due to clean up of the gas as a result of sputtering.

REFERENCES

1. WALSH, J. W. T., *Photometry*, Constable, London, (1958)
2. GAVANIN, V. A., PERELMUTER, V. S., RYBKINA, E. I. and SOKOLOVA, N. S., *Instr. Exp. Tech.*, **5**, 1024 (Sept.-Oct. 1965)
3. LOTHALLER, W. E. and VAN VLODROP, P. H. G., *Electron. Appl.*, **23**, 89, (1962/3)

4. KERR, G., POORTER, T. and VAN VLODROP, P.H.G., *Mullard Tech. Commun.*, **9,** 17, (1966)
5. OMI, H., FUKUKAWA, Y. and NAKAJO, T., *I.R.E. Trans. Electron Devices*, **ED9,** 266, (1962)
6. FUKUKAWA, Y., *Symposium on 'Cold Cathode Tubes and their Applications'*, Cambridge, (1964) [I.E.R.E. Conference proceedings Number 3]
7. BOTDEN, TH. P. J., *Philips Tech. Rev.*, **21,** 267, (1959/60)
8. MALONY, T. C. and GLASER, D., *Symposium on 'Cold Cathode Tubes and their Applications'*, Cambridge, (1964) [I.E.R.E. Conference proceedings Number 3]
9. MCLOUGHLIN, N., REANEY, D. and TURNER, A. W., *Electron. Eng.*, **32,** 140, (1960)
10. NEALE, D. M., *Cold Cathode Tube Circuit Design*, Chapman and Hall, London, p. 219, (1964)
11. HIGGINS, G., *J. Brit. I.R.E.*, **22,** 133, (1961)
12. HODGSON, D. G., *Electronics*, **33,** 80, 1st April, (1960)
13. MACDOUGALL, M. A., *Electronics*, **38,** 78, 22nd March, (1965)
14. DURAND, M., *Les Thyratrons à cathode froide*, Philips, Holland, p. 179, (1963)
15. SOMLYODY, A., *Electronics*, **33,** 48, 8th Jan., (1960)
16. JEYNES, G. F., *Electronics*, **38,** 80, 19th April, (1965)
17. REANEY, D., *Electron. Eng.*, **34,** 372, (1962)
18. ALFVÉN, H. and ROMANUS, H., *Nature*, **160,** 614, (1947)
19. BJÖRKMAN, J. and LINDBERG, L., 'Development of Trochotrons', *Trans. Roy. Inst. Technol. Stockholm*, **No. 80,** (1954)
20. ADBY, P. R., *Electron. Eng.*, **38,** 147, (1966)
21. POORTER, T., VAN VLODROP, P. H. G. and KERR, G., *Electron. Appl.*, **25,** 20, (1964/5)
22. MASH, D. H., *J. Sci. Instr.*, **37,** 47, (1960)
23. GREENBERG, I., *Electronics*, **34,** 31, March 24th (1961)
24. SEIFERT, H. E., *Rad. Electron. Engr.*, **28,** 201, (1964)
25. SEIFERT, H. E., *Electronik*, **11,** 237, (1962)
26. *Electronic News*, p. 5, July 26th (1965)
27. WESTON, G. F., and SCHAGEN, P., *Electron. Engr.*, **36,** 21, (1964)
28. REES, D., *Symposium 'Cold Cathode Tubes and their Applications'*, Cambridge, (1964) [I.E.R.E. conference proceedings Number 3]

Index

Abnormal glow region 336
After effect 100
Afterglow, effect of neon 177
  electron density in 173–74
  measurements in hydrogen 181
  microwave studies of 173–81
Alpha-numeric display tubes 355
  drive circuit 357
'AND' gate 235, 268, 269, 271, 276, 310, 351
Anode fall 69, 78, 88, 92, 97
Anode glow, transition of 92
Auger process, single step 31

Binary-coded-decimal counting systems 351–55
Blocking oscillator 295
Boltzmann's constant 105, 171
Breakdown, at low pressures 41–66
  in non-uniform fields 57
Breakdown characteristics 345
  of inert gases 228
    uniform field 43–53
Breakdown mechanism 160
Breakdown potential 22, 41, 42, 46, 47, 151, 192, 253, 257, 258, 259, 278, 283, 345, 365
  distribution 285
  effect of Penning reaction 54
  for corona discharge 218
  for inert gas mixtures 53–57
  lowering of 65
  minimum 42, 48–53, 59
  priming current effect on 62–65
Breakdown voltage 192, 256
  minimum 48, 49

Capacitor triggering 242, 243
Cathode contamination 258, 293–94, 303, 334
Cathode dark space 70, 72, 110
Cathode discharge, hollow 107–13
  impedance of 111–13
  negative *V-i* characteristic 111
  see also Hollow cathode effect
Cathode fall 69
  normal, correlation with work function 82
    effect of temperature 92
    for inert gases 78–87
  theory of 71–78
Cathode follower circuit 196
Cathode material, trigger tubes 251
Cathode zone of glow discharge 70–87
Cathodic sputtering. *See* Sputtering
Cavity cathode 112
Continuity equation 71, 161, 162
Control circuits 249, 263–68
Corona discharge 62, 218
Corona stabiliser 218–24
  current range 222–24
  design of 218–24
  hydrogen-filled 223
  incremental resistance 219–20
Counters 271
  bi-directional 285
  reversible 312–13
Counting circuits 233–35
Counting systems, binary-coded-decimal 351-55
Counting tubes, beam-switching 350
Coupling circuit, stepping tube 306
  transistor 307, 345
Current indicating tubes 362–65

Decoder, diode matrix 351
Decoder/encoder 360
Degression time 99
Deionisation, effect of argon or krypton 178
  effect of metastable atoms 176
  effect of xenon and krypton 180
  equation for 172
  measurements of 172
  theory of 169–72
Deionisation curve for switching diode 227
Deionisation time 151, 168–89
Dekatron 282, 286, 310
Delay effects 101
Delay times, classification of 151
Detection circuit 264
Dielectric loss due to gas molecules 101
Diffusion 170–72, 181
  ambipolar 171, 178, 180, 184, 188
Diffusion coefficient 171
  ambipolar 171
Diode matrix decoder 351
Discharge, voltage-current characteristics 1
Display devices, miscellaneous 360–66
Display panels, large area 360–62
Display tubes 329–60
  alpha-numeric 355
    drive circuit 357
  general characteristics 329–30
  matrix 330
  numerical or character 316
  segmented-cathode 360
  stacked-cathode 330
  visual output 316
'Drift' of experimental tubes 203

Electrode geometry, of stabilisers 204–08
  stepping tubes 294
  trigger tubes 252–56
Electrode material, for stabilisers 201–04
  indicator tubes 326
  stepping tubes 293
Electrode structure, switching diodes 227
  trigger tubes 256
Electroluminescent panels 360
Electron balance equation 72
Electron density in afterglow 173–74
Energy balance equation 5
Energy distribution 37
Energy distribution function 4
Escape probability 37

Faraday dark space 68, 70, 89, 91, 107
Flame-failure circuit 264
Flip-flop binary counter 233
Formative lag 151, 157–68, 241
  analysis 156–64
  experimental 163
  measurement of 164–68
Frequency division, stepping tubes 309

Gas clean-up in glow discharge 146
Gas filling, indicator tube 326, 332
  stabilisers 198–201
  stepping tubes 291–93
  trigger tubes 250
Glow/current characteristic 365
Glow discharge 67–114
  abnormal 67
  cathode zone 70–78
  gas clean-up 146
  general features 67–70
  impedance of 97–104
  negative 107
  noise in 104–07
  normal 67
  potential between electrodes 67
  sputtering in. *See* Sputtering
  temperature effects 92–97
  temporal growth and decay 151–90
  V-i characteristic 67
  voltage across 97–98
Glow-discharge diode 225
Glow modulator tubes 365

Hollow cathode effect 108–10

Impedance, glow-discharge 97–104
  hollow cathode discharge 111–13
  stabilisers 216–18
Indicator lamps 316
Indicator tubes, gas filling 332
  glow position 325–29
    electrode materials 326
    gas filling 326
    low voltage 327
  glow-transfer 322-25
  numerical 330

Indicator tubes, gas filling *continued*
numerical *continued*
current limits 336–38
current-sharing of 'off' cathodes 338–39
design and construction 332–36
direct drive by stepping tube with auxiliary anodes 348
drive circuits 340–55
drive from binary counting system 351
life of 334, 339–40
operation and performance 336–40
stepping tube drive 342–48
trigger-tube ring counter drive 340
trochotron drive 350
V-i characteristics 323–24, 338
Indicators 227
glow-position 315
on-off 315, 316
transistor circuits 319
Inert gas-hydrogen mixtures 292
Inert gas mixtures 101, 188
breakdown characteristics of 228
Inert gases, breakdown characteristics, uniform-field 43–53
breakdown potentials for 53–57
deionisation 176–77
gas clean-up of 146
normal cathode fall 78–87
Information store, stepping tubes 314
Ion drift velocities 28
Ion transit time 101
Ionisation coefficients, Townsend. *See* Townsend coefficients
Ionisation cross sections 176
Ionisation potential 31

Logic circuits 235, 249
stepping tubes 311
trigger circuits 268-76
Luminance 316
Luminous flux distribution 318

Maintaining potential 192, 258, 362, 365
as function of envelope temperature 92
as function of *pd* 91
drifts in 210
effect of pressure 92–94
Maintaining potential *continued*
minimum 91
temperature coefficient 94–95, 214
Maxwellian energy distribution 5
Memory circuits 235
Mercury vapour 318
Metastable atoms 24–26, 38–39
Microwave studies of afterglow 173–81
Multiplication factor 152, 153, 158 160
Multiplication processes 106, 107

Negative glow 315, 316
Neon lamps 316
Neon-xenon mixture 323
Noise, in glow discharge 104–07
in stabiliser tubes 105, 217–18
shot 104–05
thermal 105
'white' 105
Noise mechanisms 106
Noise sources 107
Normal glow region 336
'NOT' gate 236, 269, 271, 276

'OR' gate 269, 271, 276
Overvoltage detector 264
Overvoltage/strike delay characteristic 166, 225

Parallel-plate discharge, *V-i* characteristics of 62
Paschen characteristics 250
Paschen curve 41, 43, 45, 47, 56, 199, 328
Paschen minimum 46, 56, 228, 292, 326
Paschen's law 12, 41–43, 47, 48, 56, 59, 71
Penning effect 16, 18-20, 54, 60, 110
Penning gauge 362
Penning mixture 53, 56, 84, 176, 181, 199, 250, 292, 326, 328
Penning reaction 54, 85, 86, 178, 187, 199, 201
Photoelectric work function 36
Photoemission 28, 159, 161, 162
Photoform glass 361
Photoionisation 20–21, 160–61
Photon absorption 56
Photons 24–27, 29
Poisson's equation 64–65, 71, 72, 162

Positive column 88
Positive ions 24–27
Potential difference 3, 338
Potential distribution 67
Potential drop 69
Priming, stabilisers 208–10
  stepping tube 281
  trigger tube 242, 255
Priming current, effect of breakdown potential 62–65
Protection circuit 264
Pulse-plus-bias circuits 268
Pulse-plus-bias triggering 245
Pulse programming device, steeping tube 312
Pulse-train generation, stepping tubes 311

Reactance 102–03
Read-out, in line 329
Read-out circuit 320
Read-out diodes 227, 228
Recombination 169, 172, 178–79
  processes 170, 175
Recombination coefficient 169, 174, 181
  and mobilities 180
Recombination losses, correcting for 174
Recovery curve, voltage-time 185
Recovery time, measurements of 181–89
  pressure dependence 182
Reference tube 191, 195–97
Reference voltage 195
Register tubes 315
Relaxation diode 227
Relaxation oscillations 229, 276
Relay circuits 263–68
Resistance 102–03
Reversible counter 312–13
Ring counter 234, 274
Ring counter drive, numerical indicator tube 340
Routing guides 313

Scaler circuits, stepping tubes 306
Screening of envelope from deposited layer 146–49
Self-quenching and relaxation oscillations 229
Space-charge distortion 62, 65, 157, 162, 163, 240
Sputtering 79, 112, 115–50, 166, 198, 202, 227, 259, 294, 334
  cleaning of electrodes by 144–46
  collision theories 133–37
  deposition rate 140
  effect of geometry and degree of 157
  effect of ion mass 132
  evaporation theory 132, 140
  gas clean-up in glow discharge 146
  in glow discharge 137–49
  measurements at low pressures 116–32
  nature of process 115–16
  of insulating materials 130
  of single crystals 130–32
  rate of 142–45
  screening of envelope from deposited layer 146–49
  theories of 132–37
  threshold energy 133–35
  yield, as function of ion energy 121
    for low energy ion bombardment 125
  yield determination 116
  yield formulae 134, 136
  yield measurements, in arc discharge 118
    in glow discharge 138–44
    in magnetic field 117
    using ion beams 128–30
Stabilisation 191–224, 276–79
  direct 191
  thermionic valve 195
Stabilisation ratio 195, 196
Stabiliser circuits, trigger tubes 276–79
Stabiliser tubes, noise in 105
Stabilisers 191, 226
  corona 218–24
    current range 222–24
    design of 218–24
    hydrogen-filled 223
    incremental resistance 219–30
  design 197–210
  drifts in maintaining potential 210–14
  effect of cathode shape 206
  electrode geometry of 204–08
  electrode material for 201–04
  gas filling of 198–201
  impedance 216–18
  impedance-frequency curves 216
  life behaviour 210–14
  noise 217–18
  performance 210–18

Stabilisers *continued*
priming of 208–10
principle of operation 192–97
sealing 208
temperature coefficient of maintaining potential 214
Statistical breakdown delay of switching diodes 227
Statistical lag 151–57, 241
effect of tritium 157
mean 152, 153
experimental investigation 155
method of measuring 154
theory 152
Stepping tube drive, numerical indicator tube 342–48
with auxiliary anodes 348
Stepping tubes 281–314
applications 306–14
coupling circuits 296, 306
design 290–95
double-pulse, circuit design and performance 295–306
principle of operation 283–86
drive circuits 295
blocking oscillator 295
differentiated and integrated pulse 301
rectangular and integrated pulse 296
sine-wave 302
drive pulses 295–302
electrode geometry 294
electrode materials 293
frequency division 309
gas filling 291–93
information store 314
life of 305
logic circuits 311
output signal 303
positive guide bias range 302
priming 281
principle of operation 281–90
double-pulse 283–86
single-pulse 286–90
pulse amplitude 303
pulse programming device 312
pulse-train generation 311
recommended operating ranges 305
requirements 290
scaler circuits 306
single-pulse, principle of operation 286–90
transfer characteristics 291, 294
Stepping tubes *continued*
waveform generation 314
Streamer mechanism 160–61, 165
Supply voltage 192
Switching diodes 225–38
applications of 229–38
counting circuits 233–35
deionisation curve 227
electrode structure 227
logic applications 235–38
practical design 226–29
required characteristics of 225
self-quenching and relaxation oscillations 229
statistical breakdown delay of 227
timers and tone generators 231
with metal cathodes 227
with sputtered metal cathode 228
Switching tubes, grid controlled 360

Telephone exchange, trigger-tube 274–76
Temperature coefficient of maintaining potential 94–95, 214
Temperature effects in glow discharge 92–97
Thermionic triodes 343
Thermionic valve stabilisation 195
Time delays 241
Time-lag curves 23–24
Timers 231, 265–68, 310
Timing circuits 263–68
Timing period 266, 268
Tone generator in electronic organs 233
Townsend coefficient 1–39
first 2–21, 53, 70
dependence on field and pressure 3
effect of gas mixtures 16
measured values in pure gases 8
second 21–39, 257, 290, 293, 364
direct measurement 31–39
dynamic studies of mechanism 23
effect of adsorbed gas layers 33
Townsend discharge 157
Townsend equation 62, 64
Transfer characteristics, stepping tubes 291, 294
trigger tube 253–54
Transistor circuits, indicators for 319
logic 268–76
Transistor coupling circuit 307, 345

Trigger tubes 238–99
application 249
capacitor triggering 242, 243
cathode material 251
circuit diagrams 242
circuits, and their applications 263–79
operation 242–47
constructions 255
contamination 258
control circuits 249, 263–68
counters 271
design of 249–56
development 238
electrode configuration 252–56
electrode structure 256
extinguishing the discharge 248
firing methods 242
gas filling 250
life of 258–63
logic circuits 249
output voltage 250
performance 256–63
priming 242, 255
principle of operation 239–42
Trigger tubes *continued*
private telephone exchange 274–76
pulse-plus-bias circuits 268
pulse-plus-bias triggering 245
relay circuits 263–68
relay operation 255
ring counter drive, numerical indicator tube 340
sensitivity 253–55
stabiliser circuits 276–79
time delays 241
timing circuits 263–68
transfer characteristics 244, 253–54
Trochotron drive, numerical indicator tube 350

Voltage-current characteristics of discharge 1
Voltage stabilisation 276

Waveform generation, stepping tube 314
Work function 50, 82, 83, 120