The transition to conduction in diffusion-controlled gas breakdown

Boyer, M. F.

THE TRANSITION TO CONDUCTION IN DIFFUSION-CONTROLLED GAS BREAKDOWN.

by

M. F. BOYER, B.Sc.

being a thesis submitted for the Degree of Doctor of Philosophy in the University of Durham.

June 1965

P. 247

ABSTRACT

Measurements have been made of ultra-high-frequency (180Mc/s) breakdown fields in hydrogen, nitrogen and neon, in a Rogowski-profiled spark gap, under conditions of pressure in which both the electron ambit and the mean-free-path were less than the gap width (0.494cm). The measurements were made with sustained fields and also with pulsed fields in overvolted conditions to obtain data on the variation of breakdown delays with overvoltage. The delays were measured by photographing an oscilloscope trace which displayed the gap voltage, and a range of 400 to 0.2 microseconds was observed with overvoltages up to 100%. Breakdown was initiated during a pulse by irradiating the gas in the mid-gap region with a short (about 0.1 microsecond) burst of ultra-violet photons.

A theory is proposed which determines the shape of the oscillograms, and hence gives breakdown delays. Good agreement was obtained between theoretical and experimental delays for hydrogen and nitrogen, but predicted delays were too short by factors up to ten for neon. A qualitative explanation is given for the anomalous results for neon.

Maintaining voltages have been measured in hydrogen and nitrogen, as a function of the gas pressure, but were too low to measure in neon. In all three gases it is shown that the maintaining fields are consistent with losses by ambi-polar

diffusion, although electrons which approach within one ambit of an electrode are driven into that electrode by the action of the field.

Immediately after breakdown in overvolted conditions the voltage approached zero and then recovered to approach the maintaining voltage asymptotically. The recovery time-constants were of the order which might be expected if an initial excess electron population were being diminished by ambi-polar diffusion.

PREFACE

The work which is described in this Thesis is a continuation of some experiments which were begun by Mr. J. R. Rowbotham and Mr. P. G. Monk in the University of Durham and under the supervision of Dr. W. A. Prowse. The present experiments were carried out by myself and Mr. R. G. Earl in close co-operation, also under the supervision of Dr. Prowse. Mr. Earl is also a candidate for a Ph.D degree and therefore the work has been divided between us, with Mr Earl attacking the problem of formative and statistical delays in gases subjected to pulsed high frequency fields, while I concentrate on the changes which take place in the spark gap between the initiation of a discharge and the establishment of a steady glow discharge. Inevitably there will be some overlap, since we both use essentially the same results, but for this reason very little mention of the statistical element of the delay is made here while maintaining voltages and ambi-polar diffusion are discussed in some detail.

CONTENTS

LIST OF FIGURES.

CHAPTER I

INTRODUCTION

The qualitative study of electrical conduction in gases was begun in the nineteenth century but little progress could be made theoretically until a physical model of the atom evolved. The necessary model was provided by the discovery of free electrons by J.J. Thompson in 1897 and Rutherford's hypothesis of a nuclear atom with electrostatically bonded electrons in 1911. Townsend was then able to apply the concept of ionization by collision of negative particles with atoms to give the first quantitative account of the buildup of electrons to give breakdown in a gas.

Townsend's initial publications[1] assumed that these negative particles were ions, but by about 1910 he recognised that electrons were responsible for ionization. Townsend considered an electron moving in a gas making random collisions with molecules: when an electric field is applied the electron still makes random collisions but has an overall tendency to move in the direction of the field. The average velocity of an electron in the field direction gradually increases towards a terminal value, at which time the average gain in energy between collisions equals the average loss per collision. If the field is sufficiently high the electron may then have sufficient energy to ionize a molecule, at a collision.

Suppose n molecules travel a distance dx in the field

direction, then the number of new electrons produced, dn, is proportional to n and also to dx, so we may write

$$dn = \alpha n \, dx$$

where α is a constant of proportionality, which is known as Townsend's first ionization coefficient.

Integrating, we have $\log n/n_0 = \alpha d$

where $n = n_0$ at $x = 0$, and d is the gap width

Therefore, $n = n_0 \exp(\alpha d)$

If we think of n as the number of electrons arriving at the anode in unit time, the current $i = ne$ where e is the charge on one electron,

$$i = i_0 \exp(\alpha d).$$

This is precisely the law of variation of i with gap width that Townsend observed experimentally when the electric field was maintained constant. However, breakdown is not predicted by this equation, because removal of the initial current (i_0) would reduce i to zero, whereas in practice once breakdown has been observed, the current is self sustaining. Townsend therefore postulated a secondary emission process due to which a factor γ enters. γ is the number of electrons produced at the cathode by the impact of a positive ion, or photon. With the same notation as before, the anode current is now

$$i = \frac{i_0 \exp \alpha d}{1 \mp \gamma(1 - \exp \alpha d)}$$

The current now tends to infinity as exp (αd) tends to $(1+\gamma)/\gamma$ without the need for i_o to remain. Although this equation does not make allowance for the effects of space charge, which become important at relatively large current it is apparent that a process of this kind could lead to breakdown. In fact this equation fits experimental results very well, even within a few per cent of the breakdown gap width, and with currents up to 10^{-7} amps[2].

Let us now consider the application of these ideas to the case of a sinusoidally varying electric field.

As the frequency is increased to a few thousand cycles per second, the field still remains high for long enough during the peak of one half cycle to allow many transits of particles and photons across the gap. Therefore provided i_o is maintained to initiate a discharge, the full breakdown can still take place at the same value of peak voltage as with a unidirectional field.

At a few tens of kilocycles the field must be raised to allow breakdown within one half cycle. At about 100 kilocycles a reduction in the breakdown field occurs, when all the positive ions are no longer swept into the electrodes. However an appreciable number still moves close to the temporary cathode and maintains the secondary process. Further increase of frequency results in a second maximum in the breakdown field. This is because the reduced amplitude

of oscillation (ambit) of positive ions tends to keep them in the mid-gap region and γ processes are prevented. These departures of the breakdown field from the d.c. value would be typically not more than 20%.

At a few megacycles per second, the breakdown field decreases rapidly to about 40% of the unidirectional value, when electrons are no longer swept out of the gap by the field. Although there are now no secondary processes to maintain a flow of electrons into the gap, the electrons are lost only by diffusion, and a much smaller field can establish breakdown. This type of breakdown is called "ultra high frequency" (u.h.f).

Obviously the precise frequencies and fields depend to a large extent on the actual gas, pressure, and electrodes used, and the figures given[3] would apply to a typical hydrogen discharge at a pressure of about 8 torr and a gap width of 10mm.

Finally, when the electron ambit becomes less than the mean free path, (which is at microwave frequencies for pressures greater than 1 torr,) the breakdown field rises, because the phase of the oscillation is then in quadrature with the applied field and therefore the electron gains no energy when this is averaged over a complete cycle of the field. A sinusoidal oscillation is such that an electron spends the bulk of its time with nearly zero velocity, and

therefore the majority of molecular collisions result in no ionization.

Microwave breakdown has been extensively investigated[4,5,6,7] in the years since the Second World War. This has been made possible by techniques, which were developed as part of the wartime radar effort, being adapted for precise measurements of many aspects of gas discharge breakdown.

It has been shown by Professors A.D. MacDonald and S.C. Brown[5,6,7] and their collaborators that it is possible to calculate the breakdown fields for some gases quite accurately. This was accomplished by taking published ionization efficiencies and excitation potentials of the gas used, together with the gas pressure and gap geometry, to calculate the electron energy distribution function and hence the rates of electron production and loss, as a function of the applied field. When the production and loss terms are equal, the breakdown occurs. Theoretically an infinite time would elapse before breakdown, but in practice, a field very slightly greater than the breakdown value is sufficient to cause the gap current to increase rapidly to a value limited only by the external circuit. The breakdown criterion is analogous to that of Townsend, discussed above, because in this case also, if the external generator of the initiating electrons is removed, the discharge is self-sustained.

When a field which is in excess of that required to give

breakdown is applied to a spark gap, the gap is said to be overvolted. In these circumstances it is found, as might be expected, that the application of a high overvoltage causes breakdown sooner than when a lower overvoltage is applied. In the general case, for a fixed overvoltage and pressure, a certain minimum delay, called the formative lag, occurs between the application of the field and breakdown, which is generally supposed to be the time required for the development of the discharge. A second component of the total observed delay is called the statistical lag. This is random and is thought to be the time elapsing from the application of the field, until an event occurs which can initiate the discharge.

It will be clear that in an overvolted gap the breakdown criteria which we have already mentioned are inapplicable. There is, in this case, no possibility of a quasi-equilibrium condition where electron losses and gains are equal. Instead, there is a net rate of gain of electrons and an unstable state exists. It is the principal purpose of this work to present results which have been obtained in an experimental study of this unstable state, which occurs in the transition to conduction. Therefore it is appropriate to consider further the definition of breakdown. When the case of "infinite" delays was considered, the only criterion needed for practical purposes was that the gas should become

conducting. However, a study of the approach to the conducting state raises the question : at what time does the change-over take place? There is no single answer, because the gas can be said to be conducting from the time of the event which produces one free electron, right through the processes of electron multiplication till the external field is removed, and further, till the last charged particle has disappeared from the gap.

It is therefore necessary to define some sort of level at which we can consider breakdown to have occurred, and it is clear that several characteristic features of the discharge could be used. For instance, we could choose a current or a voltage level, either of which would be completely arbitrary. Perhaps more satisfactory would be a definition in terms of basic properties of the discharge. Such a property is that of ambi-polar diffusion: consideration of a homogeneous neutral collection of positive ions and electrons undergoing random thermal motion, shows that the system is unstable, because electrons, by virtue of their small mass, diffuse outwards faster than the heavy ions, and leave behind a net positive space charge. The attraction of the space charge retards the outward motion of electrons, and accelerates the ions, and results in the net diffusion rate being modified to the so-called ambi-polar rate. The modification to the free diffusion rate becomes significant

only when sufficient space charge is present and causes a large increase in the existing excess rate of ionization over loss. This marks a point-of-no-return which it is possible to define as breakdown. However, this point was not readily ascertainable in the present experiments: certainly no discontinuity appeared on oscilloscope traces showing the gap potential. Therefore in this work an arbitrary definition was used: breakdown delays were measured at a point corresponding to a 1% decrease in the gap voltage level.

The work to be described is a continuation of some experiments carried out by Mr. J.R. Rowbotham and later by Mr. P.G. Monk under the supervision of Dr. W.A. Prowse. The main object of those experiments was the determination of formative and statistical delays in fields of frequency 180 Mc/s under conditions of pressure which ensured an electron ambit less than the gap width. Their results have been published[9] but the principal observations will be summarized briefly here, for convenience:

(1) In hydrogen no delays were observed at all, the least alteration in field strength that could be detected, being sufficient to cause a change from no breakdown within the pulse duration, to no detectable delay.

(2) Breakdown in air was preceded by a statistical delay, but again there was no detectable formative time.

(3) In neon, both formative and statistical delays were appreciable, and graphs were plotted of formative delays as a function of applied voltage. The r.m.s. breakdown voltages for delays of 200μsec, and 10μsec were respectively 2% and 50% greater than the value obtained for breakdown under sustained fields.

In order to analyse the statistical variations in the breakdown delays the fraction of observations having delay less than a given value was determined, and the fraction plotted against the given value, for a constant applied voltage. These graphs were linear, indicating constant probability of breakdown at a certain delay.

The different results obtained for these three gases suggest that radically different processes caused breakdown. In contrast, the results of McDonald and Betts for neon[7], Reder and Brown for helium[6], and McDonald and Brown for hydrogen[5], using sustained microwave fields, agree very well with a theory based on ionization by collision. This seems to suggest that for sustained fields, at least, a single process determines breakdown. Taking these results at face value it might appear that (a) the processes which are important for the breakdown of overvolted gaps differ from those which are important at the breakdown threshold, (b) monatomic gases behave differently to diatomic gases when overvolted and(c) the presence of oxygen in air causes

prominent statistical delays which are not observed in hydrogen.

It must be realised, however, that although "spectroscopically pure" gases were used in those experiments[9] no attempt was made at outgassing the gas enclosure before admitting the sample, and only a rotary pump was used for the preliminary evacuation. Grease vapour evolved from the stopcocks and water vapour evolved from the glass walls would have been removed by the liquid air traps used, but nitrogen and oxygen impurities were almost certainly present. Mercury vapour from the manometer was excluded by using a differential bellows gauge for pressure measurement. Further, the measurements of breakdown delays were limited to a minimum of 10 microseconds, and voltages could be read to a relative accuracy of 2%. The uncertainties, could therefore be such as to allow the retention of the hypothesis that a single simple process can explain breakdown both under sustained fields and pulsed overvoltages.

The experiments which will be described in this thesis were designed to help to resolve the mechanism of breakdown in the case of the overvolted gap. The existing apparatus as used by Prowse, Rowbotham and Monk was developed so that greater precision could be achieved in the measurements of gap voltage and breakdown delays, and the minimum delay observable was reduced by about two orders of magnitude. In addition,

particular attention was paid to increasing the gaseous purity by the adoption of high vacuum techniques.

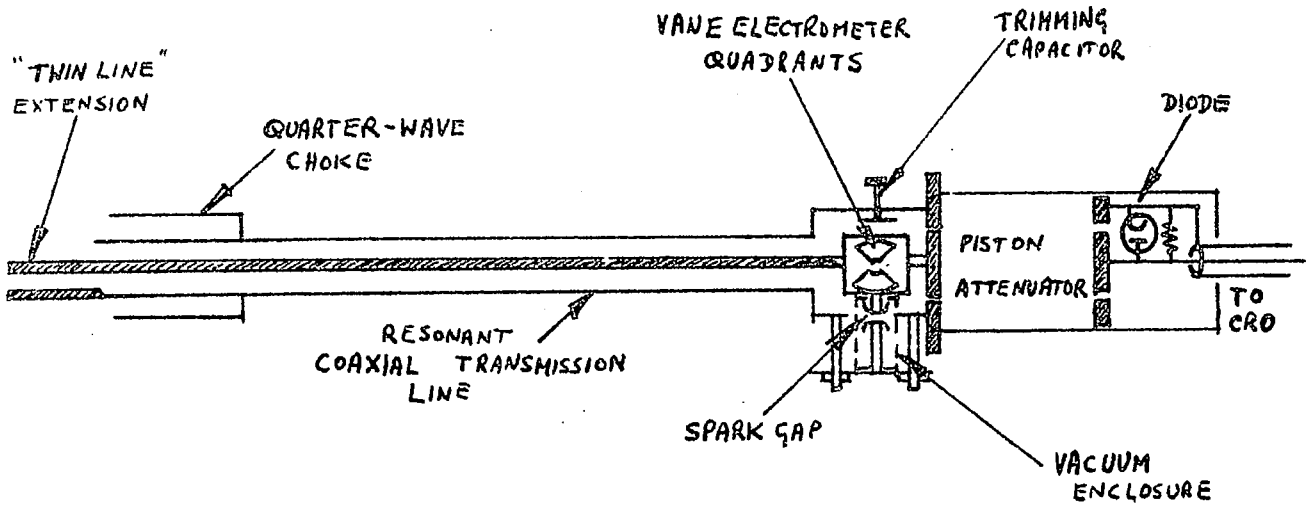


FIG 2.1 A LONGITUDINAL SECTION OF THE TRANSMISSION LINE AND PISTON ATTENUATOR ASSEMBLY

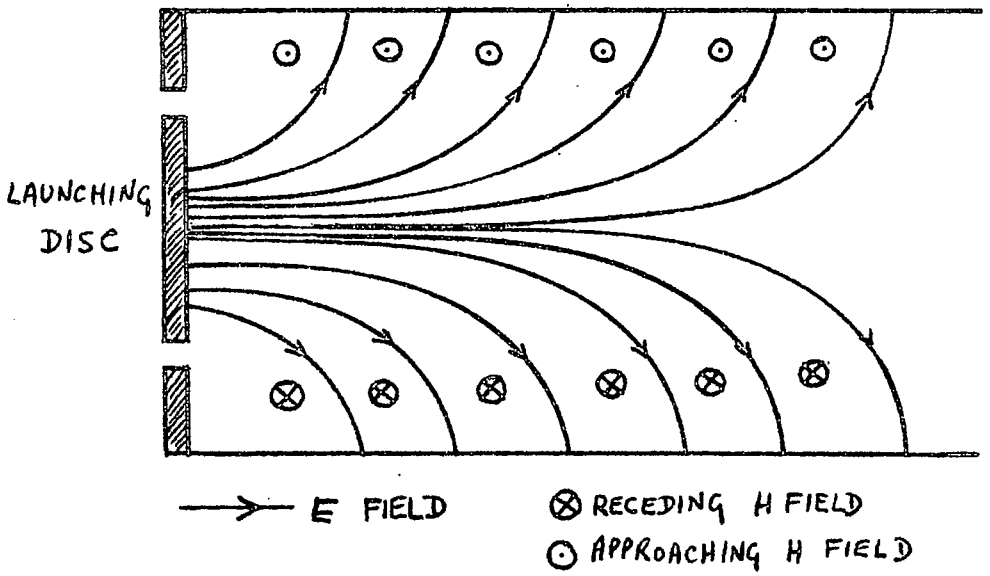


FIG 2.2 E_{01} EVANESCENT MODE IN A CYLINDRICAL PIPE

CHAPTER II

A GENERAL DESCRIPTION OF THE APPARATUS

This chapter contains a description of the basic elements and functions of the apparatus designed by Prowse, Rowbotham and Monk which was used to obtain the results[9] which were described briefly in the Introduction. The same basic elements were used for the present work, but several modifications were made for the purpose of obtaining increased accuracy and stability. These modifications will be mentioned here, but will be described in more detail in succeeding chapters.

2.1 The transmission line

The spark gap was a pair of Rogowski-profiled electrodes situated at one end of a coaxial transmission line. One of the electrodes was attached to the inner conductor, and the other to the earthed outer conductor, each by way of a short stout copper conductor which passed through the glass walls of the vacuum enclosure.

At the end of the line remote from the spark gap the inner conductor projected 40cm beyond the end of the outer. Parallel to this extension was a continuation of the outer conductor in the form of a copper tube, similar in size to the inner - see fig 2.1. These two extensions therefore formed a short length of parallel-twin transmission line. The characteristics of a twin line are such that it will pick

up power from (and radiate power into,) an electromagnetic field. In contrast the interior region of a coaxial line is effectively isolated from the influence of external electromagnetic fields and does not radiate. Therefore a voltage wave induced on the twin line could travel down the coaxial line to the spark gap with little attenuation. Such a wave was created by placing a high frequency oscillator near to the parallel-twin section of the line. This had a nominal frequency of 180 Mc/s and a maximum dissipation of 150 watts.

The overall length of the transmission line was about two wavelengths (330cm) at the oscillator frequency so that resonance occurred, and a slow-motion extension mechanism at the oscillator end of the line was used to facilitate exact tuning. Also, a small variable capacitor was connected across the line near the spark gap to fulfil a similar function. This was initially incorporated to compensate for the variable spark gap width, but it was not used here because the present experiments were carried out with a fixed gap width of 0.494cm.

The outer conductor of the coaxial line was made of copper pipe 5.0cm outside diameter, and thickness 0.15cm, and the inner conductor had outside diameter 0.9cm, and was also made of copper. The line was mounted horizontally, about 1.5 metres from the floor, and served three roles:

(i) Its length physically separated the oscillator from the measurement apparatus at the spark gap end, reducing the problem of high frequency interference.

(ii) The resonant property produced a high voltage at the spark gap whilst minimising both the power output requirement of the r.f. oscillator and the amount of energy dissipated in the spark gap.

(iii) The earthed outer conductor acted as a screen to prevent interference by external electric fields, and the highly selective nature of the line also helped to reduce interference.

In the first experiments[9] the line was short-circuit terminated at the input end but in the experiments to be described here the line was unterminated, because this arrangement was found to provide a greater overall range of voltages at the spark gap.

It is possible for an electromagnetic wave to flow down the outside of a coaxial line, and as this mode of propagation was undesirable here because of the danger of interference in the voltage measuring devices, it was prevented by mounting a concentric quarter-wavelength choke on the outside of the line. The form of this choke is illustrated in fig 2.1. The position of the choke relative to the end of the coaxial line was found to be critical in its effect on the resonant quality factor (Q) of the line, and was adjusted to give maximum Q.

Reaction of the resonant line back on the oscillator could cause frequency jumping if the coupling was too close. Provision for adjustment of the coupling was therefore made by suspending the oscillator on a counterweighted system that allowed vertical movement directly below the input end of the transmission line.

2.2. The piston attenuator

The spark gap capacitance was of the order 0.5 pF and therefore this end of the line could be considered approximately as an open circuit. The spark gap was therefore situated close to a voltage maximum at resonance. The experiments required the attainment of a high voltage at this point, and its accurate measurement, so that very little power could be taken away by measuring instruments, and therefore a piston attenuator was used to connect them to the line. The geometrical form of the piston attenuator can be seen in section in fig 2.1: A circular brass disc was attached to the inner conductor of the coaxial line, immediately beyond its connection to the spark gap, which launched an electromagnetic wave into the adjacent section of cylindrical waveguide. The waveguide had dimensions less than those required for propagation of waves, so that a wave of exponentially decaying amplitude existed along its length. Such a wave, in which the electric and magnetic fields oscillate in quadrature is referred to as an evanescent mode.

The mode used is known as the E_{01} mode, and the field pattern pertaining to this is shown in fig 2.2. The E_{01} mode was used in preference to the H_{11} mode which is the dominant mode in a cylindrical waveguide,[12] because it could be launched by the high impedance disc. The H_{11} mode is launched from an inductive loop with low impedance and was therefore unsuitable for this application.

A second disc, similar to the launching disc, placed a little further down the waveguide served to pick up the attenuated waves. The resulting voltage was rectified by a miniature thermionic diode placed immediately behind the pickup disc. The remaining high frequency component was filtered out and the voltage output obtained approximated closely to a linear attenuation of the envelope of the sinusoidal voltage at the spark gap.

The electromagnetic waves induced on the transmission line by the oscillator suffered reflections at both ends of the line, with phase changes such as to produce a steady buildup of energy on the line from the instant that the oscillator was switched on. Energy was removed from the waves by (a) resistive losses in the walls of the line, (b) radiation from the parallel twin section, (c) dissipation in the terminations. A constant amplitude would be reached when the losses became equal to the rate of energy supply from the oscillator. Gas breakdown in the spark gap upset this

equilibrium, and could be observed as a decrease of voltage at the output terminals of the piston attenuator. The piston attenuator output after rectification was observed on a cathode ray oscilloscope (CRO), the display of which could be photographed, when required.

The application of an electric field to a region containing gas produces no effect unless charged particles are present. It has already been mentioned that Townsend observed the exponential increase of a number of electrons in his measurements of the coefficient α. The initial electrons in his experiments were produced by the action of photons on the cathode. An alternative method which has been used successfully in these laboratories[11,13] is to liberate electrons from a hot filament. Discharges in the experiments to be described here, however, were initiated by electrons generated in mid-gap by the photo-ionization of the test gas. For this purpose, ultra-violet photons were produced by an auxiliary discharge - called the "irradiator" - in the same gas enclosure as the test spark gap, and in line-of-sight with it. A small collimating aperture was placed between the irradiator and the spark gap to ensure that no photons could hit either electrode. This technique has been used by Bainbridge and Prowse[10,14] to obtain photon absorption coefficients in nitrogen, oxygen, air and argon.

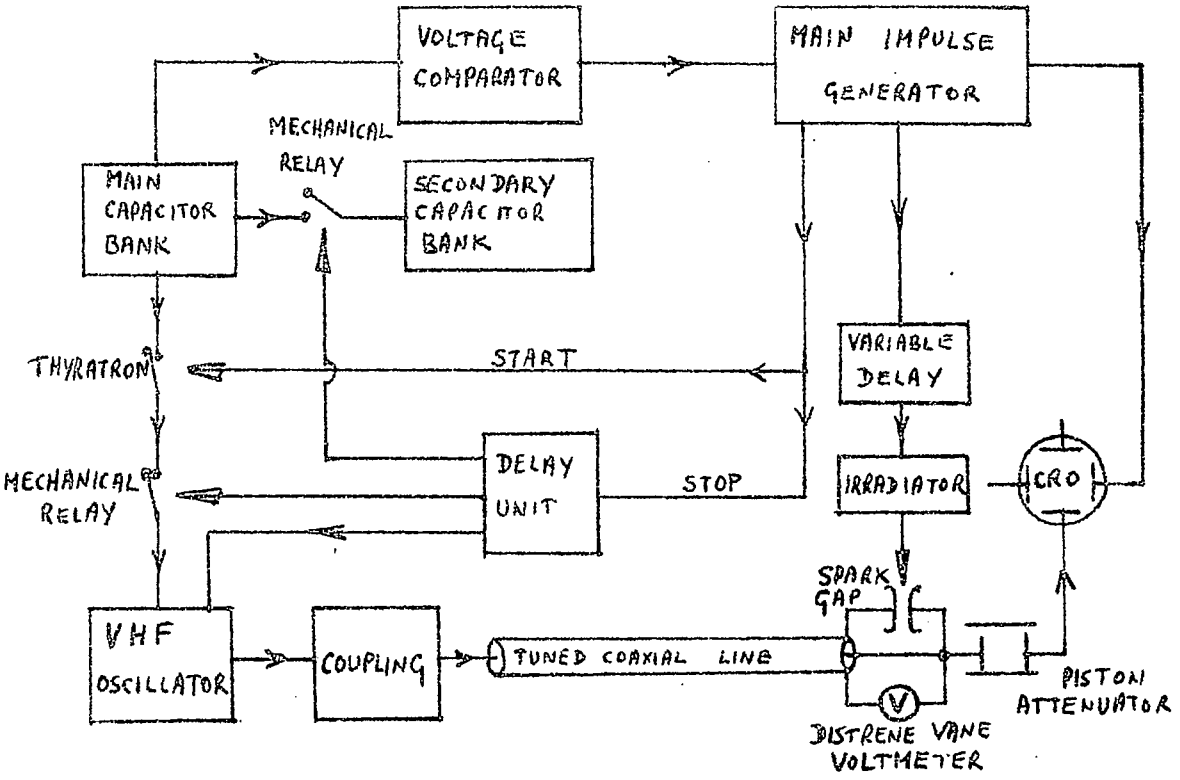


FIG 2.3 BLOCK DIAGRAM OF ORIGINAL APPARATUS (9)

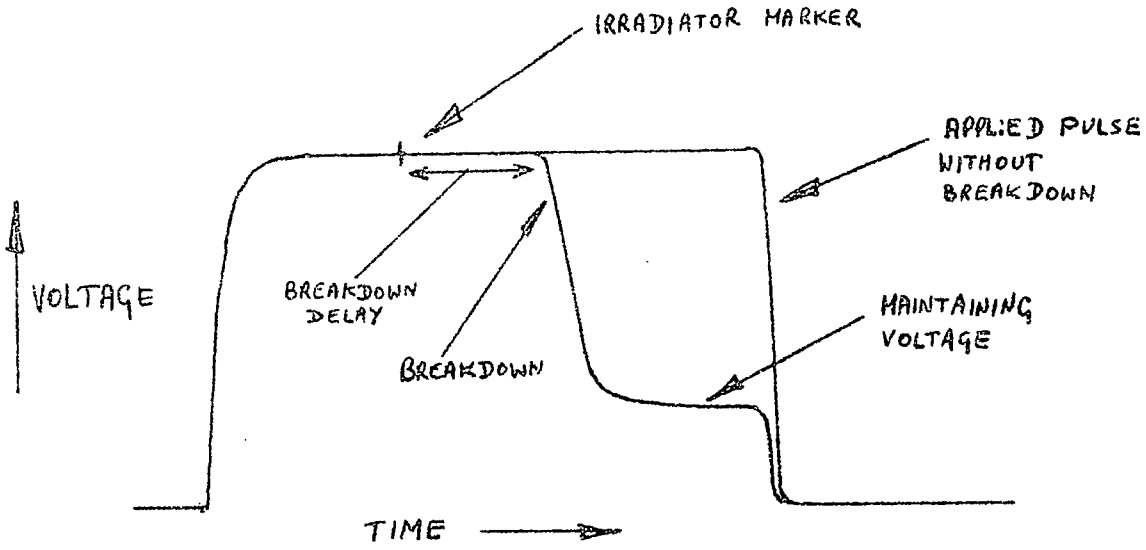


FIG 2.4 IDEALISED GAS BREAKDOWN OSCILLOGRAMS

2.3. Electronic timing control

Two functions were required of the apparatus: measurements were required firstly of breakdown fields under sustained oscillations, and secondly of the time delays to breakdown, for which the repeated application of a pulsed field was necessary.

Continuous oscillations were obtained by applying a steady voltage to the anodes of the oscillator, for which the source of supply was a d.c. generator driven by a mains electric motor. Fine adjustments of the output voltage could be made by varying the direct current through the field coils. Up to 1000 volts could be delivered from this supply with a load current of 0.2 amps.

Arrangements for supplying pulsed fields to the spark gap were considerably more complicated. The block diagram of the original circuitry is shown in fig 2.3. The anode supply for the oscillator in the pulsed field configuration was required at a lower mean current level, and at higher voltage than that used for production of continuous oscillations. Therefore a transformer and half-wave rectifier system was used to charge a bank of capacitors which provided the immediate current supply for the oscillator, during a pulse.

A voltage comparator circuit produced an output pulse when the capacitor bank reached a predetermined potential, and

this was used to trigger the "main impulse generator". The four sequential functions of the circuit were as follows: firstly the oscilloscope timebase was triggered, and then the oscillator was energised, putting a high r.f. voltage across the spark gap. After a predetermined delay the irradiator was fired to initiate a discharge, and the supply voltage was removed from the oscillator a little later. In order to see the instant of irradiation on the oscilloscope a few inches of unscreened wire were placed near the irradiator discharge and connected to the Y input terminals and this picked up enough voltage to deflect the trace momentarily. Fig 2.4 shows the types of oscillogram which were obtained, both with and without breakdown occurring during the pulse.

An assessment of the existing control circuitry indicated that it would not be practicable to obtain the desired accuracy and stability, although the basic functions were correct. It was therefore decided to completely redesign the electronic circuitry. The same basic functions were necessary but careful attention was given to the stability of the EHT voltage, and also greater repeatability of the timing of the irradiator pulse and the oscilloscope trigger were required to allow the use of faster sweep speeds. A description of the modifications which were made forms the subject of the next two chapters.

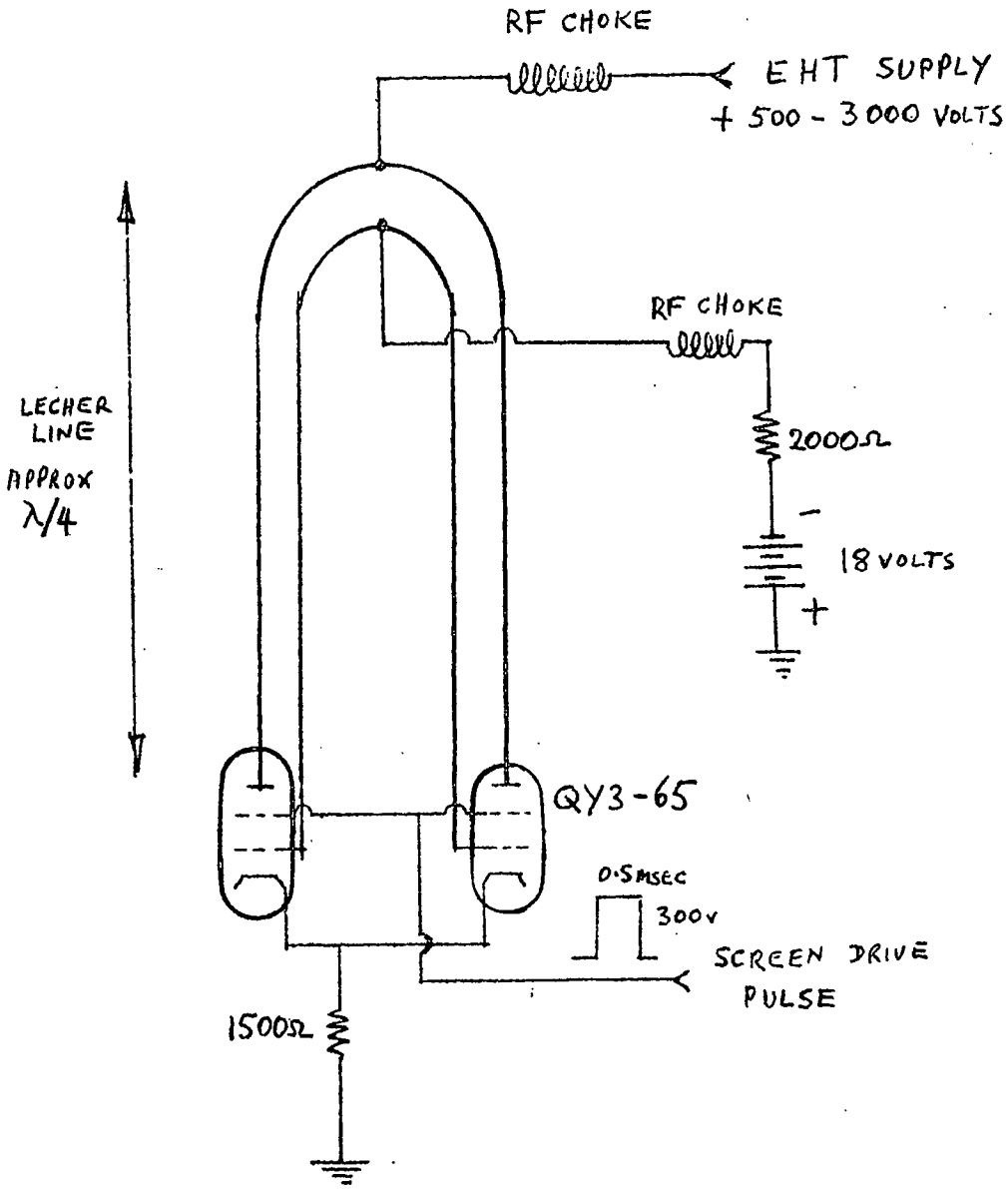


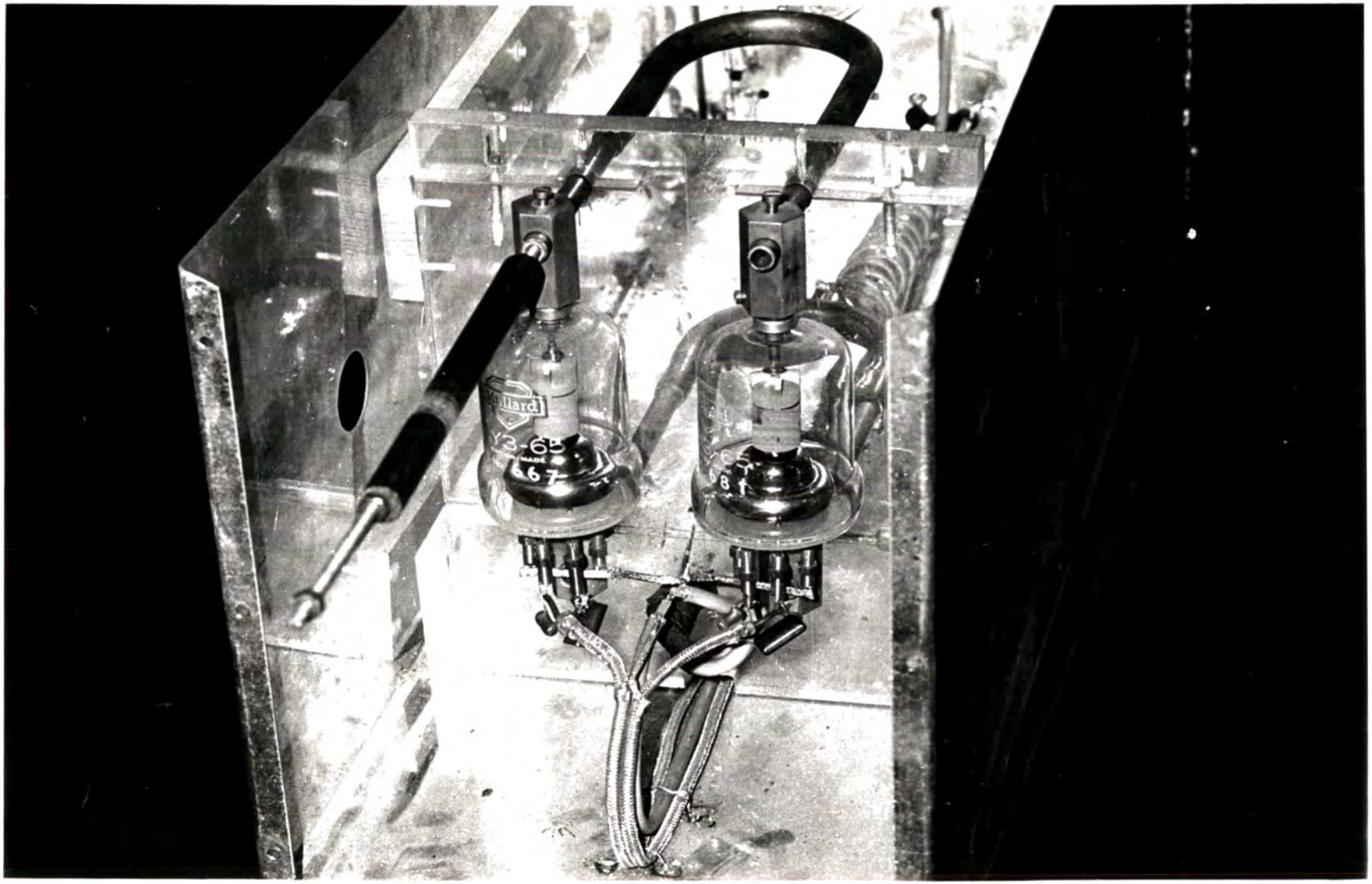
FIG 3.1 CIRCUIT DIAGRAM OF VHF OSCILLATOR

CHAPTER III

THE OSCILLATOR

The apparatus already in existence in this laboratory, used an oscillator from an ex-W.D. v.h.f transmitter. For tuning purposes there was a sliding short circuit bar on lecher lines attached to anodes and grids of two triodes oscillating in push-pull. Time and long use had corroded these lecher lines so that no amount of cleaning would give a steady output to the spark gap. Moreover, it was intended to extend the measurements of Prowse, Rowbotham and Monk[9] into the region of much shorter delays, so that more power would be required. A new oscillator was therefore built to operate at 180 Mc/s using larger v.h.f power tetrodes in a push-pull tuned-anode-tuned-grid circuit fig. 3.1. In this both anode and grid tuned circuits were of the lecher line type, constructed from half inch bore copper pipe. They were tuneable by means of sliding "trombone" loops. In this case, unlike that of the old oscillator, the sliding contacts were well removed from the current maximum, and no trouble was experienced due to sparking at these points.

The oscillator valves were of type QY3-65, which have special bases, and the anode connection emerges from the top of the glass envelope. Special sleeves were made which clamped tightly onto the base pins, for soldering to the d.c.


Mullard
V3-6B
6T7
6T7

supply leads, thus avoiding possible loose contacts in a conventional base. The anode and grid lecher lines were connected to solid brass posts which were drilled to exactly the right diameter to take the respective pins. The only mechanical support for the oscillator assembly was a square piece of polystyrene sheet, half an inch thick, through which the lecher lines passed in closely fitting holes. The anode line was parallel to the grid line, and directly above it, so that the valves could be hung on the ends of the lines by their anode and grid pins, as can be seen in fig 3.2.

This assembly was mounted in an open-topped tin-plate box which acted as a partial screen thus preventing proximity detuning effects. All the joints in the box were carefully soldered, to avoid parasitic discharges.

The modulation of the oscillator was effected originally by means of a thyratron in series with the e.h.t supply to the anodes. This system was inconvenient for two reasons: firstly the oscillations had to be terminated by an independent quenching circuit, and secondly a mechanical relay was required to open the anode circuit of the thyratron to allow it to de-ionize. The mechanical relay was troublesome in practice as high voltages were being switched. A modulator was therefore devised using hard valves, to overcome these defects in the original system. This is shown in fig 3.3. This circuit operates in the fashion of a "bootstrap" time-

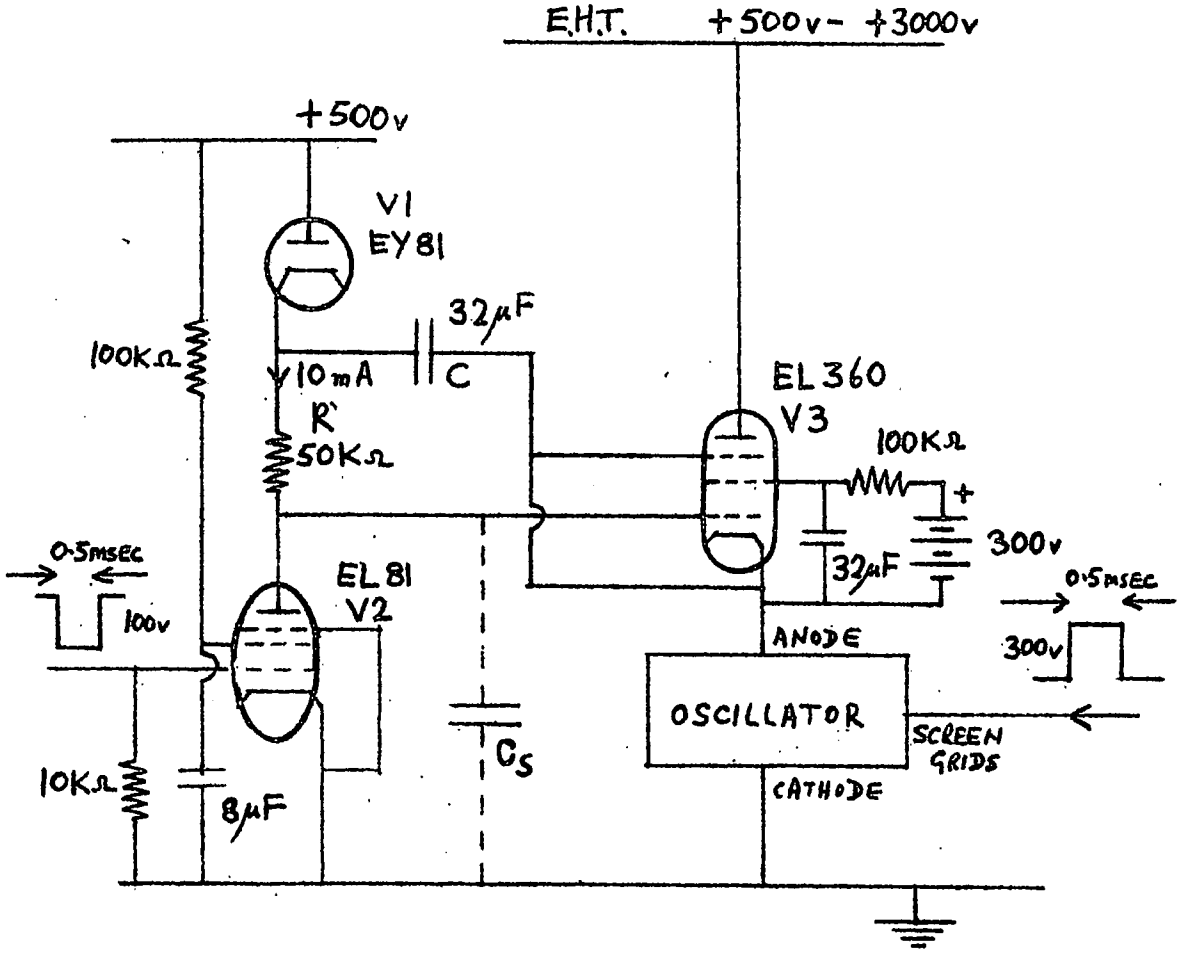


Fig 3.3 MODULATOR CIRCUIT FOR OSCILLATOR

base[15], the timing capacitor C_s in this case being reduced to stray capacitance. The positive feedback capacitor C linearises the rise of potential at the oscillator anodes, and allows the use of a comparatively low value of H.T. on V_2 anode. But for the bootstrap connection, the resistor R would have to be taken to full e.h.t. with consequent need for a large current e.h.t. supply, because V_2 is normally conducting, in the bottomed state.

To avoid excessive dissipation at the oscillator screen grids, between pulses, a positive potential of magnitude 300 volts was applied to these electrodes only for the duration of the pulse.

The modulator circuit operated well but it was later found that just as satisfactory pulsed operation could be obtained by applying the e.h.t. continuously to the oscillator anodes, and a positive switching pulse to the screen grids as before. The only modification to the oscillator then necessary was the inclusion in the grid circuit of a negative bias, to prevent the flow of current from the e.h.t. supply in the interval between pulses.

Occasionally, when the anode voltage was very high, the oscillations did not terminate at the end of the modulator pulse, but continued at reduced amplitude until the e.h.t. reservoir capacitor was discharged considerably. This was

because the screen grids were fed from a cathode follower: although such a circuit has a very low output impedance, (about 50 ohms), some electrostatic coupling between anode and grid still existed and the valves could oscillate as triodes. These oscillations did not affect the measurements in any way because they only occurred after the end of the pulse.

In operation the oscillator anodes, and the anode lecher line became heated to about 100°C. This caused an increase in the length of the line and therefore a decrease in frequency of 0.2%. Provided time was allowed for warming up frequency variation was not a source of trouble, despite the sharp resonance (Q = 1000 approx) of the transmission line. To compensate for long term fluctuations in frequency, a threaded brass rod was inserted coaxially into one end of the anode line, to allow a fine tuning adjustment to be made. A total change in frequency of about 1% could be made with this screw, and this was sufficient to enable it to be used to measure the Q of the transmission line, after calibration. Q was determined by plotting a resonance curve giving the value 1200±100.

Fig 3.4. shows the variation of the oscillator frequency with the fine tuning control: while in fig 3.5. the frequency is given as a function of the length of the anode lecher line

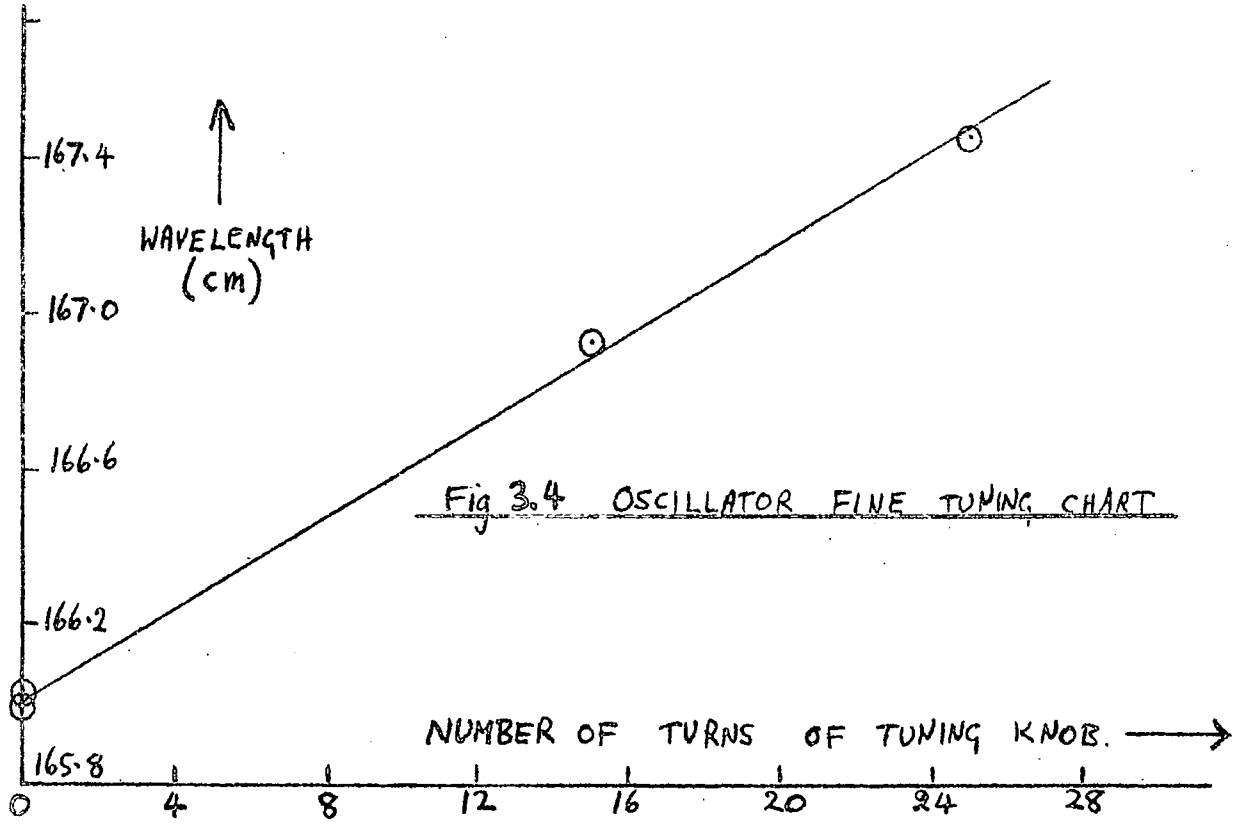


Fig 3.4 OSCILLATOR FINE TUNING CHART

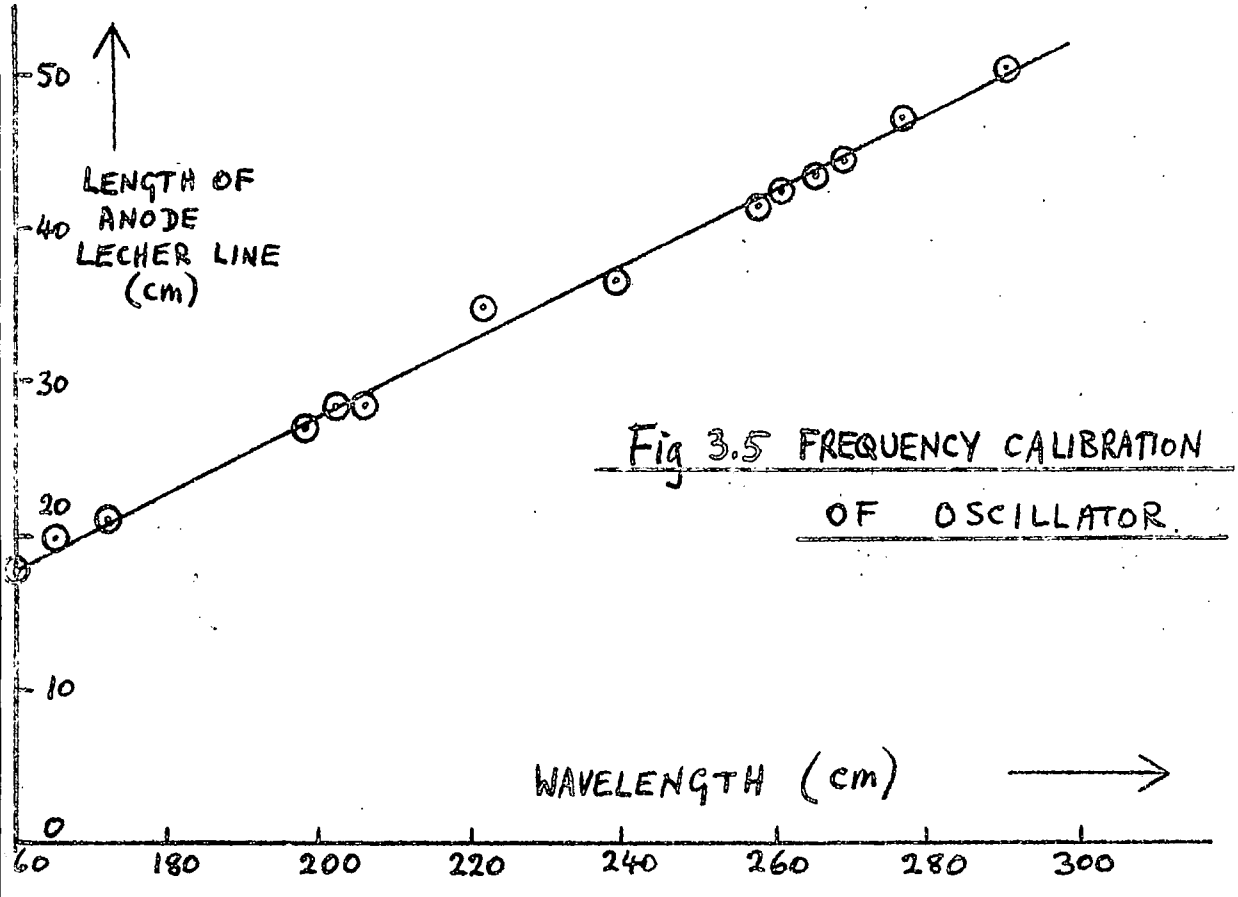


Fig 3.5 FREQUENCY CALIBRATION OF OSCILLATOR.

on a coarser scale. The frequency was measured with a pair of lecher lines about five wavelengths long. About 0.1% accuracy could be obtained by finding the successive positions of a shorting bar giving a minimum reading on a galvanometer which was connected between the lines.

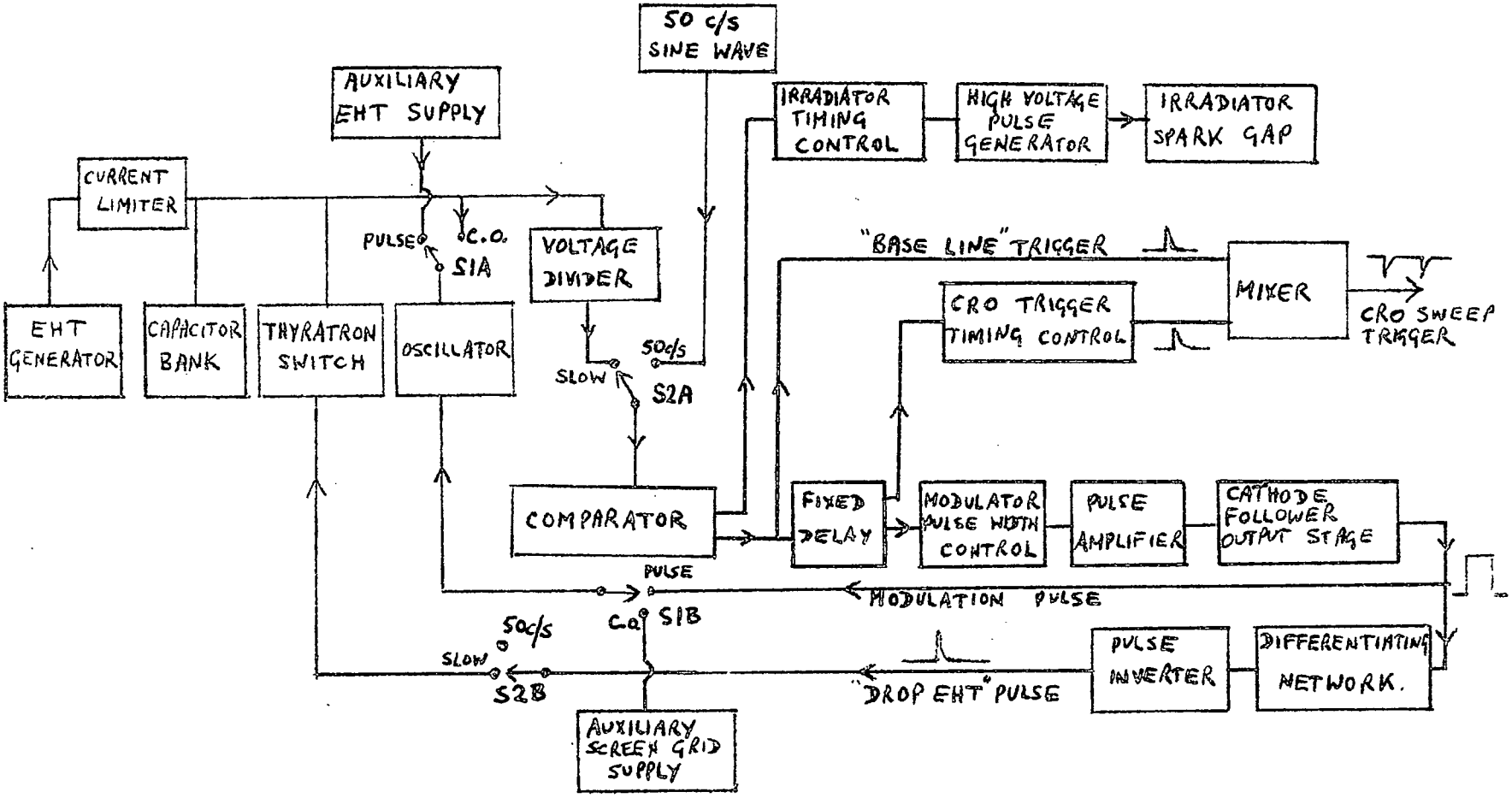


FIG 4.1 BLOCK DIAGRAM OF THE SEQUENCING CONTROL CIRCUITRY

CHAPTER IV

SEQUENCING CIRCUITRY

4.1. A functional description

The description which follows, of the circuitry, may be followed initially by reference to the block diagram opposite; fig 4.1, and later to the circuit diagrams figs 4.3, 4.4, 4.5.

To provide a record of the time delay between the irradiator pulse and the occurrence of breakdown in the spark gap, a photograph of the oscilloscope trace, was required. For this purpose, the switches S1A and S1B were set to the "pulse" position (fig 4.1) while S2A and S2B were set to "slow".

The circuits all remained in their quiescent state until the e.h.t. supply was switched on. The main capacitor bank then charged on a time constant of approximately 4 seconds. The voltage appearing across this capacitor also appeared across the oscillator, a mercury thyratron, and the comparator circuit (via a voltage divider). The function of the comparator was to provide an output pulse at the instant a certain fraction of the e.h.t. line voltage became equal to that across a neon voltage-regulator tube. This output pulse initiated all the succeeding events.

Firstly, the "fixed delay" circuit was triggered by the comparator output to produce a positive-going rectangular

pulse. The "modulator pulse-width control" circuit was triggered from the final edge of this pulse, to produce a square pulse of variable duration. This variable-width pulse, after being "squared" by a saturating amplifier stage, was fed into a cathode follower. The cathode follower output was a positive pulse of magnitude 300 volts and of sufficiently low source impedance for modulating the oscillator valves on their screen grids.

The square modulation pulse was also fed via a differentiating network to the control grid of a pentode. This resulted in a negative "pip" followed by a positive "pip" at the pentode anode, coinciding with the leading and trailing edges of the modulation pulse, respectively. These were taken to the control grid of the thyratron which therefore fired at the termination of the modulation pulse. The firing of the thyratron virtually short circuited the e.h.t. line, so that its voltage fell quickly to a level at which the comparator circuit reset itself. As soon as deionization had taken place, the capacitor bank began to recharge to repeat the above action indefinitely.

A second output from the comparator output was used to trigger the "irradiator timing control" circuit. This circuit provided a high voltage, short duration pulse to the irradiator spark gap, after an interval which could be adjusted to make irradiation occur at any given instant

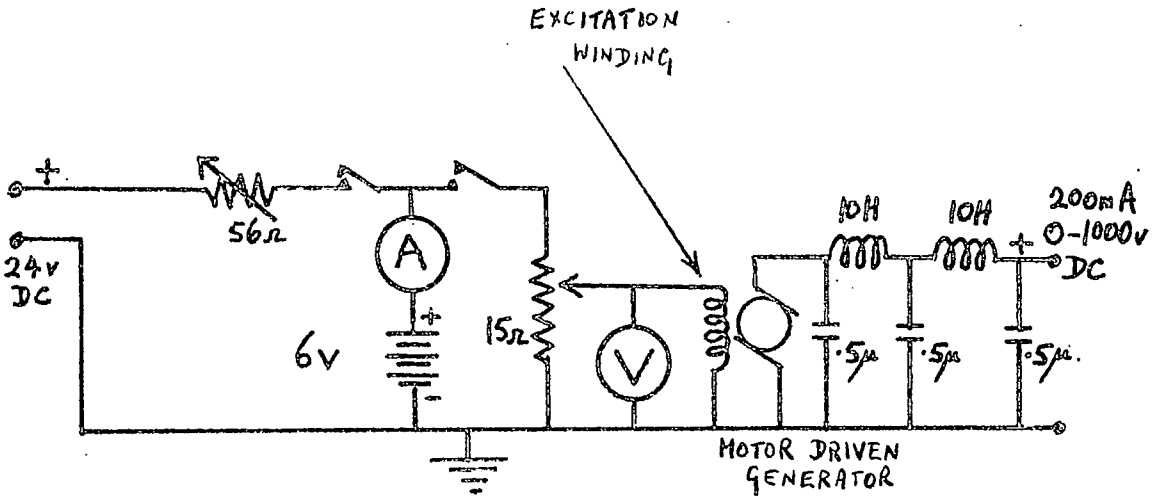


Fig.4.2 AUXILIARY EHT SUPPLY

during the oscillator pulse.

A variable delay circuit similar to that in the irradiator circuit was triggered from the tail end of the fixed delay pulse, the output of which was used to trigger the oscilloscope timebase at a predetermined instant during the oscillator pulse. Since on fast timebase sweeps, (less than about 300 microseconds duration) only the top of the pulse would therefore be shown on the oscilloscope screen, a second sweep was arranged to occur before the oscillator was switched on, (i.e. during the "fixed delay" interval) to give a base-line to the pulses. This second sweep was triggered by the comparator output pulse, and to avoid undesirable feedback between the two circuits, the two sweep trigger pulses were combined by a mixer circuit. If the timebase lasted more than about 300 microseconds, the second trigger pulse arrived during the first sweep, and therefore had no effect. The resulting picture on the screen was then of the beginning of the oscillator pulse preceeded by zero output for the duration of the fixed delay (150 microseconds). When the sweep was of duration greater than 1 millisecond the whole of the oscillator pulse (400 microseconds) appeared on the screen.

For measurements of breakdown fields with very long ("infinite") delays, the auxiliary e.h.t. and screen grid

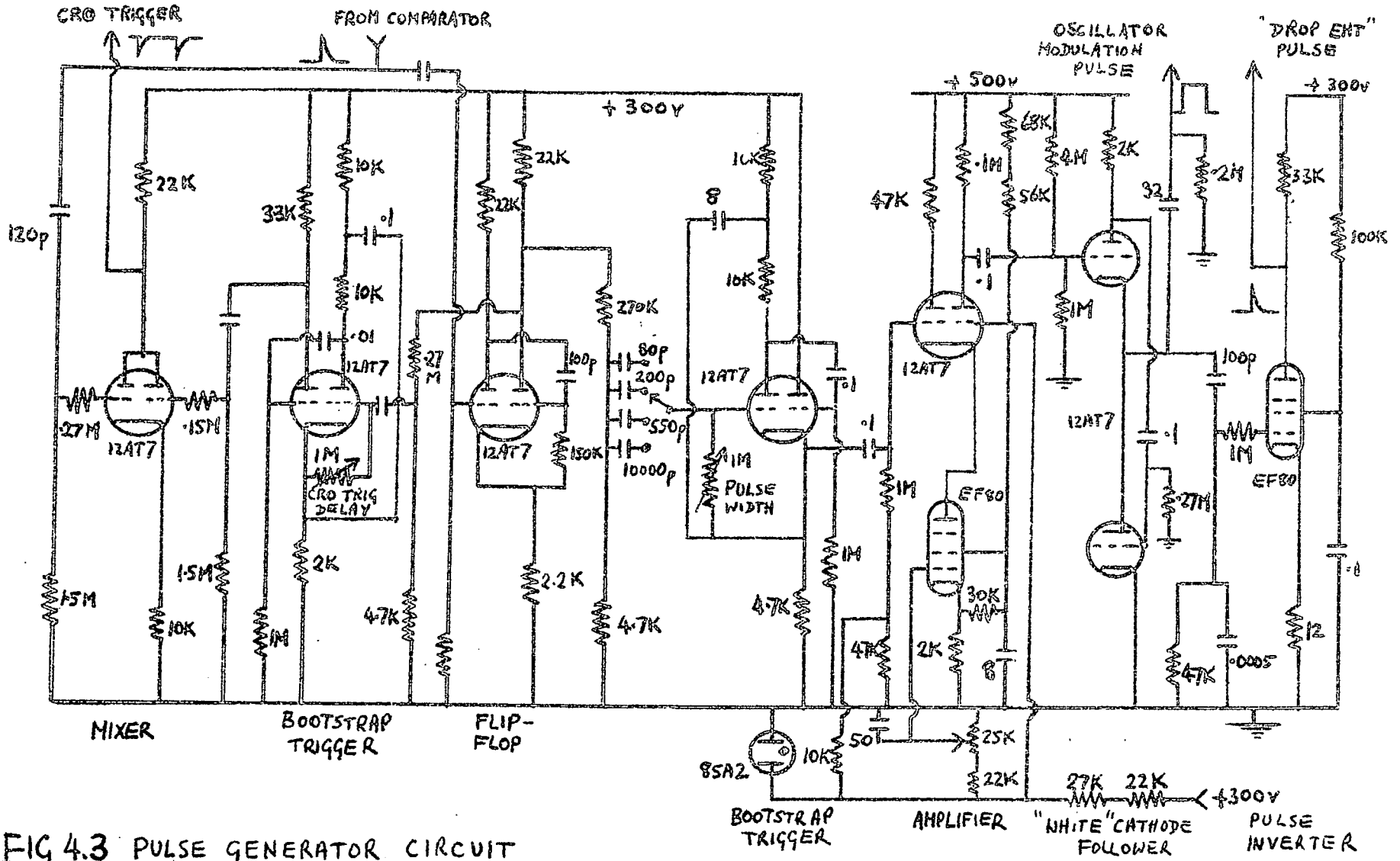


FIG 4.3 PULSE GENERATOR CIRCUIT

supplies were used, to run the oscillator continuously. Fig 4.2 shows the circuitry of the auxiliary e.h.t. generator. Switches S1A and S1B on fig 4.1 were moved to "continuous oscillations" (C.O.), and switch S2A was put to "50 c/s", to fire the irradiator 50 times per second, ensuring an adequate supply of electrons in the test gap to initiate breakdown. This arrangement was also used for making an absolute calibration of the CRO voltage in terms of the spark gap voltage, but in this case, of course, it would not be necessary to operate the irradiator.

The "50 c/s" function of switch S2A was of great value in setting-up and testing the pulse circuitry, in that a visually continuous picture of the waveform at any point could be displayed on a monitoring oscilloscope. The monitor could be triggered from the same point as the measuring oscilloscope, at 50 c/s, to give a perfectly locked display.

For the most part the circuitry consisted of conventional trigger circuits, namely monostable "bootstrap" and "flip-flop" square wave generators. The choice between these two was governed by the available trigger pulse, and the desired output polarity. The bootstrap circuit is more suitable for negative, and the flip-flop for positive, trigger pulses. Output pulses were obtained from the trigger circuits by bearing in mind the general rule that it is desirable to take

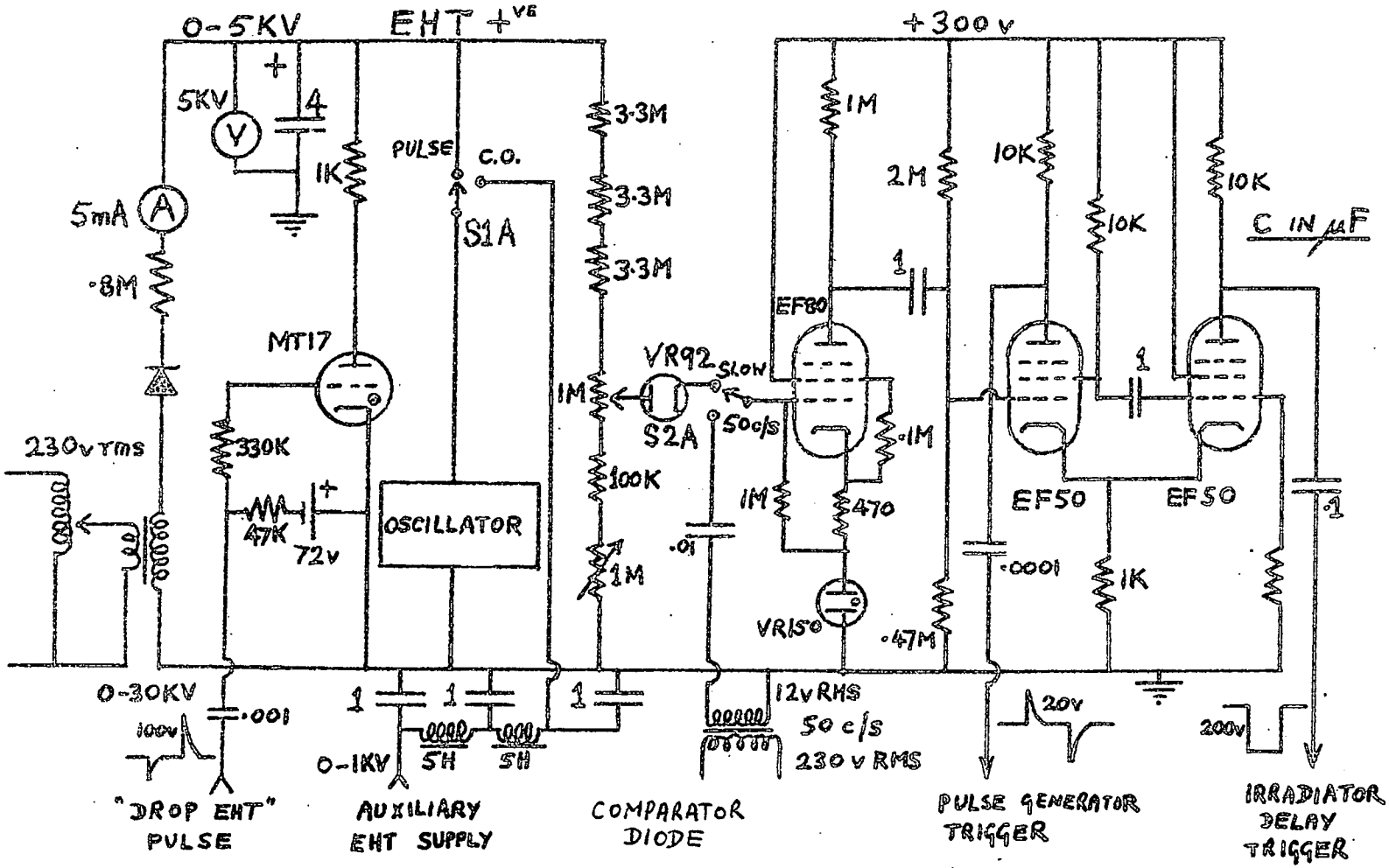


Fig 4.4 EHT AND COMPARATOR CIRCUIT

inter-circuit connections to or from electrodes which play no part in the regenerative feedback. When two pentodes are used, their screen grids can be used effectively as anodes for the internal regenerative output, leaving a "free" anode to deliver the external output pulse. In this case the flip-flop circuit can be designed for use with either polarity trigger.

4.2. The comparator circuit.

The action of the comparator circuit fig 4.4 was as follows: as the e.h.t. voltage rose on a time constant of 4 seconds, towards a potential of the order 5KV, the potential on the anode of the VR 92 diode rose, typically, at a rate of 150 volts/second. After one second, the potential at this point became comparable to the voltage across the regulator tube, enabling the pentode to conduct. The pentode had a large load resistor, and therefore a high gain, and served to increase the rate of rise of the voltage applied to the following Schmidt trigger circuit. In this way the output risetime of the Schmidt trigger was improved, and time jitter in the dependent circuits was reduced.

4.3. The Irradiator Circuit.

The irradiator circuit fig 4.5 was triggered by the leading edge of the pulse from the comparator. The first stage of the circuit was a bootstrap trigger whose function

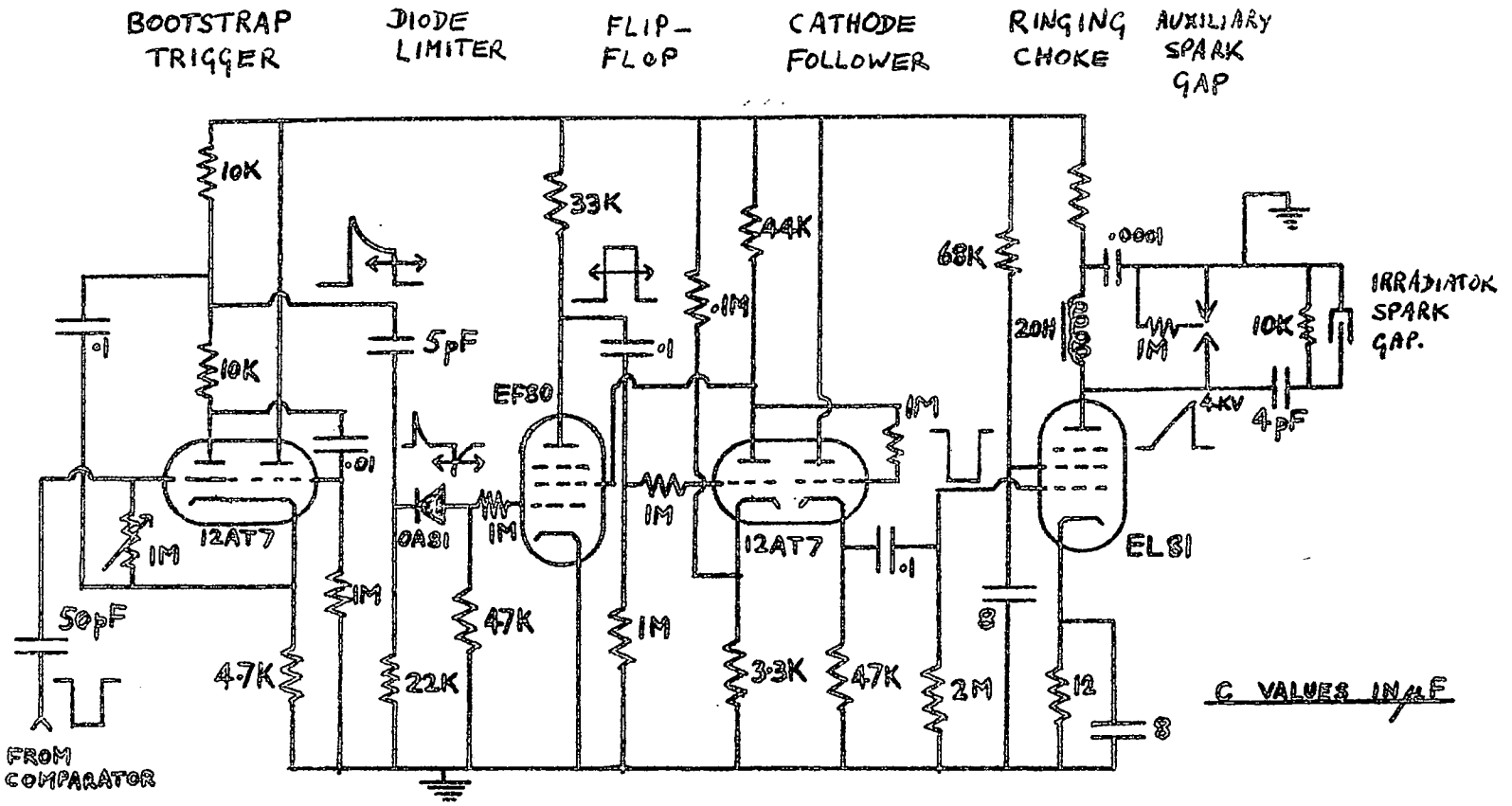


FIG. 4.5 IRRADIATOR DELAY AND OUTPUT CIRCUITS

was to provide a positive output pulse of variable duration. The final negative-going edge of this pulse was then used to trigger a "flip-flop" circuit, consisting of a pentode and one half of a double-triode valve. The screen grid of this pentode was here used as a second control grid for applying the regenerative feedback. An output pulse was taken from this electrode which was directly coupled to the grid of a cathode follower. The cathode follower output pulse, drove the grid of the final valve (EL81) about 50 volts negative, to turn if off in about 5 microseconds. With the sudden cessation of current in the 20 henry choke which formed the anode load, the anode potential started to rise sinusoidally towards about 10,000 volts. Before this high voltage was achieved, however, a spark occurred in a short air gap which was placed effectively across the terminals of the choke. The rapid fall of voltage across this gap was transformed, by use of a differentiating network, into a short negative-going pulse, which, was of sufficient amplitude to cause a discharge in the irradiator gap.

The repeatability of this sparking voltage was optimised by making the electrodes sharply pointed, and by use of a "keep-alive" electrode. The keep-alive electrode caused some ionization in the gap from a corona discharge as the gap voltage built up, thus minimising the statistical

variations, or jitter of the time at which the spark occurred. The keep-alive electrode was made from fine platinum wire which was earthed at one end and placed with its pointed end near the high potential electrode of the auxiliary spark gap.

The junction diode, OA81, was essential to the satisfactory operation of this circuit: it will be noted that although the flip-flop only triggers in response to a negative pulse on its input grid, a positive pulse at this point would be inverted and amplified at the screen grid and fed to the grid of the EL81. This would cause a premature output pulse to the irradiator.

All the valves of the irradiator circuit were mounted on a single chassis and some difficulty was experienced at first in preventing multiple triggering. The trouble, which was found to be caused by radiation from the anode of the EL81, was only cured by the introduction both of the cathode follower stage and of double screening above the chassis between this valve and the others.

CHAPTER V

MEASUREMENT OF ABSOLUTE VOLTAGE AT 180 Mc/s.

The method of measurement described in this chapter was designed for the original work the results of which were described in detail in reference 9 and outlined in Chapter II. A detailed discussion of the method here is justified by the vital nature of this measurement to the whole investigation. Moreover, the quest for improved risetime of the output voltage has led to difficulties which were not previously encountered with this apparatus.

As was seen in Chapter II, the final display of the spark gap voltage was made on the screen of an oscilloscope for which purpose the envelope of the r.f. pulse was obtained, by half-wave rectification. It will be shown in this chapter how the amplitude of this pulse was related to the absolute voltage at the spark gap.

5.1. The dielectric vane electrometer and calibration of the oscilloscope.

The calculation of the output voltage from the geometry of the piston attenuator and circuit constants would require a knowledge of (a) the absolute coupling factor of the launching disc into the piston attenuator, and (b) the rectification properties of the diode. Neither of these could be reliably calculated, and therefore an absolute calibration was carried out.

Calorimetric methods could not be used because of their inherent power dissipation: the voltage to be measured in this case was across a high quality resonant circuit, and would have been lowered prohibitively by the provision of power to a measuring instrument.

Thermionic instruments depart from their low frequency calibration at high frequencies due to skin effect and the effects of lead inductance and transit time. Electrostatic methods are well suited to this application: the presence of a small capacitative load has only a detuning effect which can easily be compensated. The direct connection of the Y plates of a cathode ray tube to the spark gap might be one solution. This method was not used however, because of the difficulty of obtaining a clear trace at the high writing speeds involved. Furthermore an attenuator would be needed, if a wide range of voltages were to be used, so that again a high frequency calibration would be required.

Calibration was in fact carried out with a device which was used by Prowse Rowbotham and Monk[8] and which superficially resembles a quadrant electrometer. The electrometer was built into the coaxial line, at a point close to the spark gap, as is shown in fig. 2.1. Figs 5.1(a) and (b) show the device in more detail.

The central conductor of the resonant coaxial line was terminated in a brass cube, of side 2.5cm. Upon the upper

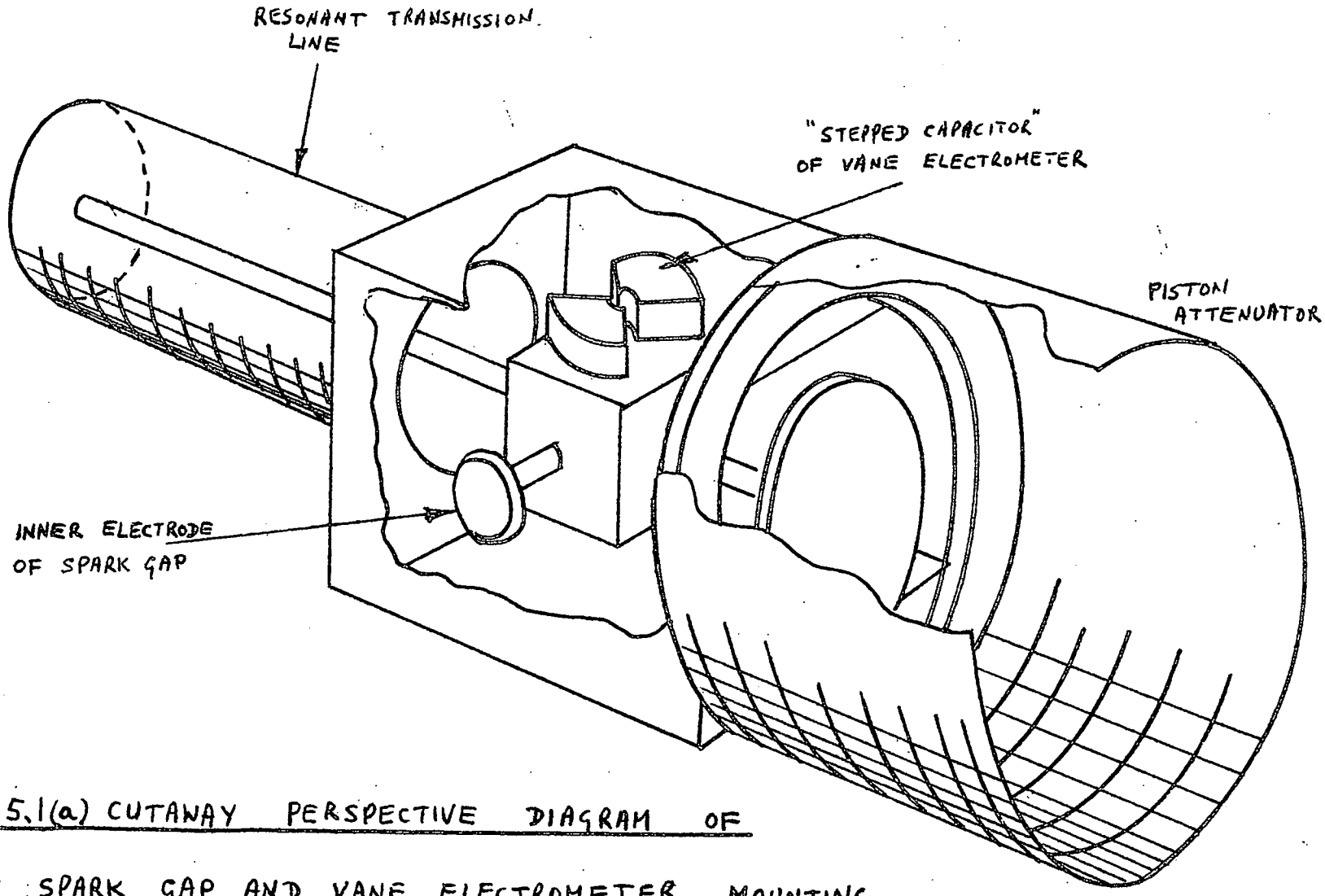
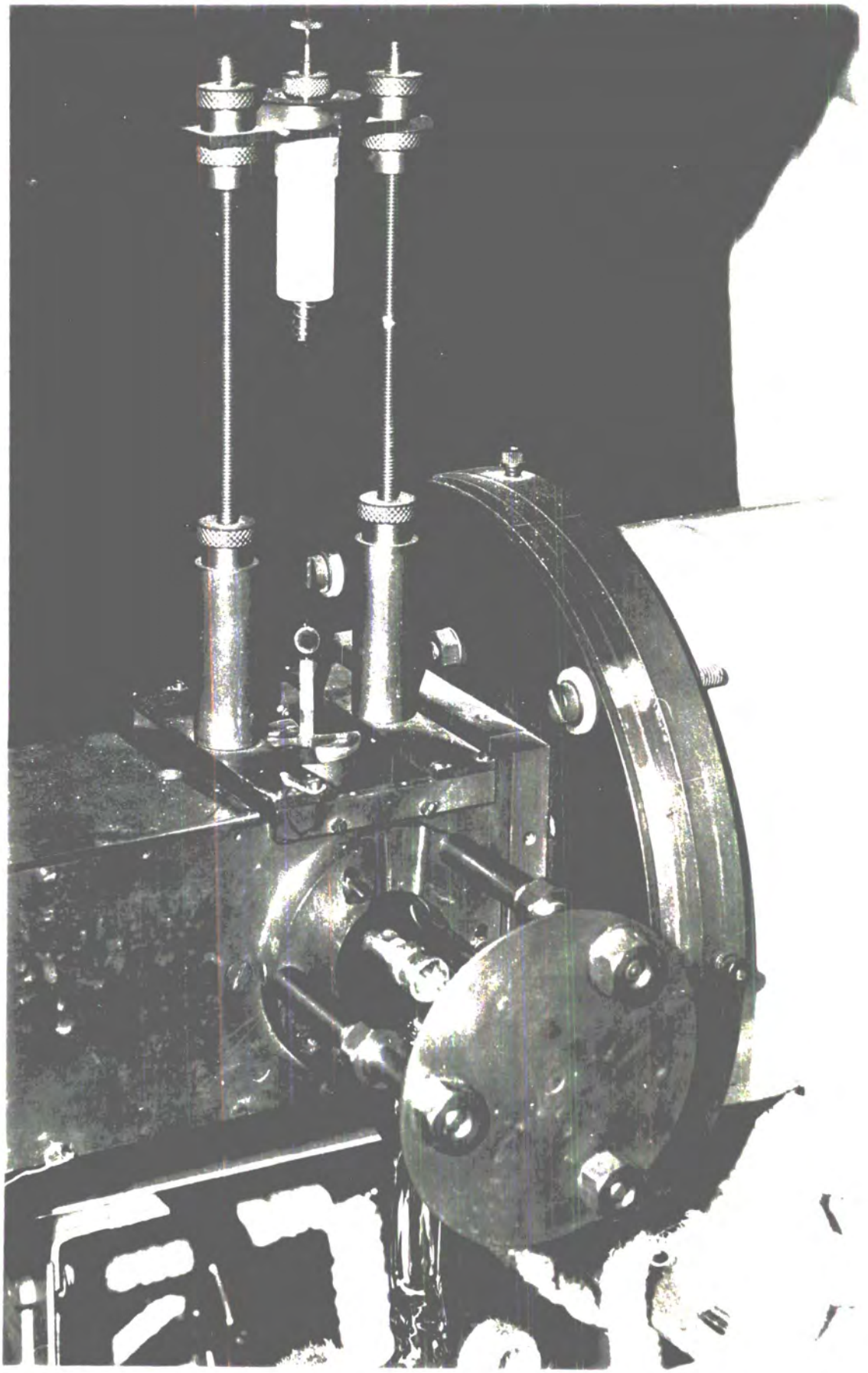


FIG 5.1(a) CUTAWAY PERSPECTIVE DIAGRAM OF THE SPARK GAP AND VANE ELECTROMETER MOUNTING



face of this cube were attached two flat slabs of brass, 0.2cm thick, which were shaped in plan like the two opposite quadrants of a circle, 2.5cm in diameter. A sheet of polystyrene also forming two quadrants of a circle was suspended horizontally 0.5mm above this, by means of a quartz fibre. The fibre was connected to one end of a thin polystyrene pillar which emerged perpendicularly from the centre of gravity of the vane. A sheet of brass covered the vane, having a small hole through which the pillar passed. The cover formed the upper face of a hollow box which was the expanded outer conductor of the coaxial line. The polystyrene vane was adjusted so that with zero voltage across the coaxial line, it approximately half covered the brass quadrants, and could swing freely on its suspension.

Viewed from a horizontal direction the vane could be seen to form part of the dielectric of a capacitor with a stepped plate separation. A voltage across the plates of this capacitor, caused the vane to move in a direction which increased the capacitance. The force acting on the vane was independent of the direction of the applied voltage, and a rotational deflection was obtained for alternating voltages which was proportional to the square of the applied voltage.

A galvanometer mirror attached to the vane support pillar allowed the angle of rotation to be measured.

The permittivity of polystyrene is sensibly constant from

zero frequency up to several thousand megacycles per second, and since it is this quantity which determined the force acting on the vane, an absolute calibration could be carried out at d.c. or power frequencies using a standard voltmeter. This calibration could then be used with confidence at 180 Mc/s.

One electrode of the spark gap was connected to the brass cube, by a copper lead about 3cm long, and the other electrode was connected to the outer conductor of the coaxial line opposite this point. The operating wavelength, 170cm was much larger than this dimension so that the voltage measured by the vane could be assumed to be identical to that existing across the spark gap.

Calibration of the diode-voltmeter/oscilloscope system was carried out using a sustained output from the oscillator. The resulting constant r.m.s. voltage at the spark gap produced a steady deflection of the dielectric vane, and also of the oscilloscope trace. Since the sensitivity of the vane could readily be measured in terms of absolute voltage, using a low frequency standard voltmeter, it followed that the constant of proportionality between the absolute voltage at the spark gap and the deflection of the oscilloscope trace could be determined. It was then assumed that deflections of the oscilloscope trace in response to pulsed modulation of the oscillator would bear the same relation to the spark gap voltage.

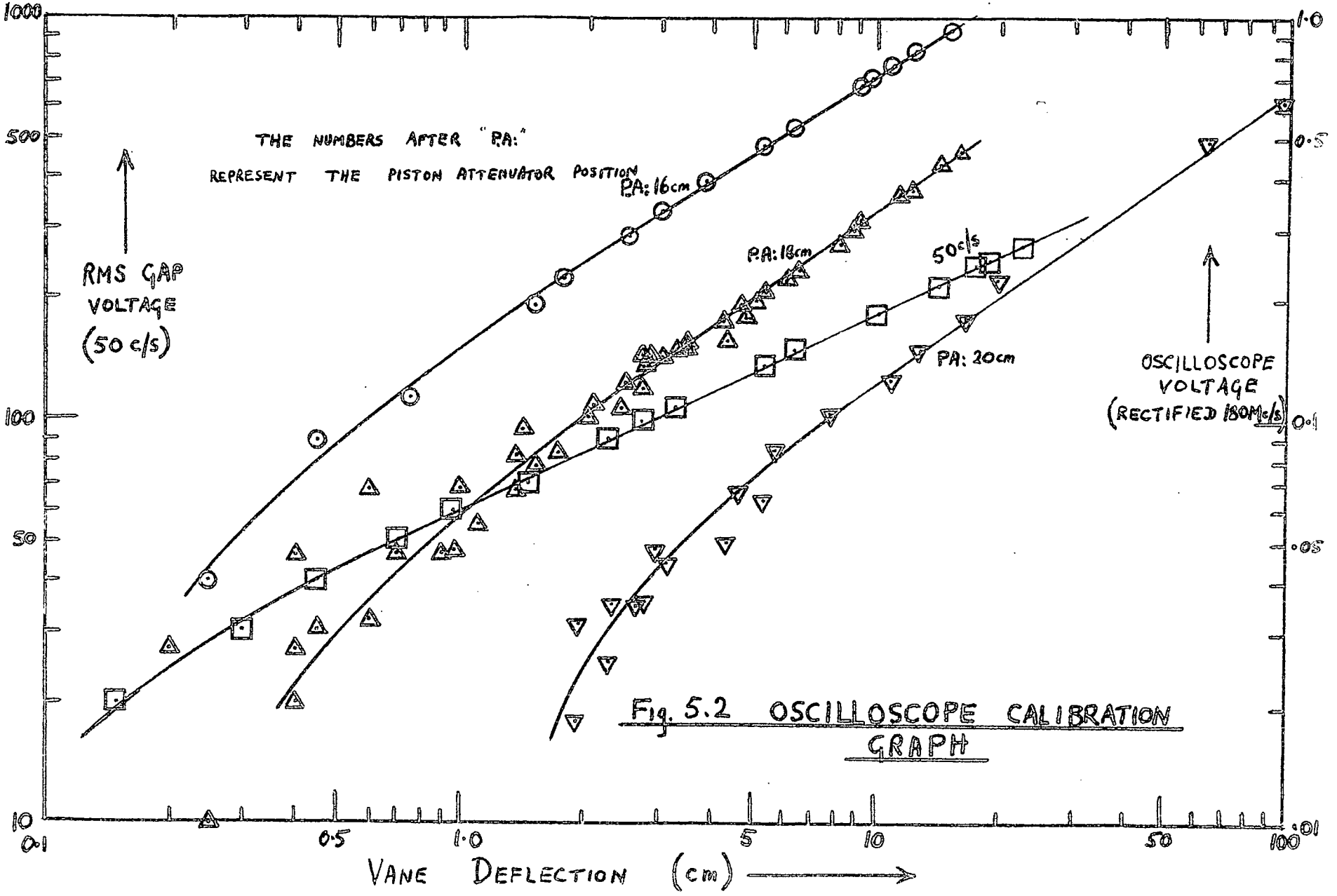


Fig. 5.2 OSCILLOSCOPE CALIBRATION GRAPH

The calibration graph for determination of spark gap voltages from oscilloscope deflection is shown in fig. 5.2 at three settings of the piston attenuator.

5.2. Performance and accuracy of the voltage measurement system.

5.2.1. Dielectric vane electrometer.

The original device as used in the experiments described by Prowse, Rowbotham and Monk,[9] had a silk bifilar suspension. This was used for a while in the present experiments, but proved rather unsatisfactory because of the change in the zero position after each deflection. The cause of this was probably the re-arrangement of individual strands in the fibre when twisted. A single quartz fibre was substituted and an improvement in stability by about a factor two was obtained.

The absolute calibration was at first tried with direct voltages. After the application of a direct voltage, however, a slow drift in the deflection was observed, superimposed upon the steady deflection, tending asymptotically towards a limit, somewhat higher than the instantaneous value, with a time constant of several minutes. A similar drift was observed tending towards the original zero when the voltage was removed. It appeared that the explanation of this phenomenon lay in terms of polarization of the dielectric. Since the time constant was so long it was anticipated that

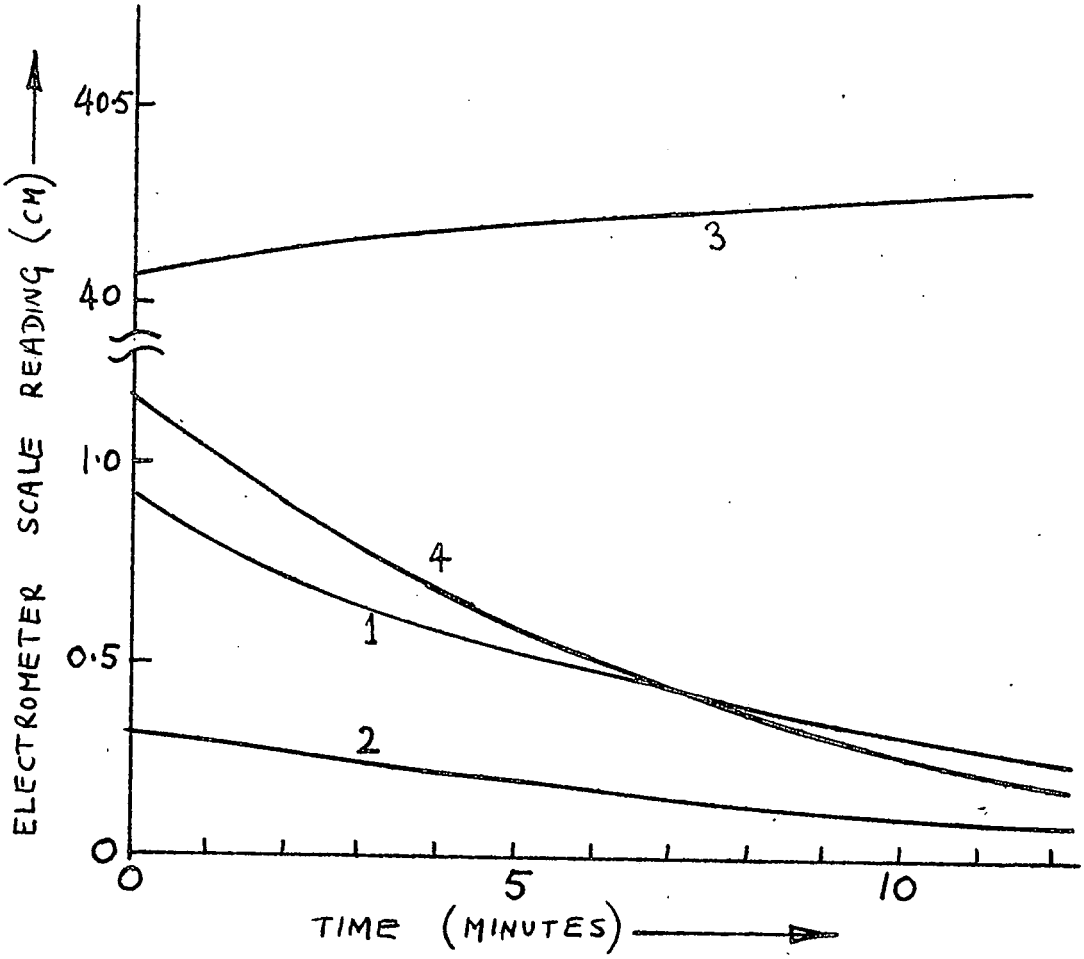


FIG 5.3 DRIFT OF THE VANE ELECTROMETER WITH TIME

CURVE 1 shows the downward drift towards zero after 100 volts had been applied continuously for 16 hours and then removed.

CURVE 2 as for curve 1 except that the voltage was applied for 1 minute.

CURVE 3 shows the approach to a steady deflection with 100 volts applied after 17 minutes at zero.

CURVE 4 shows the decay towards zero after 100 volts deflection for 13 minutes.

calibration using the a.c. mains supply would eliminate difficulty. However, the improvement achieved by this means was not more than marginal. An explanation of the drift might lie in the quality of the anchoring points: the quartz fibre was fastened at each end by means of hard vacuum-wax, which has a pitch-like consistency, i.e. brittle and very viscous at room temperature. Therefore a shear stress imposed by the twist of the fibre could have been responsible for a slow drift. However, the vane position would then have increased indefinitely with time, instead of approaching asymptotically towards a limit, as was observed, and no zero drift should then have been observed on removing the voltage. A possible explanation is that a thin coating of wax covered part of the fibre so that viscous flow caused a change in the effective fibre length with time, but care was taken to avoid this. The explanation may possibly involve the quality of the suspension; however, quartz is generally accepted to be the best dielectic material for this purpose, and certainly a pyrex suspension proved to be considerably inferior in this respect. The amount of drift which finally had to be accepted is shown in fig. 5.3.

It was found that by increasing the voltage across the vane carefully, much transient oscillation could be eliminated and readings could be taken within about ten seconds. Not much drift could occur in this short time and the voltage was

reduced to zero immediately (but carefully, to avoid oscillations of the vane) after each reading. By taking such precautions, it was possible to obtain results reproducible within 1% after a week or more.

5.2.2 The piston attenuator

The ratio of the voltages V_1 and V_2 at the distances z_1 and z_2 from an arbitrary origin along the axis of the attenuator can be shown[12] to be given by the relation

$$\log_{10}(V_1/V_2) = (1.045/r)(1-\lambda_c^2/\lambda^2)^{\frac{1}{2}}(z_1-z_2) \quad \ldots\ldots \quad 5.1$$

where r is the radius of the cylinder

λ is the actual wavelength in free space

λ_c is the cutoff wavelength for the cylinder.

This formula applies to the E_{01} mode when the pickup electrode is sufficiently far removed from the launching disc to prevent an appreciable amount of energy being reflected back to the launching disc. For this mode the cutoff wavelength is given by

$$\lambda_c = 2r/2.405$$

$$= 20.2\text{cm in this case.}$$

The factor $(1-\lambda_c^2/\lambda^2)^{\frac{1}{2}}$ in equation 5.1 was therefore equal to 0.992 so that the attenuation was very nearly independent of frequency at 180 Mc/s.

Therefore the attenuation factor, defined as

$$\alpha = \log_{10}(V_1/V_2)/(z_1-z_2) \quad \ldots\ldots \quad 5.2.$$

was 0.133 cm^{-1}.

In an evanescent mode, the E and H vectors representing the wave oscillate in quadrature, so that it is, at first glance, surprising that power can be propagated by such a field. It can be shown, however,[12] that the presence of the pickup electrode distorts the field so as to produce a small in-phase component, with consequent transfer of sufficient power to the output circuit to be displayed on the oscilloscope.

The attenuation factor was determined experimentally by measuring the voltages at the spark gap which produced a given deflection of the CRO trace for different settings of the piston attenuator. This can in fact be carried out using the calibration graph, fig. 5.2: for example a given value of CRO deflection, say 0.2 volts, determines a vane deflection for each of the three attenuator settings. Moving vertically on the graph along the lines representing these deflections until the intersection with the 50 c/s calibration line is reached provides the corresponding values of r.m.s. gap voltage. Substituting pairs of values into equation 5.2 gives the values 0.131cm^{-1} and 0.135cm^{-1} for the attenuation coefficient, from attenuator settings 16-18cm and 18-20cm respectively, in reasonable agreement with theory. The curvature at the lower ends of the calibration graphs was due

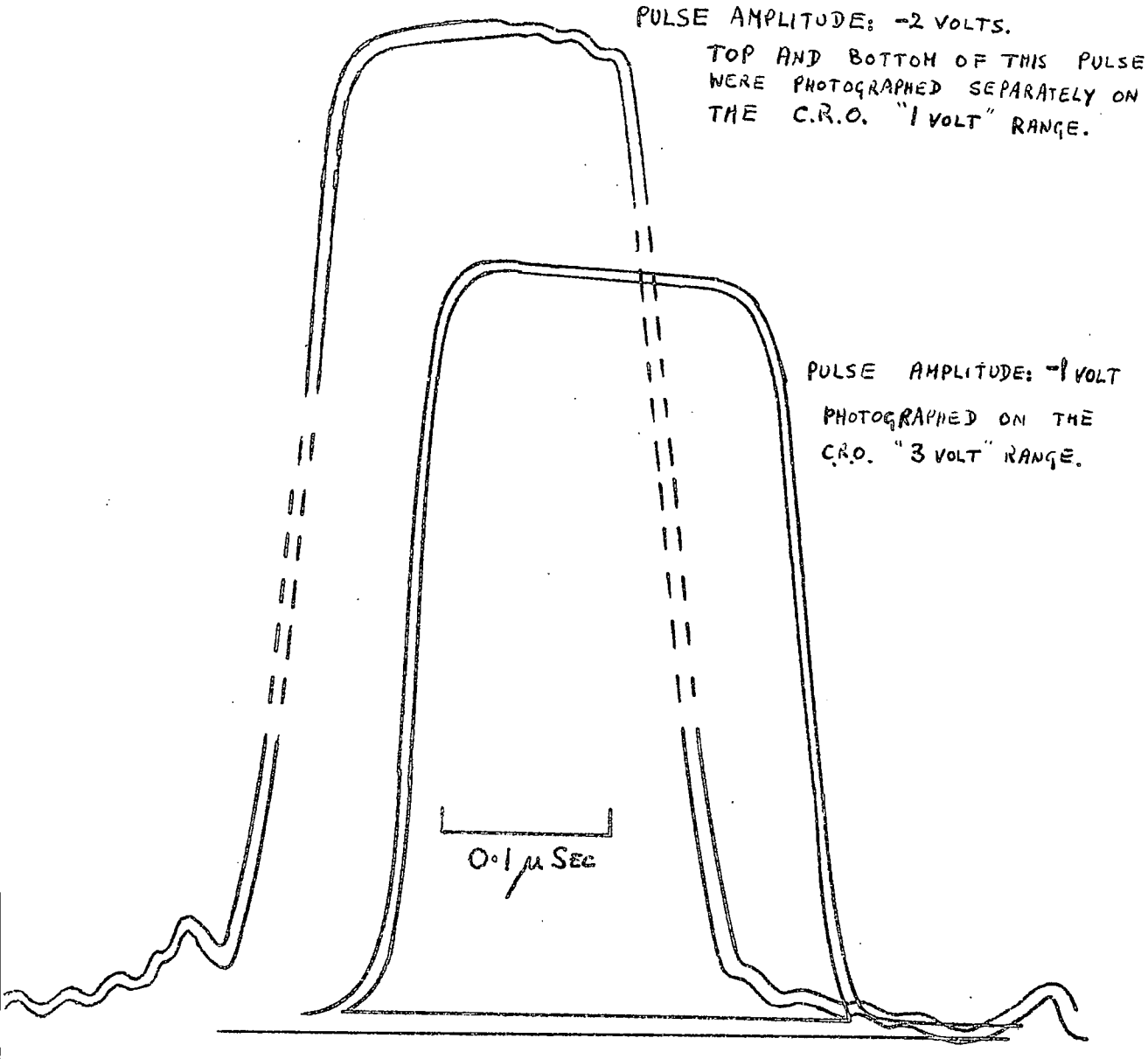


FIG. 5.4 OSCILLOGRAMS OF TEST PULSES ON TWO SETTINGS OF THE "Y" AMPLIFIER INPUT ATTENUATOR

to curvature of the diode characteristic.

All the measurements of breakdown voltages to be presented here were made at one of the piston attenuator settings 16, 18 or 20 cm, and so could be obtained without the explicit use of the attenuation factor. This agreement with well established theory does, however, give confirmation of the correct functioning of this vital part of the measuring apparatus.

5.2.3. The oscilloscope.

This was a Marconi model 1330/2. The rise time was specified as 20 n.sec. and test pulses(fig. 5.4) showed that this was achieved, for a pulse of 1 volt amplitude. The Y amplifier had a uniform response from d.c. to 15 Mc/s. A feature of this model was the stability of the Y amplifier gain: to check its calibration, a Weston standard cell was connected across the input terminals from time to time but no deviation from the factory calibration by more than 1% was ever observed.

For measurements of the transient voltages associated with gas breakdown, the oscilloscope trace was photographed, and the resulting negatives, on 35mm film, were projected by a photographic enlarger onto a sheet of graph paper. The projected height of a typical oscillogram was 10cm., and the width of the trace at this enlargement was between 1 and 2mm, enabling estimates of pulse height to be made to an accuracy

of 0.5%. The pulse height was referred to the voltage of the Weston cell, 1.018 volts. This was accomplished by connecting the cell across the Y amplifier input terminals and photographing the trace twice on one frame, firstly with a.c. coupling and secondly with direct coupling. The a.c. coupling position of the CRO input control merely placed a capacitor in series with the input terminals, so that the resulting photograph showed two traces, separated by a distance representing 1.018 volts. A time calibration waveform was superimposed upon each of the voltage calibration traces. This was of frequency 100 Kc/s or 1 Mc/s chosen according to the sweep velocity range being used, and was generated by a crystal-controlled oscillator having two alternative crystals operating at their fundamental parallel resonant frequencies.

The voltage and time calibration was made on every length of film, at each of the timebase settings used, so that the voltage of every pulse was known to an accuracy of about 0.5% at the oscilloscope terminals.

5.2.4. The coupling between piston attenuator and oscilloscope.

The observed effect of applying to the oscilloscope an amplitude-modulated sine wave of frequency greater than the cut-off frequency of the Y amplifier is shown schematically in fig. 5.5. As the frequency was raised above 15 Mc/s, the

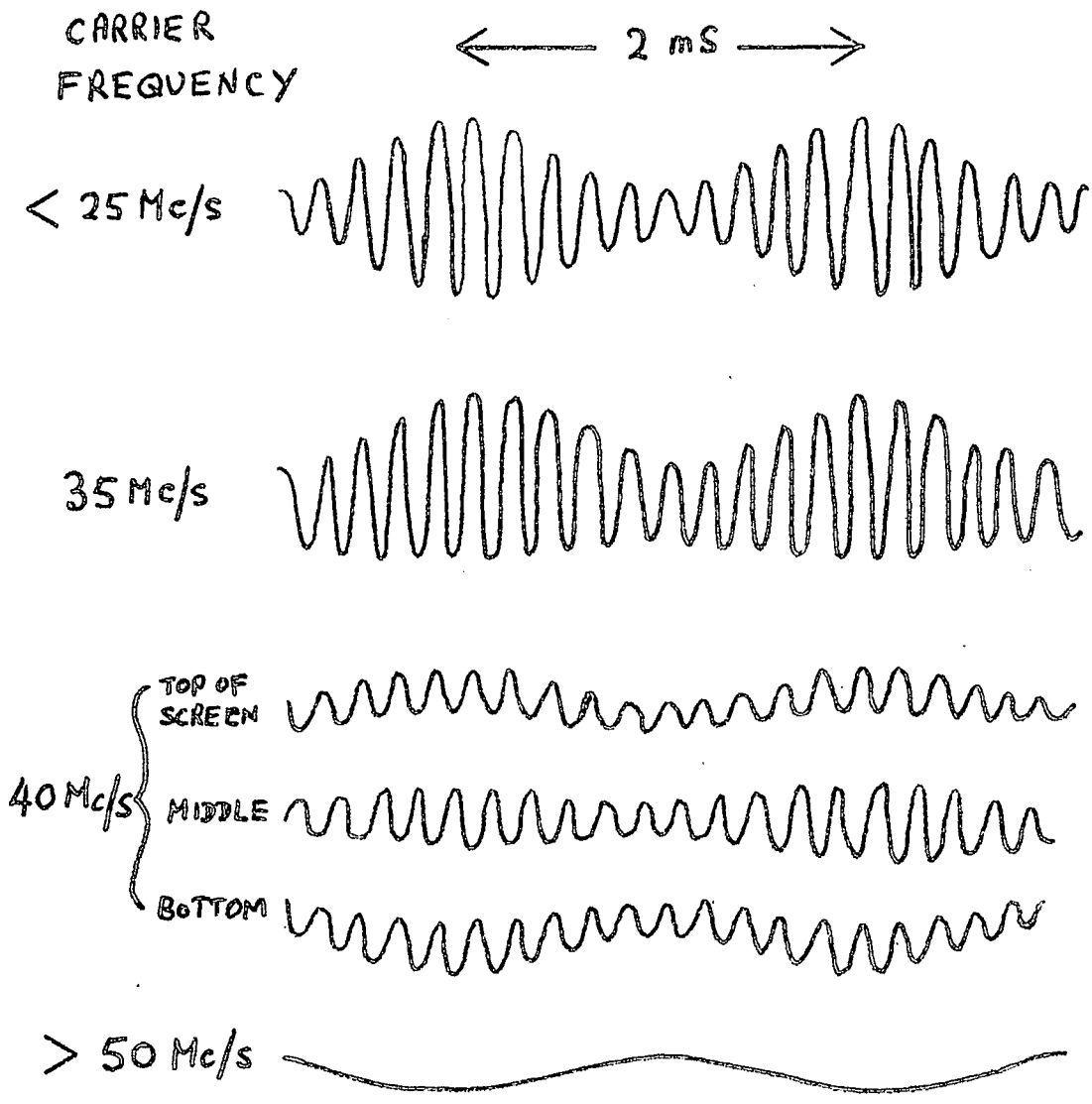


Fig 5.5 THE RESPONSE OF THE OSCILLOSCOPE TO FREQUENCIES ABOVE THE AMPLIFIER CUT-OFF FREQUENCY

amplitude first decreased, then became distorted above 25 Mc/s, the top or bottom of the modulation being evened out according to the position of the Y shift control. At 50 Mc/s only the modulation remained. These facts indicate rectification occuring within the oscilloscope amplifier itself, which, since the effect varied with the Y shift, appears to have been occuring at the grid of the input cathode follower. These observations indicated the necessity of eliminating all traces of signals having frequency above 25 Mc/s from the oscilloscope input terminals, especially the residual 180 Mc/s ripple from the half-wave rectifier in the piston attenuator.

In the original experiments, a two stage C.R. network with a time constant of 10 micro-seconds had been used to filter out the high frequency component of the rectified R.F. voltage - see fig 5.6(a). Preliminary measurements on oscillograms showing the decay of gap voltage at breakdown indicated that the decay was exponential with a time-constant which was slightly dependent upon overvoltage. The exponential shape suggested that discharging of the filter capacitors was causing the observed decay, rather than a gas discharge effect. It was clearly necessary to reduce the time-constant associated with this part of the circuit. Toward this end, the r.f. filter was removed, and the time constant was thereby reduced to 2 microseconds. This removal allowed an

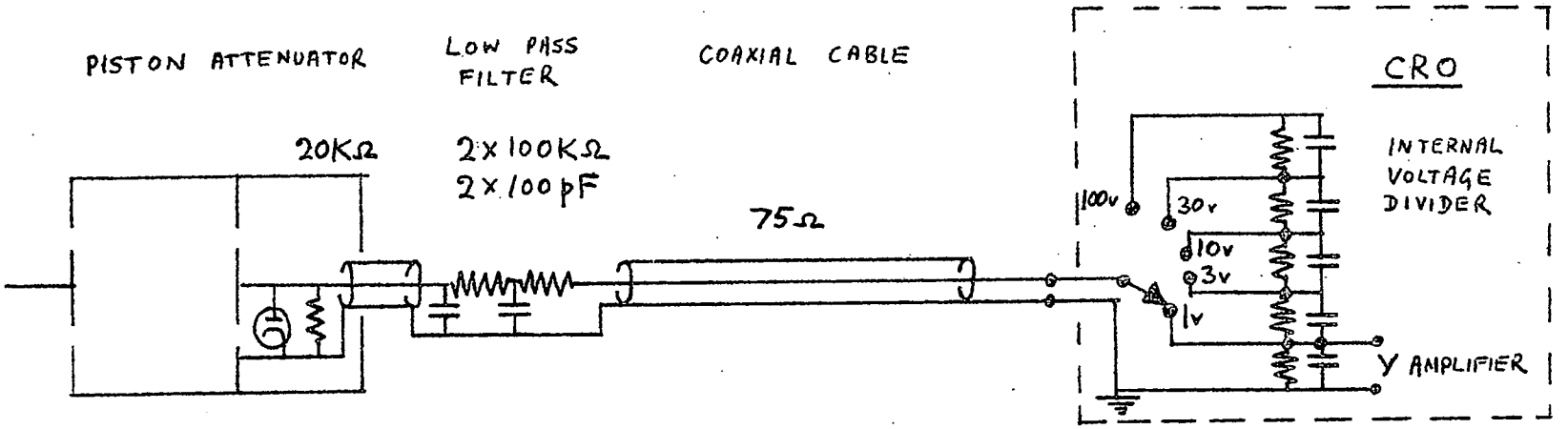
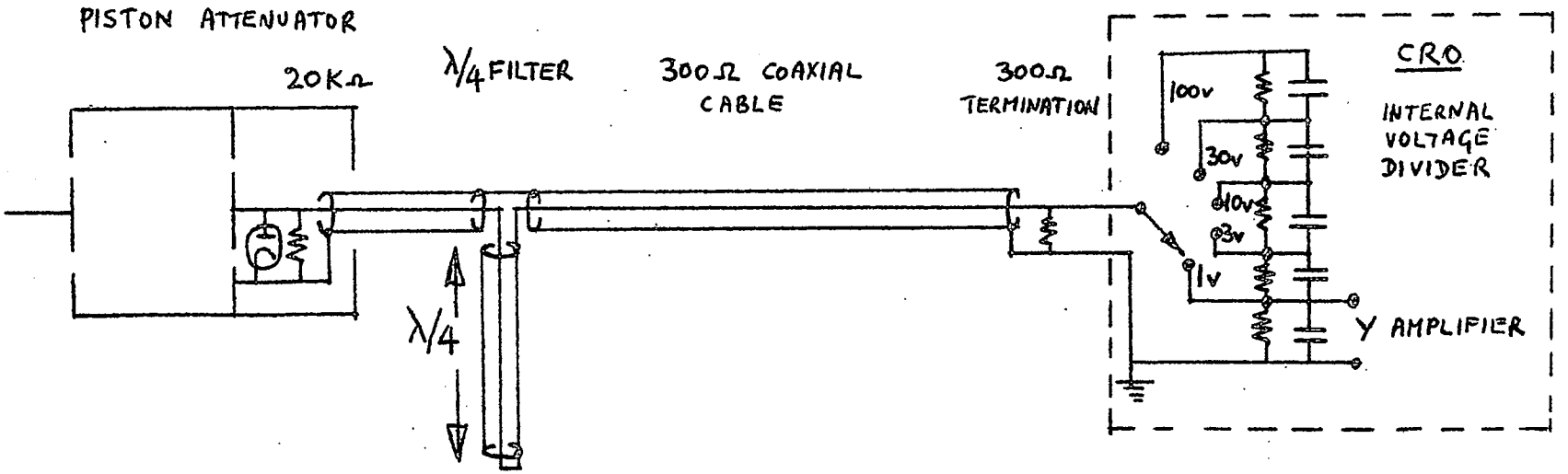


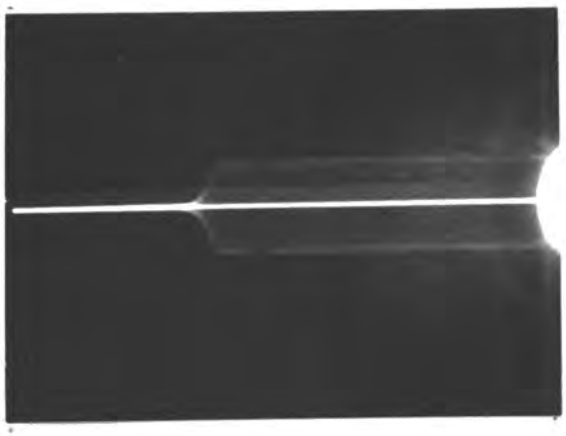
FIG 5.6 (a) ORIGINAL RF FILTER AND COUPLING TO OSCILLOSCOPE

FIG 5.6 (b) COUPLING TO OSCILLOSCOPE MODIFIED FOR OBSERVATION OF FAST TRANSIENTS



r.f. component access to the oscilloscope, with consequent rectification as described at the beginning of this section. However, oscillograms of the breakdown under these conditions showed that variation of the time-constant with overvoltage was still being masked, to a large degree, by the comparatively long time-constant of the coupling network. This time-constant was further reduced to 0.03 microseconds by using low capacitance cable for the connection between the piston attenuator and oscilloscope, and by terminating the cable at the CRO input terminals with a 300 ohm carbon resistor, to match its characteristic impedance. The 300 ohm resistor shunted the 20000 ohm resistor inside the piston attenuator, and both were effectively in parallel with the combined cable and stray capacitance.

To attenuate the r.f. component of the output from the piston attenuator without increasing the time-constant, a quarter-wave choke was used. This was made from a piece of polythene-dielectric coaxial cable which was short-circuit terminated. The terminals of this choke were joined in series with the central conductor of the low-capacitance cable at the point where it emerged from the piston attenuator as is shown in fig 5.6(b). The length of the quarter-wave choke was found to be critical and an accuracy of about 1mm was required. As this accuracy could not readily be attained in practice, the length was adjusted to be as nearly correct as possible, and then the oscillator frequency was altered



(a) Unrectified gap voltage.
decay time 0.18 microseconds.

← time

(b) Rectified gap voltage.
decay time 0.20 microseconds.

Fig 5.7. Comparison of unrectified and rectified spark gap voltages.

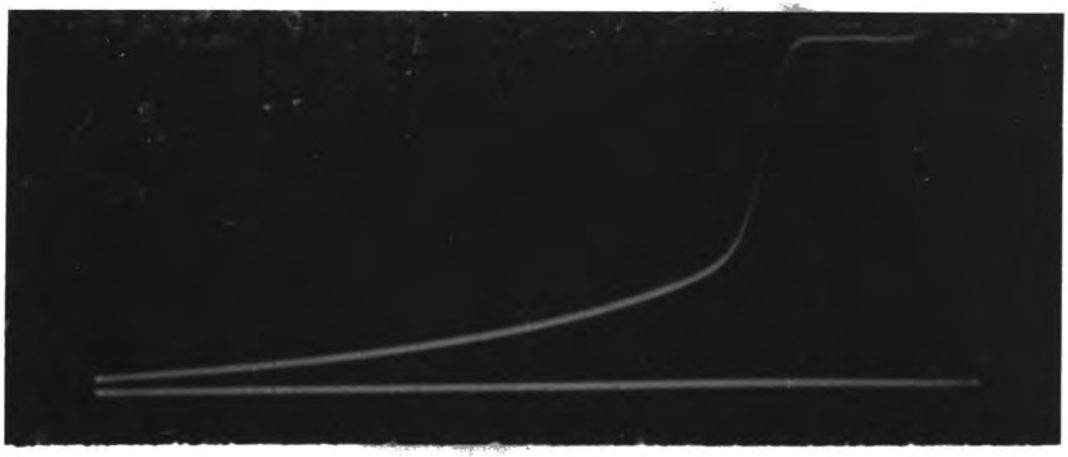
slightly (in conjunction with the length of the transmission line) until no residual r.f. appeared at the terminals of the oscilloscope. An r.f. voltage at the Y input terminals was rectified, as we have seen above, by some unidentified agency within the oscilloscope, and therefore could be observed as a steady deflection of the trace, even when the diode in the piston attenuator was inoperative.

In order to show that the breakdown transient was now being observed properly, unmasked by the effects of time constants in the associated circuitry, a photograph was taken of the unrectified 180 Mc/s oscillations. In order to accomplish this, the heater of the diode in the piston attenuator was switched off and the piston attenuator output was applied directly to the Y plates of the oscilloscope tube. The piston attenuator had to be closed up more than usual to provide sufficient output to deflect the beam appreciably. With this input to the cathode ray tube the spot moved too rapidly for most of each cycle to produce a continuous trace on the photograph, but was slow enough near the peaks of the sine wave for a faint trace to be discernible. The formative delay on this photograph (fig 5.7a) was about 0.5 microseconds, and the decay had a time constant of 0.18 microseconds. This was compared to the photograph (fig 5.7b) of the rectified wave, taken soon afterwards, in which the formative lag was adjusted to be the same as before. The

decay time constant here was 0.20 microseconds, in good agreement.

5.2.5 The effect of the voltage divider within the oscilloscope.

Figs 5.8(a) and (b) show oscillograms of two nominally identical discharges in neon. The only difference between the experimental conditions for these two pictures was that (a) was photographed on the 3 volt range and (b) was photographed on the 1 volt range of the oscilloscope. Clearly the first pulse has a long "tail", not present on the second. Reference to the circuit diagram of the oscilloscope (see fig 5.6b) showed that only on the 1 volt range was there a direct connection to the grid of the first valve. On the higher ranges (3V, 10V, 30V) the signal passed through a resistive voltage divider, which was compensated capacitatively for the effects of stray capacitance. 1 volt test pulses from a commercial pulse generator passed through the voltage divider without appreciable distortion (fig 5.4) and it is not known how the discharge waveforms differed from these, so as to give rise to such an effect. Referring back to fig 5.7. it can be seen that the "tail" was not associated with the discharge. Even supposing some residual r.f. was still getting into the oscilloscope, one would not expect this difference between the two ranges because r.f. "breakthrough" caused a positive deflection, while the normal pulses observed were in a negative direction.



(a) on "3 volt" range of CRO

(b) on "1 volt" range of CRO

Fig 5.8. Breakdown oscillograms for neon (5[illegible] torr).

of readings, and it can be seen that nearly all the points lie on a smooth curve to this accuracy, when an allowance of ±0.5mm is made for the vane readings. Since the errors involved were constants, the relative accuracy depends upon the applied voltage and on the piston attenuator setting. Table I gives the percentage errors of the absolute gap voltage for representative positions in the calibration graph fig 5.2.

TABLE I

Absolute accuracy of the oscilloscope calibration.

Oscilloscope voltage	0.04	0.10	0.20
Piston attenuator setting			
16cm	17%	6%	1.9%
18cm	8.0%	3.8%	1.6%
20cm	5.1%	3.2%	1.3%

The trace was considerably narrower on "single shot" operation of the oscilloscope, as used for pulsed breakdown measurements, than on "repetitive sweep" conditions which were used for calibration. The observational error in the measurement of pulse height on oscillograms was 0.5%. It is this figure which determines the relative accuracy in a single series of observations, and which is important, for instance, in the measurement of small overvoltages.

The timebase calibration on the CRO was carried out, as described previously, by using the output voltage from a crystal-controlled oscillator. The absolute accuracy of this oscillator was of the order 0.01% so that the trace-width, 0.2mm, again determined the measurement accuracy.

It was discovered that the Y deflection on the CRO was 89^{o} to the timebase, in such a direction as to cause the decay-time of the spark-gap voltage at breakdown to be apparently shorter that it was in reality. The error due to the assumption of X and Y orthogonality was usually negligible, but could become important on very steeply rising pulse edges, when risetimes were being measured. However, in these cases the misalignment could easily be allowed for.

CHAPTER VI

Vacuum Equipment

6.1. A description of the vacuum system.

The prime requirement of a system for handling pure gases is a low outgassing rate of the walls. The achievement of an ultra-high vacuum before the admission of gas to the system is not a basic necessity, because high purity could be achieved by several flushings with the working gas. However, a good vacuum is, in practice usually obtained automatically if outgassing is reduced.

The dominant source of impurities in the vacuum system used in the experiments of Prowse, Rowbotham and Monk,[9] was the grease in the stopcocks and in the demountable seal between the metal and glass parts of the electrode chamber. Apiezon "N" grease was used but although this is of high purity, and has low vapour pressure, it was found to be unsatisfactory in the present application because it absorbs air readily and was observed to release gas steadily under vacuum as stopcocks were turned. Moreover, its low melting point prevented the use of heat to outgas the other parts of the apparatus. An attempt was therefore made here to outgas the existing system by baking under vacuum while using silicone grease for all joints. A considerable improvement was achieved as far as the ultimate pressure and the outgassing rate were concerned, but the grease proved unsatisfact-

ory for ground glass joints, especially non-rotating cone and socket connections. The trouble was that silicone grease has the properties more of a viscous liquid, than of a true grease. Thus it never "sets" and it tends to creep along the inside of glass tubes by the action of surface tension, and to flow out of stopcocks under the influence of the atmospheric pressure acting on them.

Because of the difficulties and disadvantages of "greasy" gas handling equipment, a vacuum system was constructed which contained no grease whatsoever, beyond the diffusion pump, and which was capable of being baked to the limit imposed by the softening point of glass, i.e. 400^{o}-450^{o}C. Before describing the final design of vacuum system, however, it is of interest to mention an observation made while using the silicone grease: with neon in the system at a pressure of a few torr, it was noticed that flashes of red light were emitted from the gas in the hollow interior of a stopcock as it was turned. It appears that the shearing action on the grease caused the separation of charge, and generated sufficiently high electrostatic fields to produce breakdown in the gas. The effect occurred in most of the stopcocks containing neon, but not when they were lubricated with Apiezon grease.

The diagram, Fig. 6.1 shows the final vacuum system, as closely as is possible in two dimensions, to the actual

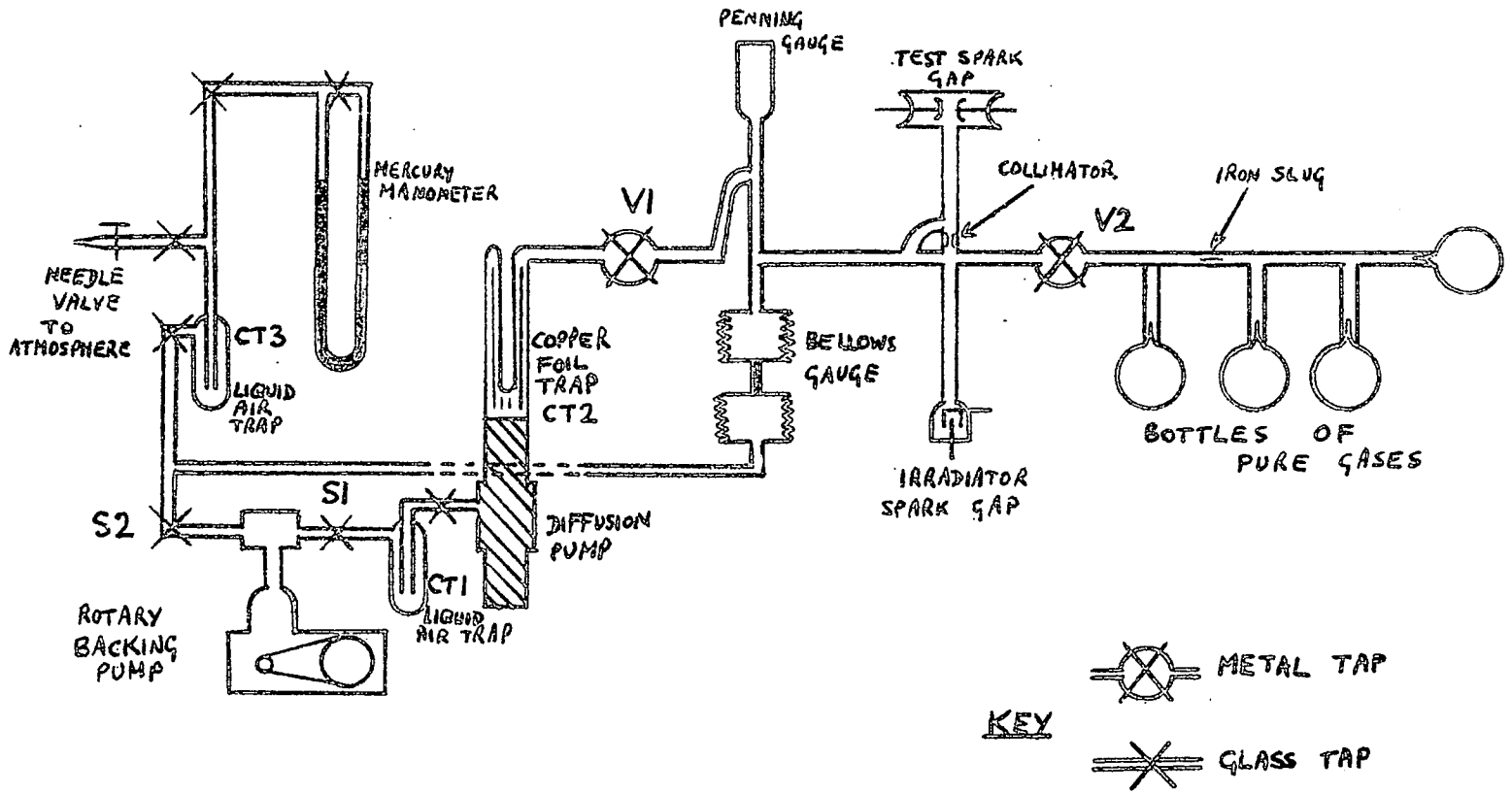


Fig 6.1 SCHEMATIC DIAGRAM OF THE GAS HANDLING SYSTEM

physical layout.

An Edwards rotary pump was used for preliminary exhaustion of the vessel, and provided a backing pressure sufficiently low to enable an oil diffusion pump to be used for evacuation of the test chamber. Since the rotary pump was incapable of pumping condensible vapours, these were removed from the region above the rotary pump by a liquid air trap CT1.

On the high vacuum side of the diffusion pump, a trap CT2 was mounted on top of the diffusion pump to prevent back streaming of oil vapour. The trap CT2 consisted of thin corrugated copper foil, coiled around the inside of a wide-bore pyrex tube, a type which has been used successfully by Alpert,[16]. The tube was made re-entrant so that it could be cooled by filling the cavity with liquid air. The outlet from the trap led via an all-metal, Alpert-type[16] tap, V1, to the main vacuum enclosure, joining the irradiator-collimator tube midway between the test gap and the irradiator. Since the collimator plug offered a high impedance to the flow of low pressure gas it was bypassed with a wide bore glass tube. Also off this enclosure were two pressure gauges: a Penning ionization gauge measured pressures in the range 10^{-3} to 10^{-7} torr whilst the system was being evacuated, and a differential bellows gauge measured from 0.1 to 200 torr in the working gas. Another metal tap, V2, separated the test enclosure from the gas storage reservoir.

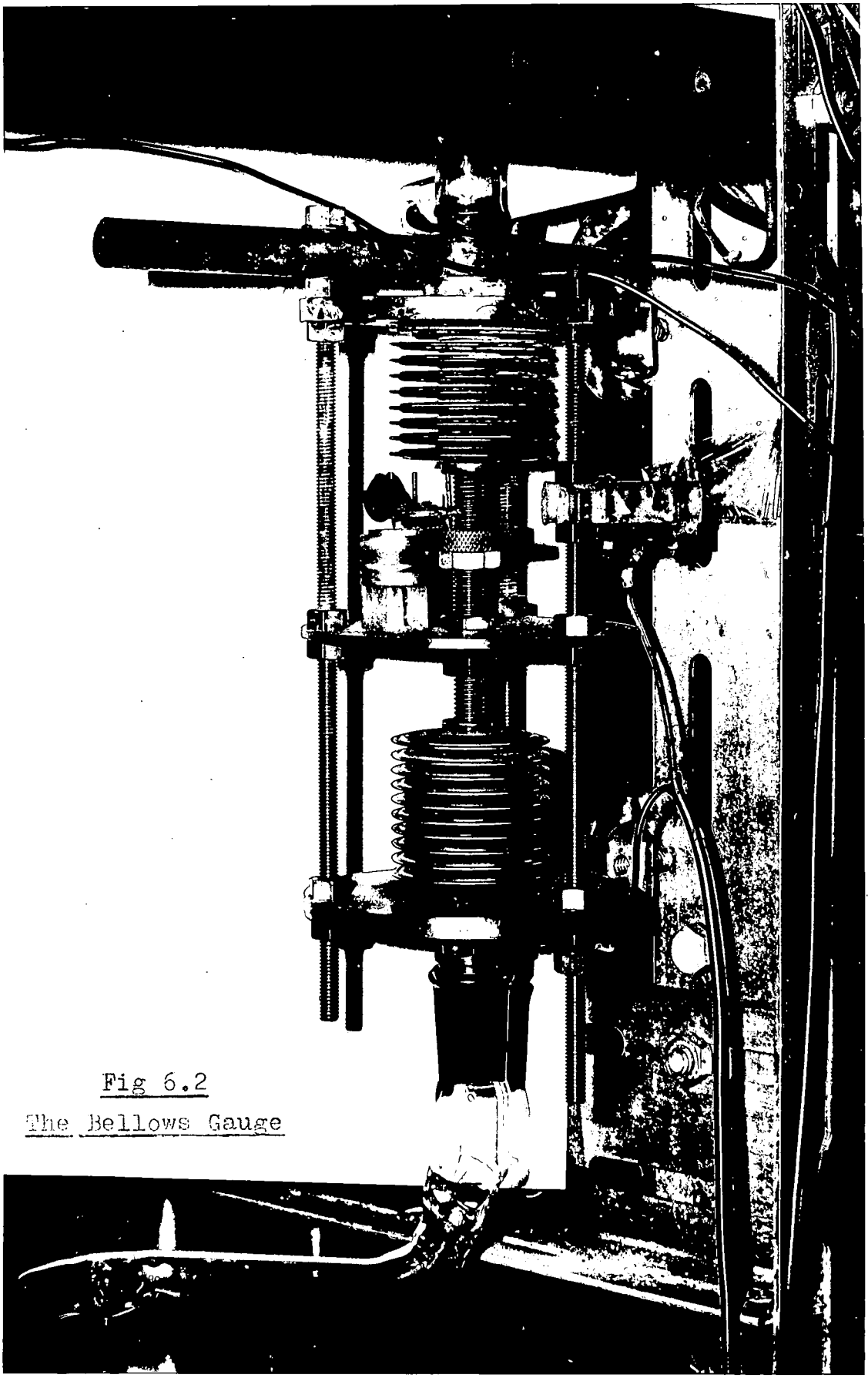


Fig 6.2

The Bellows Gauge

Gas was purchased from the British Oxygen Co. Ltd. in the form of 1 litre samples termed "spectroscopically pure", and no attempt was made to improve on the manufactured purity.

Each gas bottle was sealed off at the factory where the seal was drawn to a fine point. After the bottle was connected into the vacuum system, the glass seal was broken by the impact of a small pellet. The pellet was made of iron so that it could be manoeuvered from outside the system by means of a magnet, and was enclosed in a pyrex capsule to reduce possible gas contamination from the iron.

6.2. The bellows gauge.

This gauge, a photograph of which is shown in fig. 6.2, was used for measurement of the pressure of the working gas, and was constructed in the same style as one used by Prowse, Rowbotham and Monk.[9] It consisted of two flexible chambers, (bellows), of the type commonly used in Aneroid barometers. The bellows were cylindrical in form, mounted vertically one above the other and connected axially by a solid steel threaded rod. The outermost ends of the two units were fixed rigidly to the vacuum system framework, so that the steel rod could move axially in response to pressure differentials within the bellows. Linear movements of the rod were magnified by an optical lever. The mirror for the optical lever was a 1cm diameter galvanometer type of power+1

dioptre. It was mounted on a miniature tripod which stood with two of its legs on a horizontal platform attached to the moving steel rod, and its third on a fixed platform. To avoid the slip-stick motion caused by the tendency of the tripod to rotate, while the bellows moved linearly, the third leg was placed on a polystyrene raft which floated on mercury. By this means, perfectly reproducible deflections could be obtained. When the tripod was removed and replaced, the scale reading, at a distance of 1 metre from the mirror, did not change by more than 1mm.

The bellows gauge was calibrated by admitting the working gas to the upper side, (with the lower side evacuated.) noting the deflection, and then admitting air slowly to the lower side until the original zero position was regained. The air on the lower side also acted on the mercury manometer, so that by reading the difference in levels with a catheto-meter an absolute measure of the gas pressure in the spark gap chamber was obtained.

The gas on the manometer side of the bellows gauge was pumped out, when necessary, by the same rotary vacuum pump as was used for backing the diffusion pump on the main system. Mercury vapour has the property of considerably reducing the breakdown potentials of many gases[33] and therefore great care was taken when evacuating the manometer, to avoid the possibility of mercury vapour finding its way via the

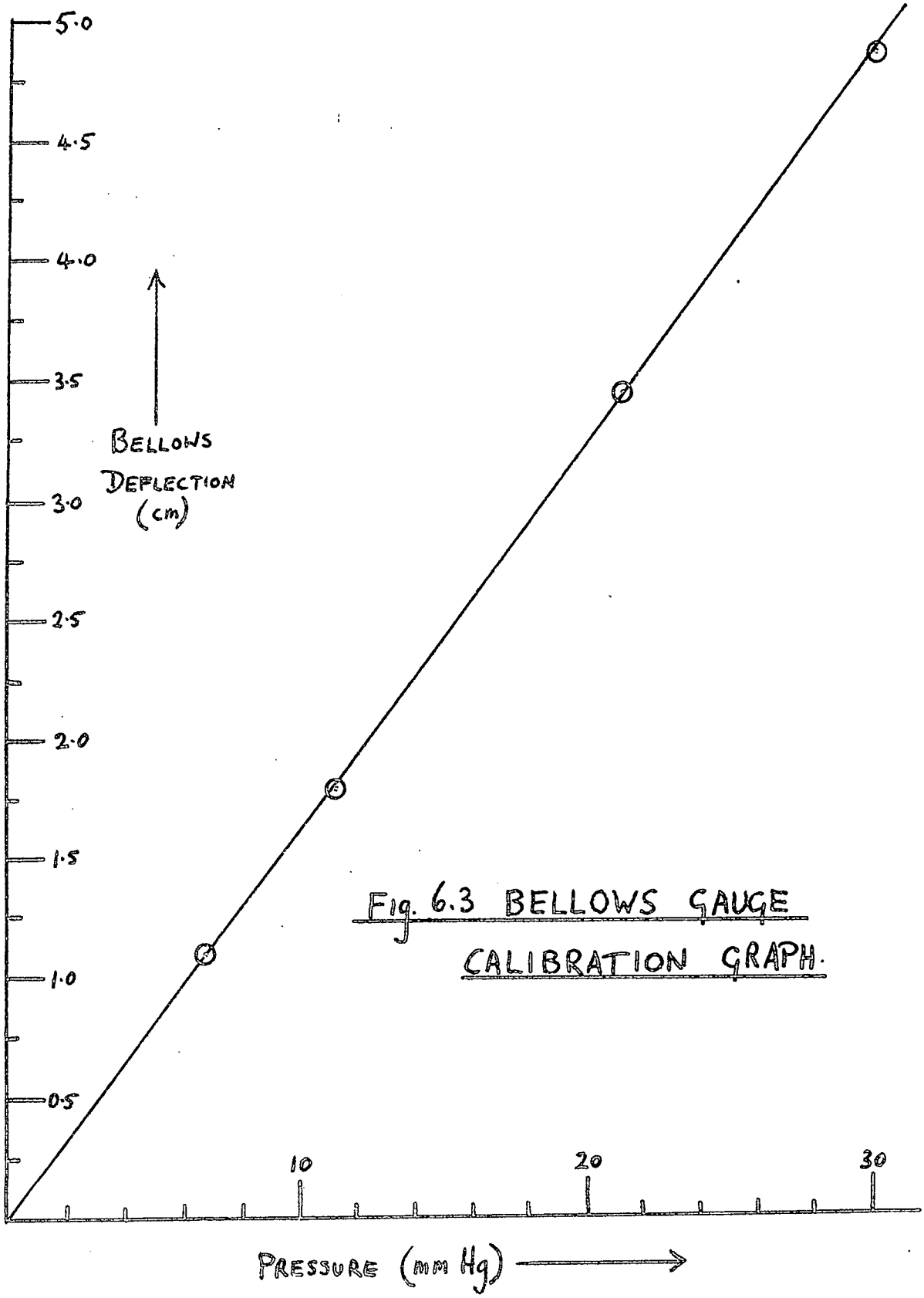


Fig. 6.3 BELLOWS GAUGE CALIBRATION GRAPH.

pump into the test chamber:

(a) The glass stopcocks S1 and S2 in fig 6.1. were never allowed to be open simultaneously.

(b) The cold trap CT3 was always filled with liquid air at least 15 minutes before S2 was opened.

(c) S2 was always closed at the earliest opportunity, i.e. as soon as the pressure was reduced below the least for which a bellows deflection could be observed - about 0.01 torr.

(d) After the manometer had been evacuated and S2 closed the rotary pump was flushed several times with air before S1 was opened.

The sensitivity as obtained from the calibration graph, fig 6.3, was 0.594 torr per mm deflection. Since estimations of 0.1 mm deflection were possible approximately 5% accuracy was achieved at 1 torr and correspondingly better at higher pressures. If needed, an increase in sensitivity of the order 100 times could readily have been secured by decreasing the separation of the tripod legs and by increasing the scale distance.

6.3. Outgassing.

Removal of gas occluded in the glass and metal walls of the vacuum vessel was effected everywhere beyond the nozzle of the diffusion pump by baking under vacuum, to the highest temperature possible, compatible with the nature of the enclosure, and for as long as possible. Most of the vacuum

enclosure was made of pyrex glass, which could be baked up to about 450°C. The metal taps could also withstand this temperature and so could the copper to glass seals of the test spark gap. All these parts were heated by means of an electrical heating tape which was capable of raising their temperature to 300°C in free air and this was further augmented by about 100°C by wrapping with asbestos paper.

The bellows gauge was manufactured using soft solder and therefore could not be baked to a very high temperature. Steam was used to heat it to 100°C and was applied for a week. Although this was not a very satisfactory or convenient expedient, the results seem to show that sufficient outgassing was obtained, as indicated by the improvement in the ultimate vacuum. The Penning gauge was calibrated down to 10^{-5} torr on its internal meter, and according to the makers could be used down to 10^{-6} torr by linear interpolation on an external galvanometer, at which pressure the current to be expected would be 10^{-6} amps. This was, in fact the minimum current obtainable before the bellows gauge was outgassed, but afterwards, a minimum current of 3×10^{-8} amps was achieved. Although this can not be interpreted as a proportionate improvement in the ultimate pressure, it seems reasonable to claim a reduction in pressure by a factor of ten.

In fig 6.1. the kink which can be seen in the pipe

between V1 and the Penning gauge was deliberately incorporated to act as an expansion link. The metal taps, V1 and V2, required a closing torque of 30 foot-pounds, and therefore had to be bolted rigidly to a steel framework. Owing to differential expansion between the steel and glass during baking, sufficient stress could otherwise have been introduced into the glass to break it.

The outgassing sequence was as follows: the whole system was pumped to as low a pressure as possible using the diffusion pump, and then the heating tapes around the whole system including the copper foil trap, were switched on, and steam was applied to the bellows gauge. When the pressure had dropped to about 10^{-5} torr. the heaters of the metal taps were switched on. After a few days, (a week if new glass had just been put in the system) the pressure as indicated on the Penning gauge would be down to 10^{-6} torr. The heater on the copper foil trap was then switched off, and the system was pumped for a few hours more before the rest of the heating was removed.

Initially, liquid air was used to cool the copper foil trap but it was discovered that no reduction in pressure resulted, and therefore cooling was discontinued.

6.4. The spark gap

Platinum electrodes were used for these experiments. They were circular and curved at the edges to give a Rogowski profile. Jones[17] has given a method for producing this profile which is of such a nature as to ensure that the maximum electric field in the vicinity of the gap occurs in the central uniform field region, rather than at the electrode edges.

Prowse, Rowbotham and Monk[9] used an ingenious arrangement with which a variable gap width was obtained. It proved impracticable to design an equivalent system which could be baked satisfactorily for the present experiments. A spark gap of fixed width, 0.494cm, was therefore used throughout this investigation, mounted in cylindrical pyrex enclosure.

The electrodes were made from platinum sheet 0.010 inches thick by simply pressing it between correctly shaped dies which had 0.008 inches clearance at each side. The tight fit at the sides caused the metal to be pulled into shape without fluting. This operation was carried out on a lathe and before the dies were removed the work was spun, and a sharp tool used to trim the rough edges off the electrodes.

The front faces of the electrodes were ground flat and then polished all over to a mirror finish, using the finest grade of emery paper available - grade "0000".

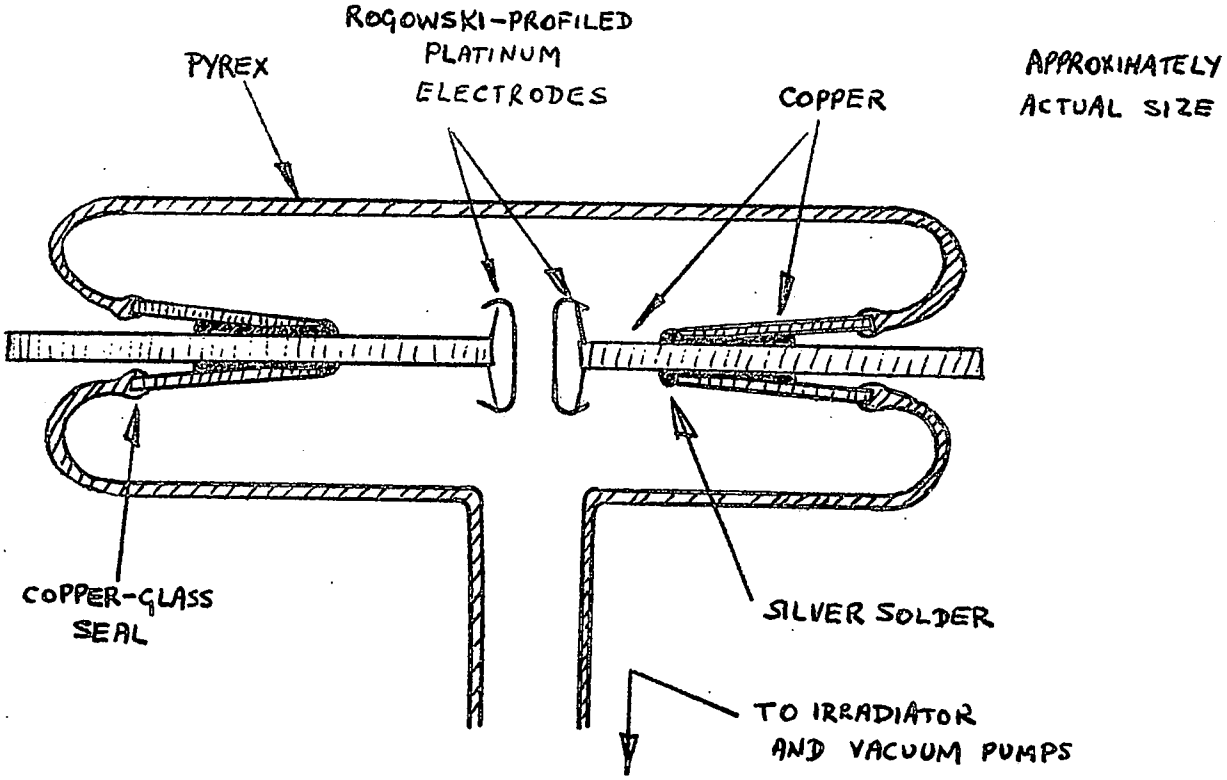


FIG 6.4 THE ELECTRODE CAPSULE

To connect the electrode to the copper leads, a thin platinum strip was hard soldered across the back of each electrode and this was soldered to the end of each lead. The leads were of $\frac{1}{4}$ inch diameter copper rod and were each hard soldered into a tubular glass-copper seal. Finally the two electrodes were sealed into a glass capsule approximately 1.5 inches in diameter. Fig 6.4. shows the final assembly.

The thin platinum strip joining the electrode to the copper lead allowed a certain amount of adjustment of the electrode plane after soldering. Final alignment of the electrodes was carried out by inserting a brass spacing disc between the electrodes and gently pushing the electrodes against this with the glass envelope heated to softening point.

The electrode capsule was evacuated through a side tube which emerged with its axis along a line in the mid-gap plane. The side tube could thus also be used to pass a beam of photons into the mid-gap plane from the irradiator.

CHAPTER VII

THE IRRADIATOR.

The need for some initial free electrons in the spark gap to initiate a discharge has already been noted in Chapter II. It was there mentioned that in these experiments the electrons were produced by photons from a nearby secondary discharge referred to here as the irradiator. In Chapter IV we saw how the voltage pulse which produced the discharge was provided and how its timing was arranged. We now consider the irradiator, and the mechanism of irradiation in more detail.

7.1. The irradiator spark gap

In fig 6.1 can be seen the physical arrangement of the irradiator in relation to the main spark gap. The separation of these two components was 30cm.

The electrodes of the irradiator were dissimilar. The "live" electrode was a pointed tungsten wire which was sealed vertically through the bottom of the glass vacuum envelope, and was coaxial with the glass tube leading to the main gap. The earthed electrode was made in the form of a hollow cylinder, from stainless steel. One end of this electrode was left open and the other end was closed, apart from a hole 0.5mm in diameter at its centre. It was mounted coaxially with the live electrode, and with this projecting into its open end. Application of a high voltage

pulse to this spark gap caused a discharge which was principally located between the tip of the live electrode and the closed end of the earth electrode. This arrangement provided photons from a "point" source so that they could be accurately collimated, and was also convenient for connection to a coaxial lead from the pulse generator.

The collimator was a "pinhole" in a cylindrical plug of stainless steel, which was positioned near the main spark gap and was bypassed for vacuum purposes with a wide-bore tube. Dimensions were such as to allow approximately one part per million of the photon flux from the irradiator to pass into the main gap. Alignment was carried out after assembly by heating the glass tube connection to the irradiator to softening point and visually aligning the two pinholes with the main gap. This was facilitated by running a discharge from a tesla coil in the irradiator gap.

7.2. Theory of the irradiation and effective photon absorption coefficients.

Suppose the irradiator discharge produces I_o photons which are in such a position, initially, and travelling in such a direction as to pass through the collimator into the main spark gap, in the absence of absorption. A gas present in the path of these photons absorbs some so that the number, I, remaining at a distance x from the irradiator is given by

$$I = I_o \exp(-\mu x)$$

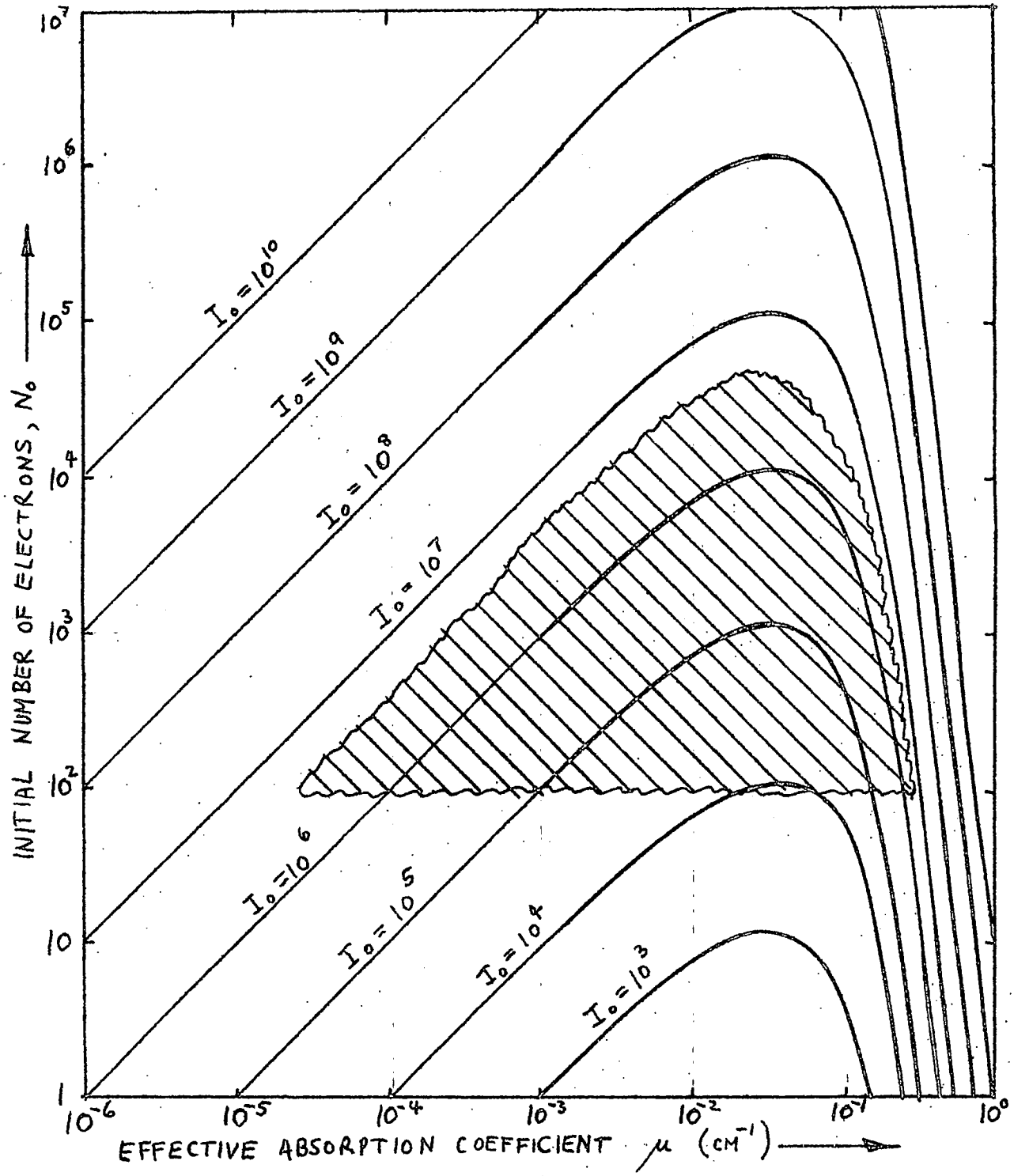


FIG 7.1 THE NUMBER OF ELECTRONS CREATED IN THE TEST GAP AS A FUNCTION OF THE ABSORPTION COEFFICIENT, FOR VARIOUS IRRADIATION INTENSITIES

where μ is an absorption coefficient, assumed constant for a given wavelength. Suppose now we assume that each photon, upon being absorbed, liberates one electron, then the number, N, of electrons produced in an elemental distance dx is

$$N = -dI = \mu I_o \exp(-\mu x)dx \qquad \ldots\ldots 7.1$$

This function is plotted in fig 7.1. for several values of I_o and with dx = 1cm and x = 30cm.

μ is unknown in this expression because of the uncertainty of the wavelength of the absorbed photon, and has been treated as a variable in the graphs of fig 7.1. The peaks of the graphs may be determined by putting $\frac{\partial N}{\delta \mu} = 0$ which gives

$$\mu x = 1$$

Therefore for maximum, $\mu = .033 cm^{-1}$ since x = 30cm.

By substitution in equation 7.1 we have, putting dx = 1cm,

$$N_{max} = \frac{I_o}{30} \cdot \exp(-1)$$

$$\approx 0.01\ I_o \qquad \ldots\ldots 7.2$$

The value of I_o may be estimated in the form of an upper limit by two independent approaches:

First consider the gas molecules within the cylindrical electrode of the irradiator to be fully ionized or excited by the discharge and assume that each molecule emits one photon. The number of molecules in the 5 mm^3, or so, of gas which can "see" the collimating aperture can easily be

determined from Avogadro's number. Avogadro's number is 6.02×10^{23} molecules per gram molecule at N.T.P. We thus arrive at the number $1.7 \times 10^{14} p$ molecules in 5 mm^3 if p is expressed in torr. Therefore $I_o \leq 1.7 \times 10^8 p$, allowing for the $1:10^6$ acceptance of the collimator.

Secondly, consider the capacitor which is discharged through the irradiator gap. If this has capacitance C and is charged to a potential V, then the number of photons of frequency ν produced by the discharge, assuming 100% conversion efficiency would be $(\frac{1}{2}CV^2)/(h\nu)$

Putting $C = 5 pF$

$V = 5000$ volts

$h\nu = 12$ eV (for wavelength 1000Å)

gives a total of 3×10^{13} photons and therefore the number starting towards the test gap is less than about 3×10^7.

Since this last result is of comparable magnitude to that obtained by the former method at p = 1 torr, for pressures greater than 1 torr, the second figure will apply, i.e. $I_o \leq 3 \times 10^7$.

Substituting this value into equation 7.2 we find

$$N \lesssim 3 \times 10^5 \text{ electrons.}$$

If we further assume that at least 100 electrons are necessary to initiate a discharge in the test gap with a small statistical delay (~5% of the formative delay), we define a region, denoted on fig. 7.1 by the shaded area, within which N and μ must lie. From this graph we can read

off the condition $6.10^{-6} \leq \mu \leq 5.10^{-1}$ which is independent of pressure. However, values of μ quoted in the literature are usually in units of cm^{-1} atmospheres^{-1} and so the limits for μ are here converted to these units, by using the expression

$$\mu_1 = \mu/p$$

where μ_1 is the absorption coefficient at atmospheric pressure and p is expressed in atmospheres.

In doing so, for any one gas, we use the lower limit of μ for the lowest pressure used, and the upper limit for the highest pressure used. The limits are given in Table II, where they may be compared with the values of μ published by various authors. These values have been collected and tabulated by Bainbridge and Prowse[14].

It will be apparent from Fig 7.1 that any error in the estimated value of I_o will cause a proportional error in the lower limit for μ, but will make little difference to the upper limit, because of the steepness of the curves in this region. Now it should be remembered that the estimate of the maximum value for I_o involved the assumption of 100% conversion efficiency, between stored energy and photons of the correct wavelength for producing ionization at the test gap. This might easily be an overestimate by up to two orders of magnitude. If this were so, both the limits on μ_1 would be brought to the order unity.

The values of μ_1 in section (f) of table II were obtained

by measurement with a vacuum spectrometer in contrast to the others which used various gas discharge methods: (a) was obtained with a very similar experimental arrangement to that described here; (b), (d) and (e) used Geiger counters while (c) was observed using a Wilson cloud chamber.

TABLE II

The upper and lower limits of the effective photon absorption coefficient as inferred from irradiator theory, compared with published data.

Gas	Hydrogen	Nitrogen	Neon
Highest pressure used (torr)	65.4	45	174
Lowest pressure used (torr)	2.4	0.6	4.7
Upper limit of μ_1 (cm^{-1} atm^{-1})	5.8	8.5	2.2
Lower limit of μ_1 (cm^{-1} atm^{-1})	.002	.008	.001
Published values of μ_1 (cm^{-1} atm^{-1})			
(a) similar to present experiment[14]	-	.25	-
(b) Geiger counter[18]	0.91	-	-
(c) Wilson cloud chamber[19]	0.84	-	-
(d) Geiger counter[20]	27.4	-	-
(e) Geiger counter[21]	1.4	-	-
(f) vacuum spectrometer at wavelength 500Å to 600Å[22,23]	50-230	100-700	130-200

It would appear that the spectrometer method is in considerable disagreement with the other methods mentioned here, including the present experiment. The present experiment, although giving an indication of the absorption coefficient to within about three orders of magnitude, still precludes a value of μ_1 greater than about 10, because of the steepness of the curves of fig 7.1 in this region.

The energies of the photons involved in the irradiation process must have been sufficient to ionize molecules, and therefore, if the possibility of two stage ionization is ignored, the wavelengths involved must have lain within the regions of high absorption as indicated by the spectrographic measurements. A possible explanation could be that very narrow "windows" exist within the absorption continuum which were not resolved by the spectrograph. A certain proportion of recombination photons from the irradiator could have wavelengths within these windows, and give the necessary initial electrons for a discharge.

CHAPTER VIII

Theoretical prediction of the shape of breakdown oscillograms.

8.1. Reflected waves on a transmission line with time-dependent terminating impedance.

The voltage as measured at one end of a resonant transmission line is the sum of a large number of component waves. In the particular case considered at present, one end of the line was open circuit and provided a means of supplying energy to the line from a nearby oscillator. The oscillator was far enough away to ensure that the presence of high r.f. voltages on the line could not disturb it appreciably. Under these conditions a sinusoidal voltage of constant amplitude was induced on the line, which travelled to the spark gap at the far end and was reflected to and fro a number of times before decaying to zero.

Because the transmission line was adjusted to be in resonance with the oscillator, waves arriving at the spark gap were always in phase and their amplitudes can therefore be added arithmetically.

The following will be clarified by reference to fig 8.1: Consider a wave of amplitude V_o induced on the line by the oscillator. This travels down the line and on arriving at the spark gap it has suffered an attenuation f, say, and has amplitude fV_o. Suppose that the discharge has already started

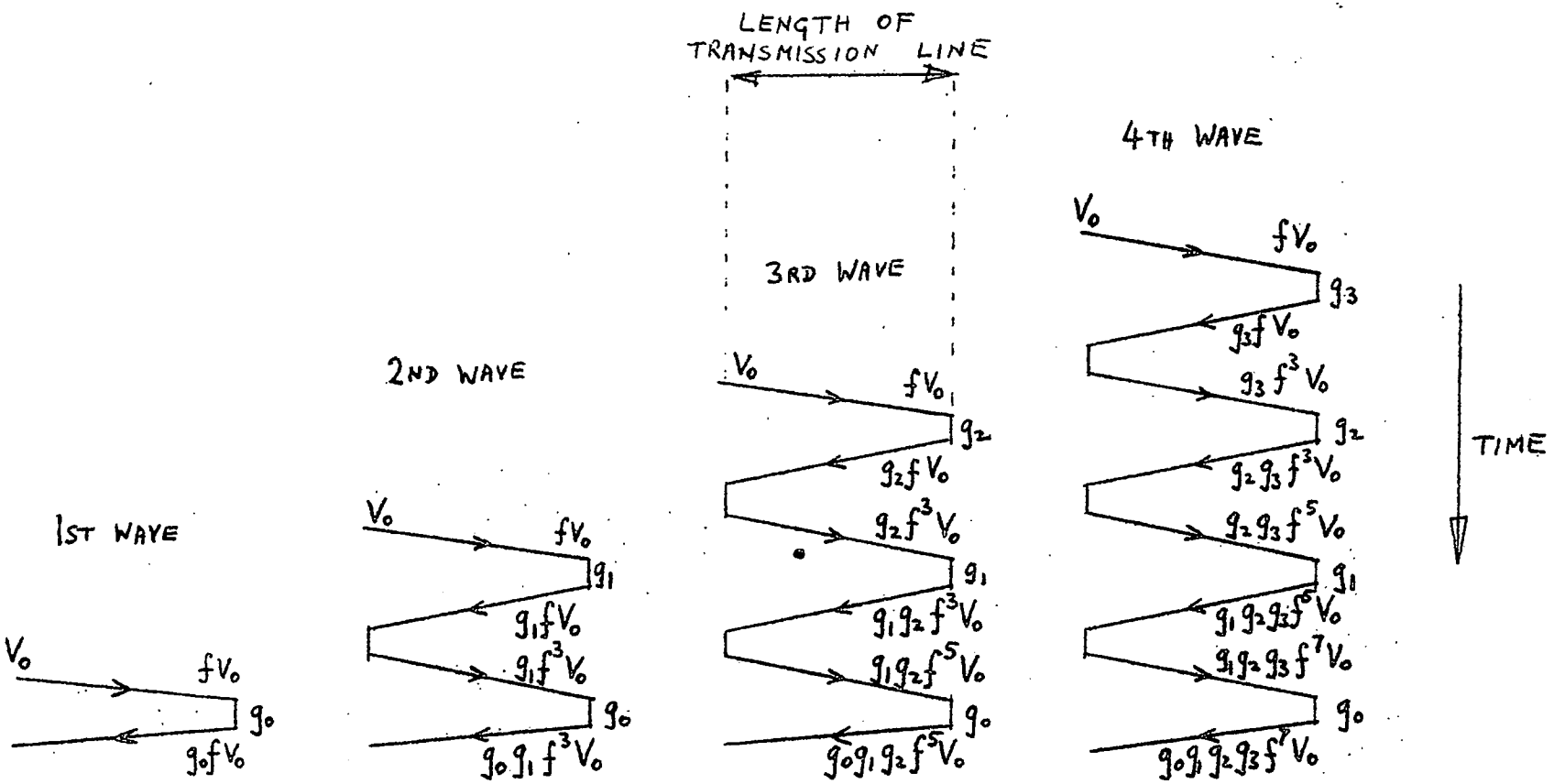


Fig 8.1 REFLECTIONS ON A RESONANT TRANSMISSION LINE : THE HISTORY OF FOUR SUCCESSIVE WAVES

and therefore the reflection coefficient at the gap is not unity, as would be the case for an open-circuit termination, but rather less. Call the coefficient at this instant g_o. The reflected wave therefore has amplitude $g_o f V_o$.

Also arriving at the spark gap at the same time is another wave which had already completed one return journey down the line when the wave previously considered was just starting. Calling the reflection coefficient g_1 when this one reached the spark gap the first time, its amplitude now is $g_1 f^3 V_o$, and is $g_o g_1 f^3 V_o$ after its second reflection at the spark gap if it is assumed that the reflection coefficient at the input end is unity.

A third wave also arriving at the same time will have amplitude $g_1 g_2 f^5 V_o$, which becomes, on reflection, $g_o g_1 g_2 f^5 V_o$.

If the oscillations have been maintained for some time, the resultant voltage, V_∞, is the sum of an effectively infinite number of such waves:

$$V_\infty = f V_o + g_o f V_o + g_1 f^3 V_o + g_o g_1 f^3 V_o + g_1 g_2 f^5 V_o +$$
$$+ g_o g_1 g_2 f^5 V_o + \cdots$$
$$= f V_o (1+g_o)(1+g_1 f^2 + g_1 g_2 f^4 + g_1 g_2 g_3 f^6 + \ldots.) \ldots\ldots \quad (8.1)$$

In the case of an overvolted spark gap, the discharge would be building up and g would be a function of time. However, a special case exists when no electrons at all are present in the gap. Then $g_n = 1$ for all n. If V_{max} represents the sum for this case, we have

$$V_{max} = 2fV_o(1+f^2+f^4+f^6+ \ldots)$$
$$= 2fV_o/(1-f^2)$$

Now, substituting for fV_o in equation 8.1. and defining a dimensionless quantity V, we have

$$V \equiv V_\infty/V_{max} = \tfrac{1}{2}(1-f^2)\ (1+g_o)\ (1+g_1f^2 + g_1g_2f^4 + g_1g_2g_3f^6 + \ldots)$$

.... (8.1a)

V is the voltage amplitude at a given instant relative to the value it would have if there were no conduction in the spark gap.

f is by definition equal to the ratio of the travelling wave amplitudes A_o, A, at beginning and end of the transmission line and therefore we may put

$$f = A/A_o = \exp(-\alpha l)$$

where l is the length of the line and α is the attenuation per unit length.

Now α is related to the quality factor Q and the propagation constant β for the line by[24]

$$Q = \beta/2\alpha$$

and β is by definition equal to $2\pi/\lambda$.

In this case l was two wavelengths long, so that resonance occurred, i.e.

$$l = 2\lambda$$

and therefore, by substitution,

$$f = \exp(-2\pi/Q)$$

which becomes approximately, for large Q

$$f = 1-2\pi/Q$$

Q can be measured by resonance methods so that f is then known.

8.2. The time dependence of the conductivity of the spark gap.

The reflection coefficient g is given in terms of the terminating admittance Y and the iterative impedance, Z_o, by the relation

$$g(t) = (1-Z_oY(t))/(1+Z_oY(t)) \qquad \ldots\ldots (8.2)$$

Z_o is known from the geometry of the transmission line, and it remains, to permit evaluation of the time dependent gap voltage, to find the function Y(t). For this, the equation of continuity for electrons in the gap, must be solved, with the appropriate boundary conditions.

The continuity equation

$$\frac{\partial \rho}{\partial t} = D\nabla^2\rho + \nu_i\rho$$

gives the growth rate of electron density due to ionization,

where D is the diffusion coefficient for electrons, assumed independent of space co-ordinates, and ν_i is the number of new electrons produced per second by one electron.

In cylindrical co-ordinates, this becomes

$$\frac{\partial \rho}{\partial t} = D\frac{\partial^2\rho}{\partial r^2} + \frac{D}{r}\frac{\partial \rho}{\partial r} + D\frac{\partial^2\rho}{\partial z^2} + \nu_i\rho$$ as ρ is

independent of the azimuthal variable.

Assuming a solution of form $\rho(r,z,t) = \rho_o R(r)Z(z)S(t)$ we have, by differentiation and substitution in the continuity

equation,

$$\frac{1}{DS}\cdot\frac{\partial S}{\partial t}=\frac{1}{R}\cdot\frac{\partial^2 R}{\partial r^2}+\frac{1}{rR}\cdot\frac{\partial R}{\partial r}+\frac{1}{Z}\cdot\frac{\partial^2 Z}{\partial z^2}+\frac{\nu_i}{D}$$

This must be satisfied by

$$\frac{1}{Z}\cdot\frac{\partial^2 Z}{\partial z^2}=a^2 \qquad \ldots\ldots (8.3)$$

$$\frac{1}{R}\cdot\frac{\partial^2 R}{\partial r^2}+\frac{1}{rR}\cdot\frac{\partial R}{\partial r}=-b \qquad \ldots\ldots (8.4)$$

$$\frac{1}{DS}\cdot\frac{\partial S}{\partial t}=c \qquad \ldots\ldots (8.5)$$

where a, b, c are constants to be determined.

Equation 8.3 gives $\frac{\partial^2 Z}{\partial z^2}-a^2 Z=0$

which has solution $Z=A\exp(az)+B\exp(-az)$

If a^2 is positive the function Z increases monotonically and cannot satisfy the boundary condition that $Z=0$ at $z=d/2$ i.e. zero electron density at the electrodes. Putting a^2 negative gives the solution $Z=A\cos kz+B\sin kz$

where $k^2=-a^2$ (k^2 is positive).

Symmetrical discharge conditions demand that B=0, and the boundary condition gives $k=(1+2j)\pi/d$ where j takes on any integral value from 0 to ∞

therefore $Z=\cos(j+\frac{1}{2})2\pi z/d$

if $Z=1$ when $z=0$

Equation 8.4 has solution $R=J_o(r\sqrt{b})$ in which J_o indicates the Bessel function of the first kind and zero order.

The boundary condition to be satisfied for evaluation of b is that R=0 at $r=\infty$. However, the J_0 function has negative values which are physically unacceptable and therefore we write R=0 at $r=R_0$, where R_0 is the electrode radius. Hence

$$b = (2.405/R_0)^2$$

and therefore $R = J_0(2.405r/R_0)$

Equation 8.5 has solution $S=S_0\exp(cDt)$ where $S=S_0$ at t=0. Now it is clear that the constants a, b, and c in equation 8.3, 8.4 and 8.5 must be related by the equation $c=a^2-b+\nu_i/D$; therefore we have

$$c = -(2\pi(1+j)/d)^2-(2.405/R_0)^2 + \nu_i/D$$

$$\equiv \nu_i/D-1/L^2$$

where L is the characteristic diffusion length,[25][26]. The question arises as to whether the value of j other than zero can have a physical interpretation. It is possible to conceive of a solution containing the sum of many harmonic terms of individual amplitudes that ensure always a positive value for ρ. However, a finite number of terms would give a multi-peaked concentration when regarded as a function of z. This would contradict Fick's Law which has the effect of smoothing out local concentrations. An infinite number of terms could give a physically acceptable solution, which would consist of the fundamental cosine distribution distorted to give a flattened peak. The corresponding higher

values of j in the time dependent part would then correspond to diffusion down the concentration gradient which would be steeper than a cosine, and give terms of short time constant. However, published data on breakdown in gases has shown[26] that the fundamental term (j=0) is sufficient to give quite accurate results, and therefore only this term will be considered here. The full solution then becomes

$$\rho = \rho_o \; J_o \; (2.405r/R) \cos (\pi z/d) \exp (\nu_i - D/L^2)t$$

where ρ_o (which includes S_o) is the initial electron concentration at the middle of the gap and L is defined by

$$L^{-2} = (\pi/D)^2 + (2.405/R_o)^2 \qquad \ldots\ldots \quad 8.6.$$

8.3. The conductance of the discharge.

Slater[27] has shown that the high frequency conductivity, σ, of a gas is given in terms of the electron density, ρ, by

$$\sigma = \frac{\rho e^2}{m}\left[\frac{\nu_c - jw}{\nu_c^2 + w^2}\right]$$

where e, m are the electronic charge and mass, respectively

ν_c is the electron collision frequency; j is here the imaginary number operator $\sqrt{-1}$,

and w is the radian frequency of the applied electric field.

For a cylindrical spark gap of width d, radius R_o, the gap conductance Y therefore becomes, to a first approximation,

$$Y = \frac{\pi R_o^2}{d}\frac{\rho e^2}{m}\left[\frac{\nu_c - jw}{\nu_c^2 + w^2}\right]$$

Thus if $Y=Y_o$ when $\rho=\rho_o$ we may write

$$Y=Y_o \exp(\varphi t) \quad \ldots\ldots\ldots\ldots \quad 8.7.$$

where $\varphi \equiv \nu_i - D/L^2$ 8.8

Y_o is the conductance caused by photo-electrons from the irradiator and t is measured from the instant of irradiation.

Further, if N_e is the actual number of electrons within the spark gap,

$$N_e = \rho \pi R_o^2 d$$

and therefore

$$Y = \frac{N_e e^2}{md^2} \left[\frac{\nu_c - jw}{\nu_c^2 + w^2} \right]$$

This shows that Y is directly proportional to the total number of free electrons present, N_e.

8.4. Numerical determination of the gap voltage as a function of time.

The only relevant values of g are the ones existing at the instants the waves appear at the spark gap, and it is therefore convenient here to express time in integral multiples of the delay time of the transmission line: Let NT correspond to the interval between the irradiator and the instant at which we wish to know the gap voltage where N is a positive integer, and T is the time it takes for a wave to travel down the line and back. Then NT is the time (measured from the instant of irradiation) at which the

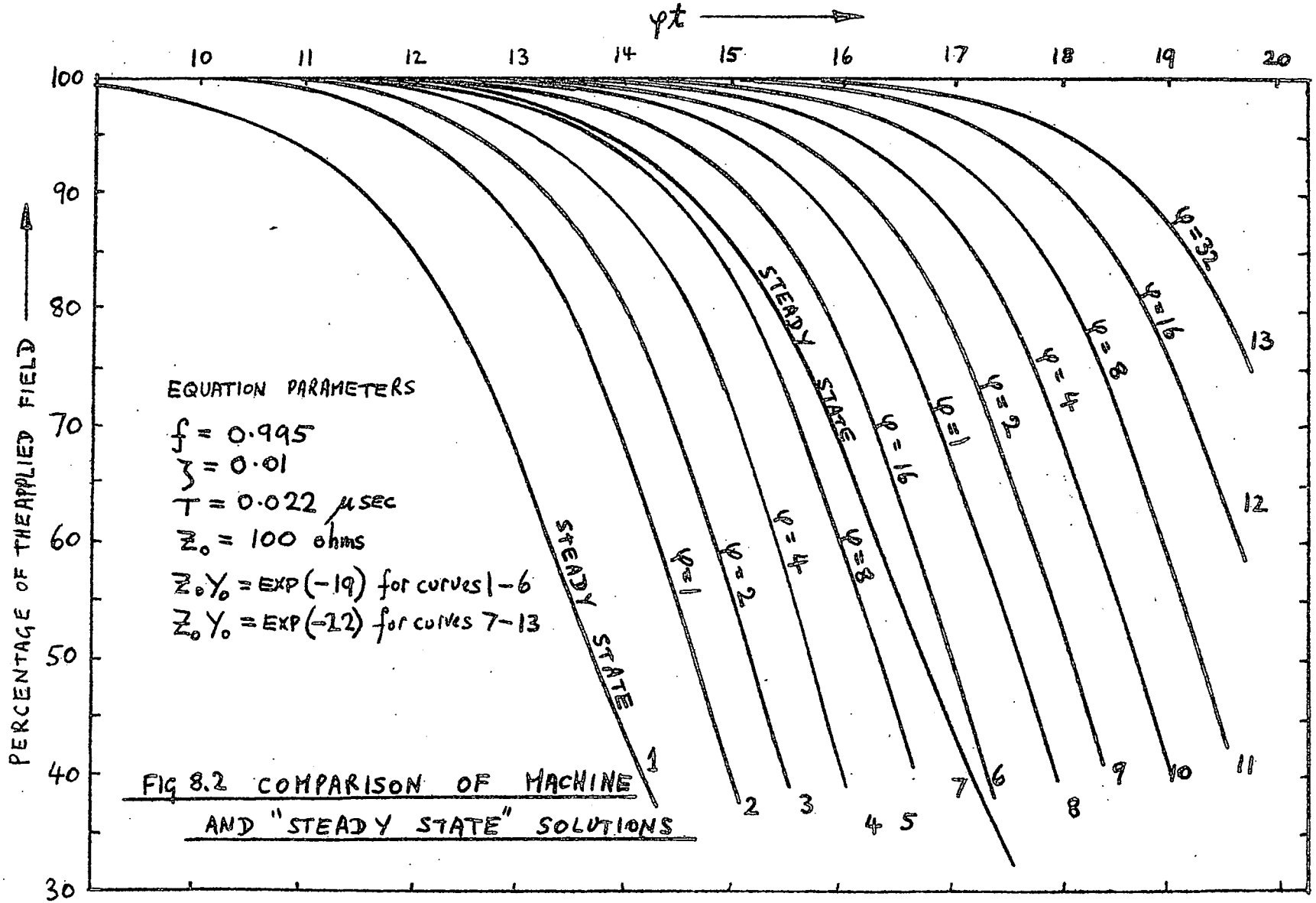


FIG 8.2 COMPARISON OF MACHINE AND "STEADY STATE" SOLUTIONS

reflection coefficient is g_o, while g_1 corresponds to time (N-1)T; g_2 to (N-2)T, and g_n to (N-n)T .

Equation 8.7 may then be written in the form

$$Y_n = Y_o \exp (N-n)\varphi T.$$

and similarly equation 8.2 can be written

$$g_n = (1-Z_oY_n)/(1+Z_oY_n)$$

Knowing T and Z_o we may now solve equation 8.1(a) to give V as a function of time, with Y_o and φ as parameters. In this expression several hundred terms must be included to give a reasonably accurate value for V. The work was therefore carried out on a computer, and the results are plotted in fig. 8.2. Details of the computer programme for this are given in Appendix 1.

It will be seen from the equations that for a fixed value of φ, plotting V as a function of N gives a graph which should correspond to an actual oscillogram. To obtain a reasonable number of graphs with both Y_o and φ used as parameters would take many hours on a computer, using the equation 8.1(a) exclusively. Therefore in practise an initial value of V was determined using equation 8.1(a), with a value of N corresponding to about $\frac{3}{4}$ of the expected formative lag and succeeding values of V were calculated by use of a recurrence formula. Details of this procedure, and of the appropriate computer programme are given in Appendix 2.

The infinite series equation 8.1(a) reduces to a simple geometrical series if we put g_n = constant=g. This may then be easily summed to give

$$V = \tfrac{1}{2}(1-f^2)(1+g)/(1-f^2g)$$

Writing $g=(1-\delta)/(1+\delta)$ where $\delta \equiv Z_o Y_o \exp\varphi t \ll 1$

and $f^2 = 1 - \zeta$ (because $f \approx 1$)

this reduces to

$$V=(1+2\delta/\zeta)^{-1}$$ when second order terms like $\delta\zeta$ are ignored by comparison with δ or ζ.

The "steady-state" values of V calculated from this formula are plotted in fig 8.2 together with the machine solutions. They are called "steady-state" values because the assumption of g constant implies that the transition to conduction occupies a very long time compared to the delay time of the transmission line, T. The machine and steady-state solutions are remarkably similar, but as would be expected, the formative delay is longer for the machine calculation. However, the steady state solution predicts a constant value of φt, whereas for the machine solution, φt increases with φ, where t is the time to a given fraction of the initial value of V.

In practice the formative delay was measured from the beginning of the irradiator marker to the first observable droop in the pulse amplitude. This corresponded to about 99% of the initial voltage which is the value taken for the theoretical predictions.

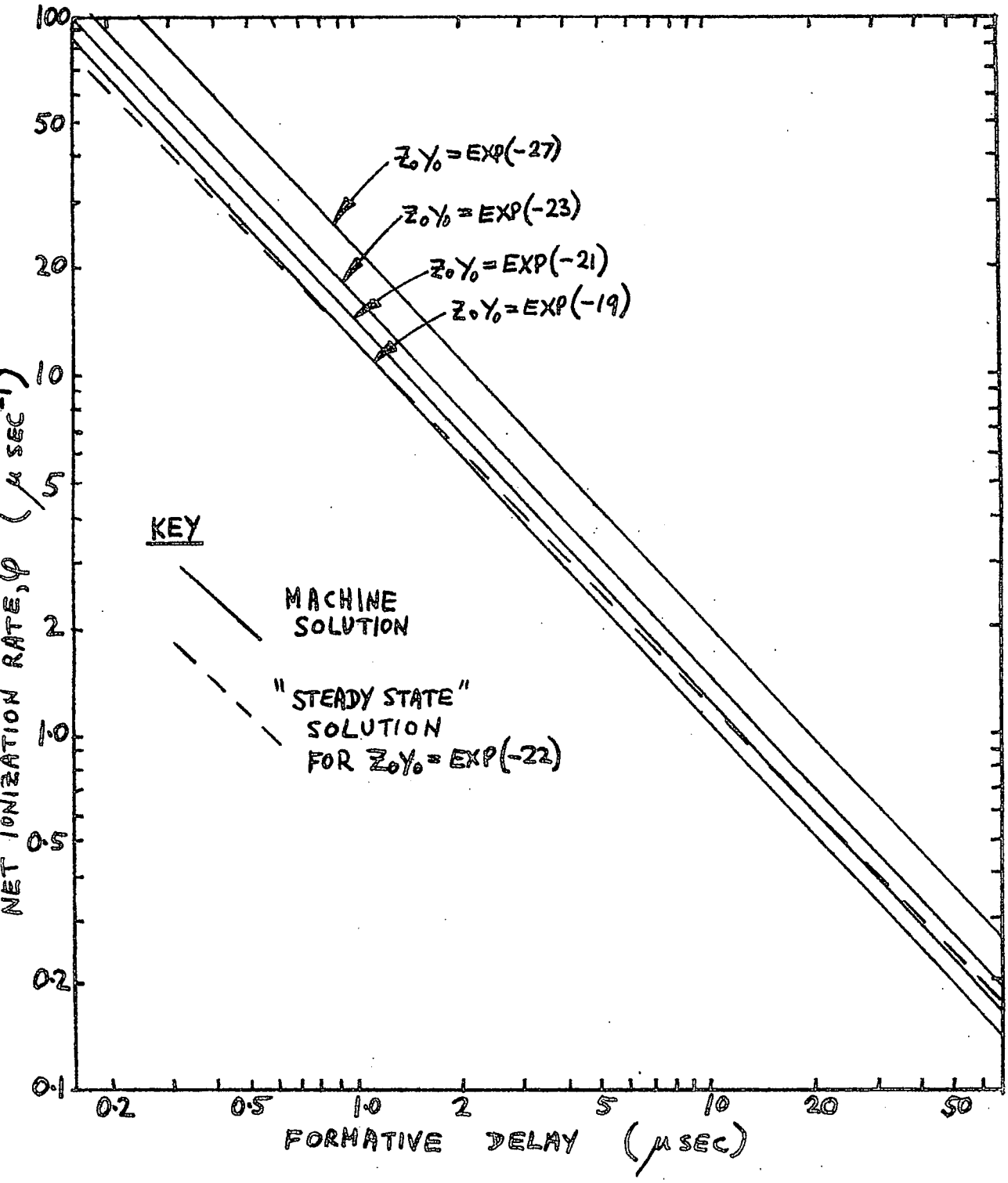


FIG 8.3 FORMATIVE DELAY AS A FUNCTION OF THE NET IONIZATION RATE φ AND THE INITIAL CONDUCTIVITY AS PARAMETER

Both methods of solution predict a decrease in formative delay by $1/\varphi$ seconds for an increase in Y_o by a factor e(=2.718). The formative delay as a function of φ, is plotted on a log-log graph in fig 8.3, for both machine and "steady-state" solutions. The principal difference between the two solutions is in their slopes, but there is little to choose between these, however, as the slopes are both negative, and of magnitude unity and 1.058 for "stead-state" and machine solution respectively.

8.5. Estimation of the net rate of ionization in an overvolted gap.

As we have observed above, the formative lags vary but slowly with variation of Y_o. It was not possible to make any measurements of the initial ionization, with the apparatus used, and therefore estimates of Y_o were made using the theory of the previous chapter. The critical parameter is φ and it has proved possible to make good estimates of this quantity in the following manner.

From the definition of φ (equation 8.8) it can be seen that we need to know both the ionization frequency and diffusion coefficient, which involve gas discharge data, and also the diffusion length which is a simple function of the gap geometry. The diffusion coefficient, D, of electrons is well known and is related to the electron mobility μ by the relation[28]

$$D/\mu = (2/3)\ (\text{average electron energy})$$

In this expression the average electron energy can be found from Townsend's factor η by multiplying by the average energy of the gas molecules. The value of electron drift velocity has been determined by several workers and is given as a function of E/p. Since $v_{drift} = \mu E$, approximately, we have, by substitution

$$Dp = (2/3)\bar{u}v_{drift}(E/p)^{-1} \qquad \ldots\ldots (8.9)$$

where p denotes the gas pressure

$\bar{u}$ is the average electron energy

E is the electric field strength

v_{drift} is the electron drift velocity.

The rate of ionization, ν_i, is given by

$$\nu_i = \alpha v_{drift} \qquad \ldots\ldots (8.10)$$

where α is Townsend's first ionization coefficient. Values of α/p have been published for many gases as a function of E/p and semi-empirical formulae exist[29] which give α/p in terms of atomic constants over limited ranges of E/p (mainly E/p > 100 volt/cm/torr). In the high frequency discharges observed in these experiments E/p was quite low, and under these conditions α/p varies very rapidly with E/p so that it was considered unwise to utilize published results in quantitative calculations. Therefore the value of α was calculated from the observed

values of breakdown under the influence of sustained high frequency fields, at zero overvoltage.

For "infinite" delay, $\varphi = 0$ so that from equations 8.8 and 8.10 we have

$$\alpha v_{drift} = D/L^2$$

and using equation 8.9 to eliminate D,

$$(\alpha/p)p_s E_s = 2\bar{u}/3L^2 \qquad \ldots\ldots 8.11$$

where the subscript s indicates values for infinite delays.

Now for the overvolted gap, substituting for ν_i and D from equations 8.10 and 8.9 respectively in equation 8.8, we have

$$\varphi = \alpha v_{drift} - (2/3)\bar{u}\, v_{drift}\, (E/p)^{-1} L^{-2}$$

Further, substituting for α from equation 8.11 we find

$$\varphi = \frac{2}{3} \frac{\bar{u}\ v_{drift}}{EL^2} \left[\frac{pE}{p_s E_s} - 1 \right] \qquad \ldots\ldots 8.12$$

The values of $\bar{u}$ and v_{drift} are to be determined for the actual applied value of E/p. E and p are also the actual values. E_s and p_s are not obvious in their interpretation. They arise as a result of substituting without first calculating an explicit value for α/p. If α/p had been first evaluated and then used in the expression for φ, then it would have been plotted as a function of E/p, using the values for breakdown under sustained fields, and its value determined from E/p applied as an overvoltage. Therefore we

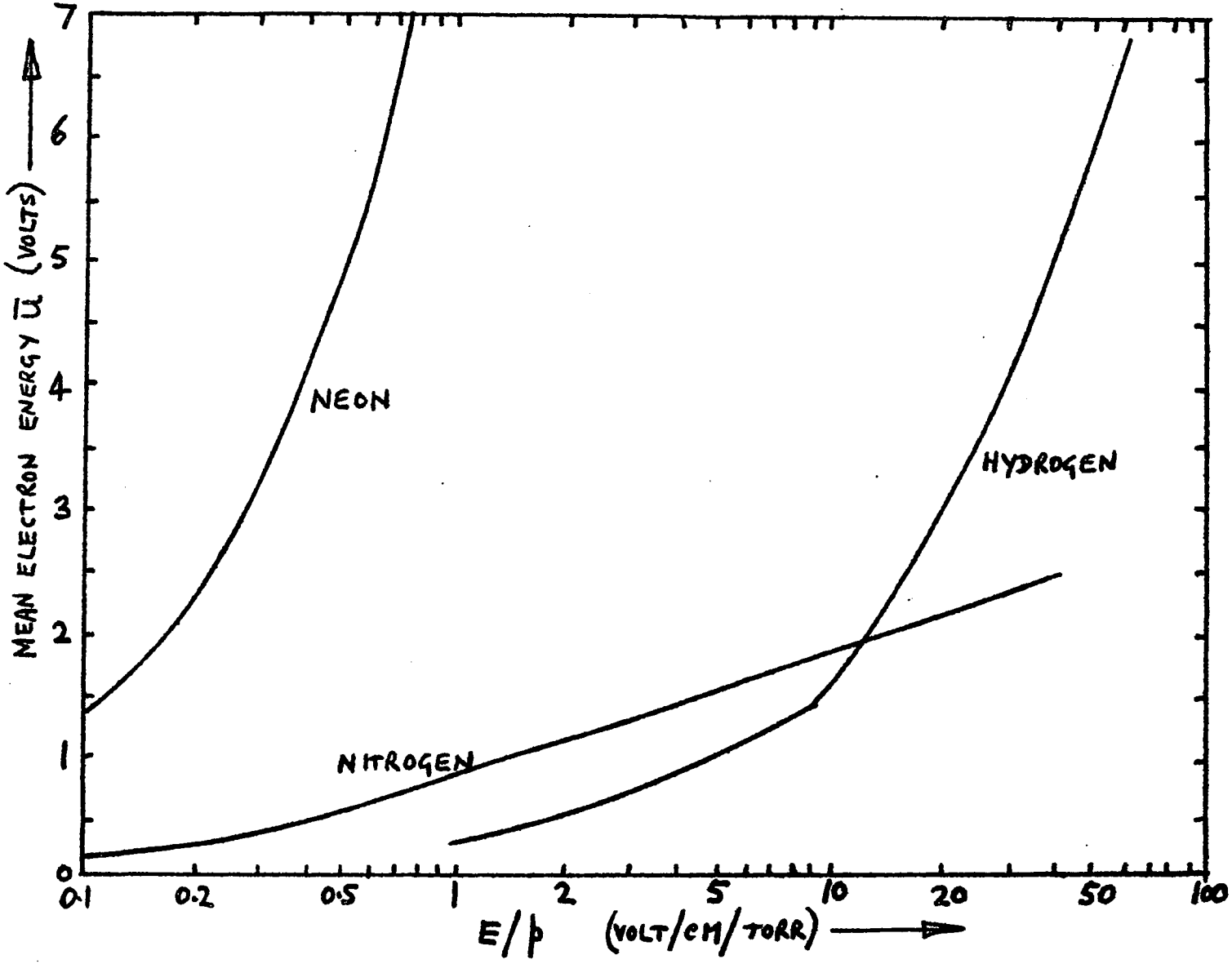


FIG 8.4 THE MEAN ELECTRON ENERGY AS A FUNCTION OF E/p IN HYDROGEN [30] NITROGEN [30,32] AND NEON [31]

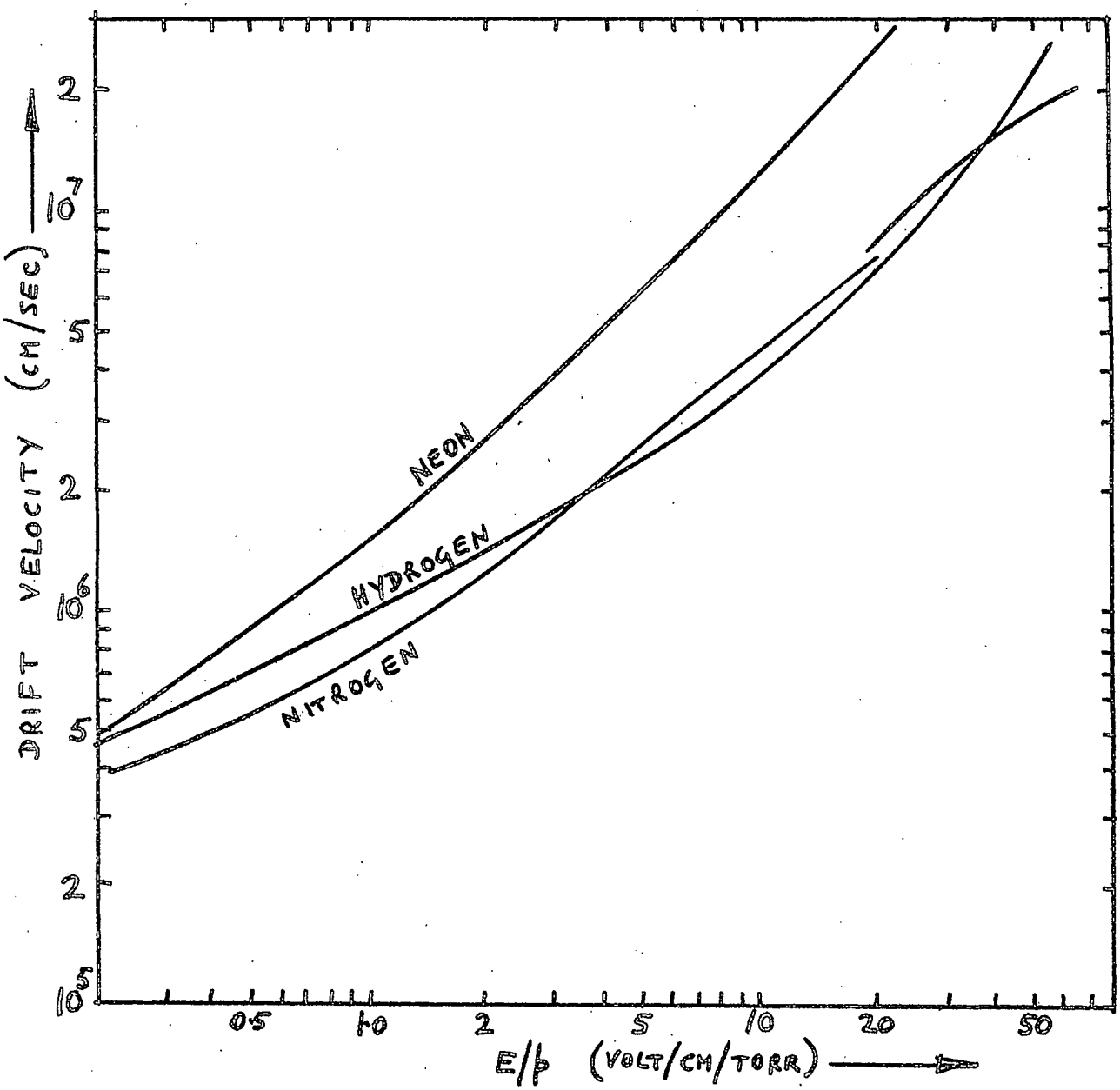


Fig 8.5 DRIFT VELOCITIES OF ELECTRONS IN HYDROGEN[30,34] NITROGEN[30,35] AND NEON[35]

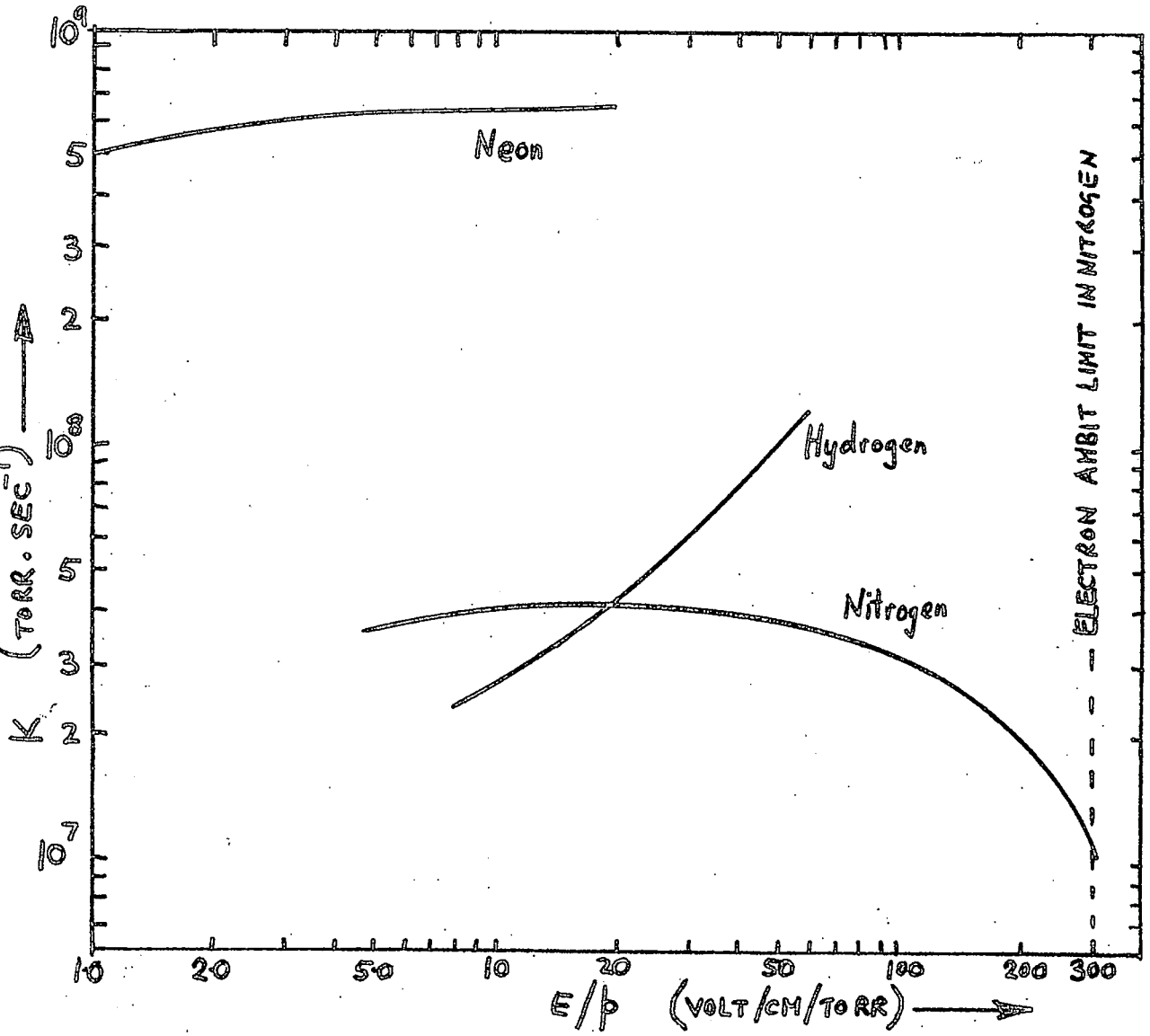


FIG 8.6 THE FUNCTION K OF EQUATION 8.14 AS DEFINED BY

$$K = \frac{2\bar{u}\, v_{drift}}{3L^2 (E/p)}$$

can see that E_s and p_s must be chosen such that

$$\frac{E_s}{p_s} = \frac{E}{p}$$

If now we substitute p_s from this equation into equation 8.12 we find

$$\varphi = \frac{2\bar{u}\, v_{drift}}{3\, EL^2} \left[\frac{E^2}{E_s^2} - 1\right] \qquad \ldots\ldots 8.13$$

Published values of $\bar{u}$ and v_{drift} are given in figs. 8.4 and 8.5 respectively.

It will be observed that φ is not a function of E/p only and therefore must be determined separately for each pressure used.

If we define a function K such that

$$K = \frac{2\bar{u}}{3L^2} \cdot \frac{v_{drift}}{E/p} \qquad \ldots\ldots 8.14$$

then equation 8.13 can be written

$$\varphi = \frac{K}{p} \left[\frac{E^2}{E_s^2} - 1\right]$$

The values of K, as a function of E/p, calculated from the data of figs 8.4 and 8.5 are shown in fig 8.6. All that is now required before the breakdown delays can be predicted is an experimental determination of E_s.

It has been assumed here that φ remained constant as the discharge evolved, although strictly ν_i decreases with the gap voltage while the diffusion coefficient D decreases

towards the ambi-polar value as the electron density rises. These effects introduce little error until the voltage has decreased appreciably, and therefore calculated values of formative delay would probably be correct, but the errors must become significant by the time the voltage has dropped to about 80% of the applied voltage. It should however be noticed that since φ is formed from the difference of two terms, and both terms decrease with gap voltage, therefore φ has a tendency to remain approximately constant as the voltage decreases.

CHAPTER IX

EXPERIMENTAL MEASUREMENTS OF BREAKDOWN FIELDS.

9.1. Preliminary observations

As the spark gap was connected across an unterminated transmission line, it was possible for a steady potential to exist between the two conductors. Such a potential could possibly be generated by one discharge and remain long enough to inhibit succeeding discharges. A measurement was therefore made of the time which had to elapse before the initial breakdown voltage was again observed after the central conductor had been charged to a potential of about 10 volts. When the line discharged through the insulators which supported the inner conductor, this time was found to be of the order of 10 minutes, and was reduced (to about 10 microseconds) by connecting a 120Kohm resistor across the line. This was carried out with negligible reduction of the resonant quality of the line by making the connection at a position of low r.f. potential.

Preliminary experiments with pulsed r.f. fields showed that there was a definite probability of breakdown occurring before the irradiator flash, if daylight was allowed to fall upon the electrode capsule. The probability of early breakdown increased with overvoltage and was particularly noticeable in hydrogen and nitrogen, and only to a minor extent in neon. As many as 50% of the hydrogen discharges

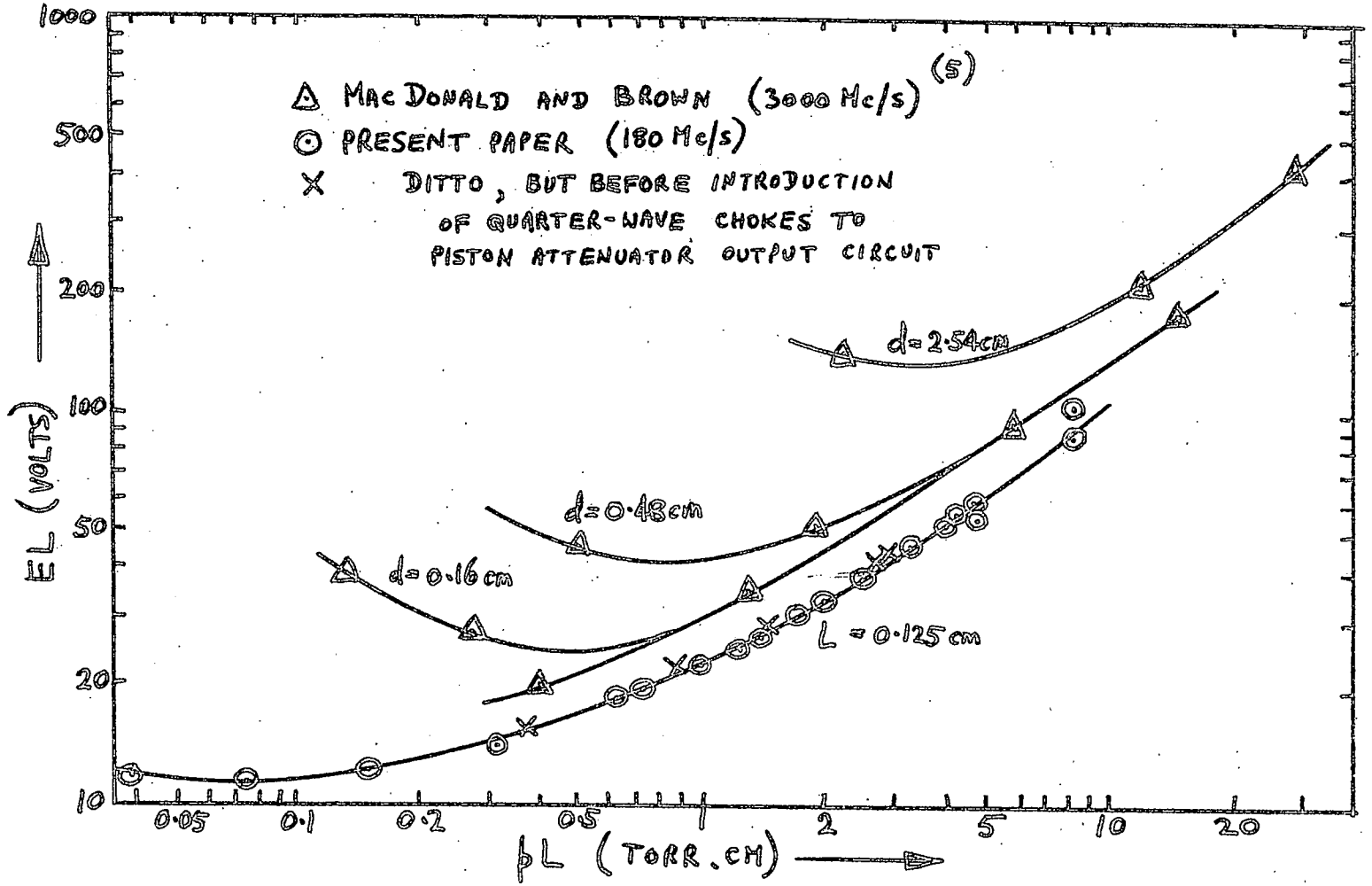


FIG 9.1 BREAKDOWN VOLTAGES IN HYDROGEN UNDER SUSTAINED HIGH FREQUENCY FIELDS

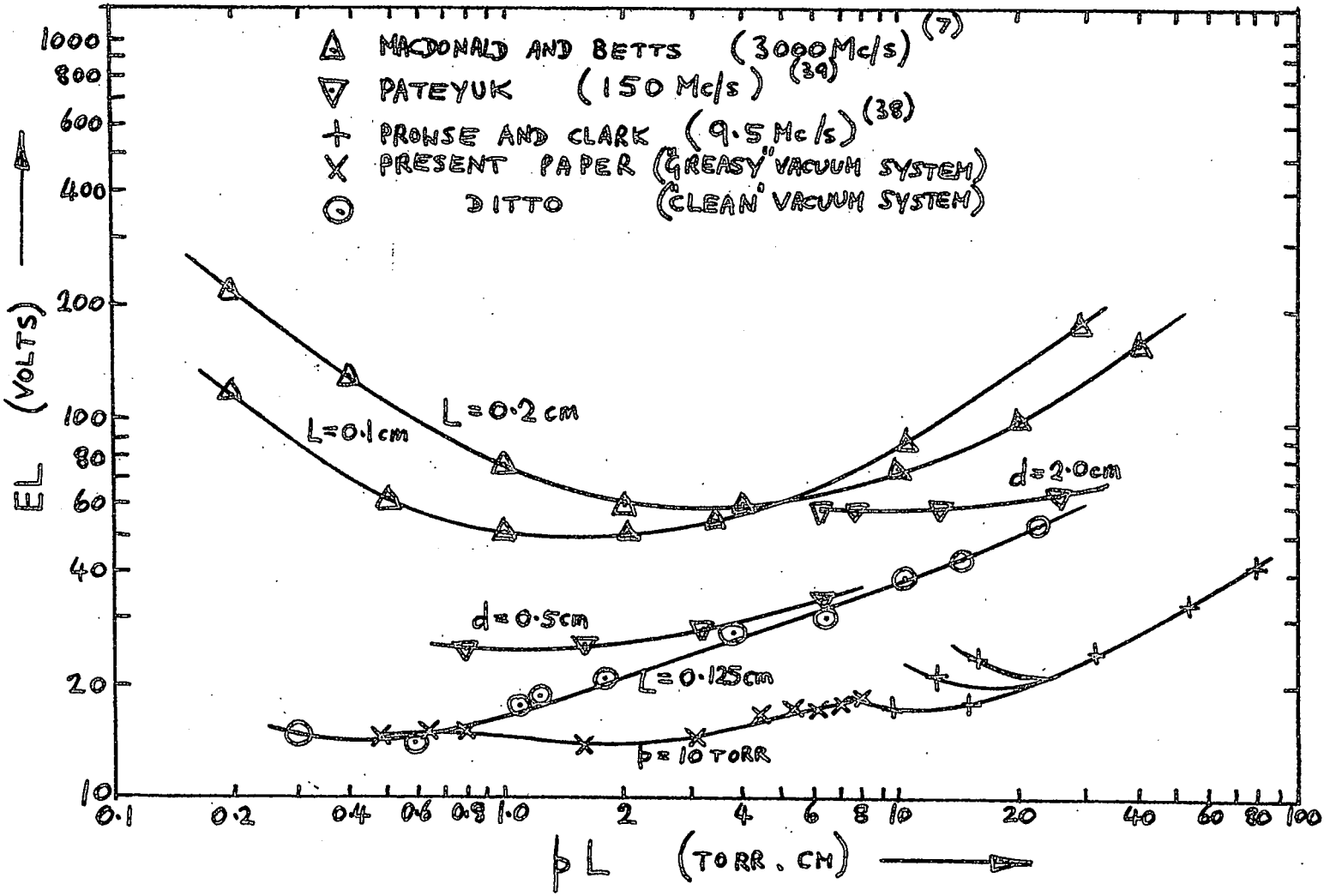


FIG 9.2 BREAKDOWN VOLTAGES IN NEON UNDER SUSTAINED HIGH FREQUENCY FIELDS

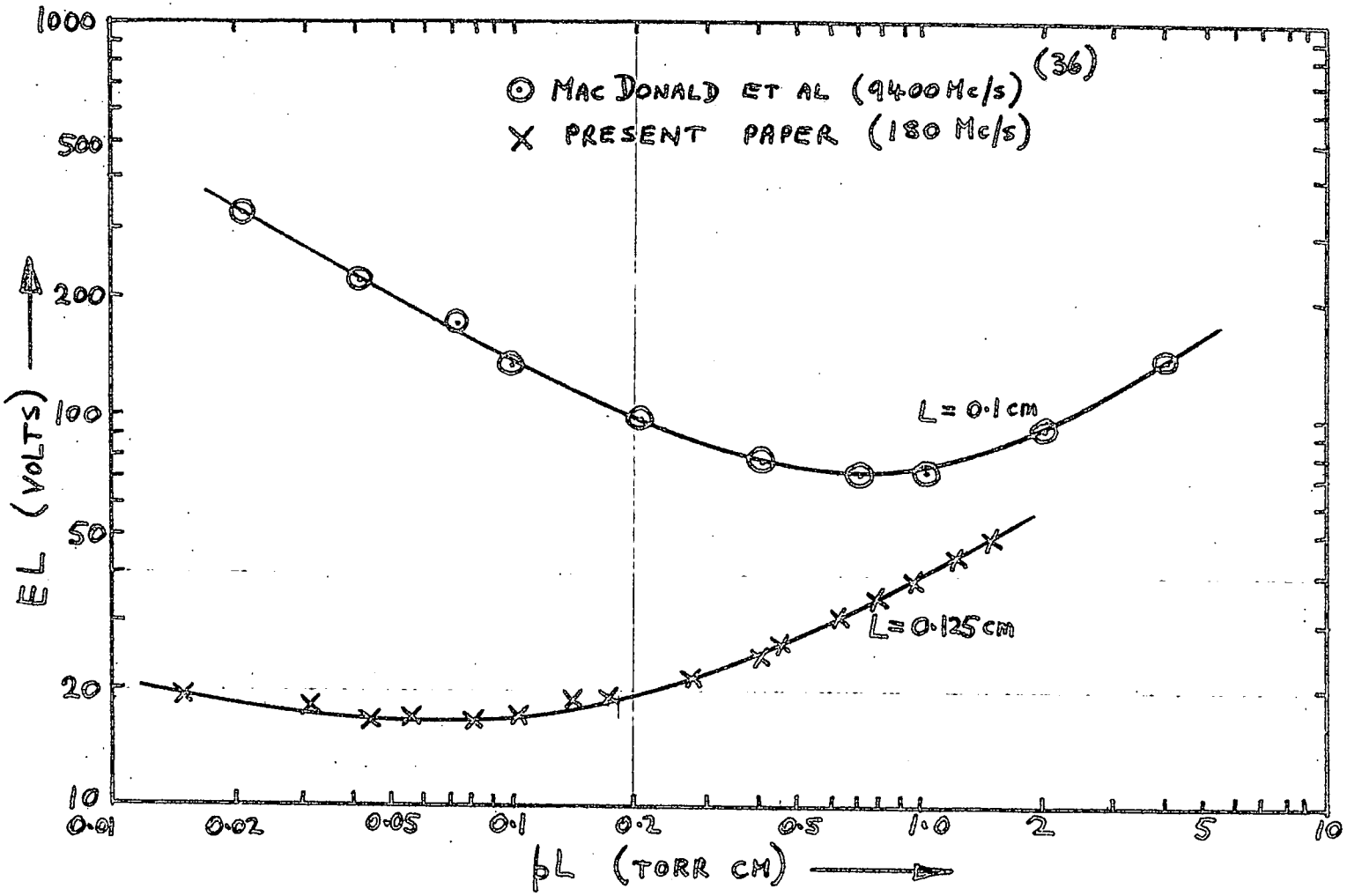


FIG 9.3 BREAKDOWN VOLTAGES IN NITROGEN UNDER SUSTAINED HIGH FREQUENCY FIELDS

were found to be spuriously initiated in some cases. The effect was made negligible (less than about 1%) by wrapping the discharge chamber and the upper part of the tube to the irradiator with black cloth. All the pulsed-field measurements quoted here were made under these conditions, but not the measurements made with sustained fields.

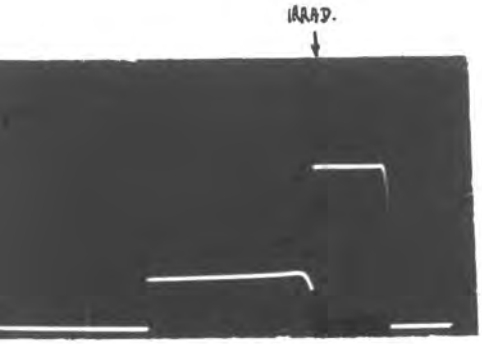
9.2. Breakdown under sustained high frequency fields.

Figs 9.1, 9.2, and 9.3 show the values of r.m.s. breakdown fields obtained with hydrogen, neon and nitrogen. The frequency used in these experiments was, as has been previously indicated, always 183 Mc/s and the gap width was fixed at 0.494cm. The diameter of the flat central part of the electrodes was 1.0cm and therefore by equation 8.6 the characteristic diffusion length, L, was 0.125cm.

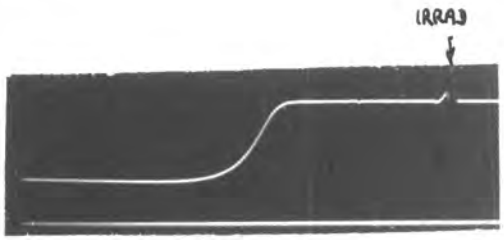
The results are presented in terms of the variables pL and EL. It can be shown[36] that breakdown may be expressed uniquely by these variables for the special case of gases in which the electron collision frequency is independent of its energy (i.e. constant mean free time). Hydrogen is such a special case, but in neon the collision frequency is directly proportional to the electron velocity (i.e. constant mean free path) for energies less than 60eV. In nitrogen the collision frequency is variable for energies below about 25eV, but above this level it becomes roughly constant.

For hydrogen, fig. 9.1, the agreement between the present observations and results published by MacDonald and Brown[5] (at microwave frequencies) and Githens[37] (at several megacycles per second) is fair, the presently measured values being some 20% lower than the published values.

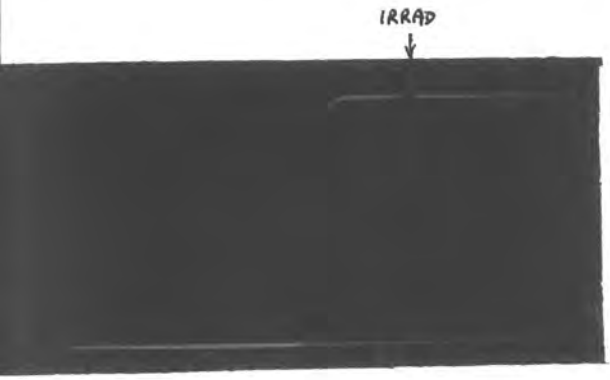
The breakdown fields observed with neon are compared with published values in fig 9.2. Wide divergences in the results of various workers are evident. Although, as indicated above, it is not necessarily expected that the EL, pL variables would give a unique curve in neon, Prowse and Clark[38] using several different gas pressures and a variable gap width at a frequency 9.5Mc/s have found such a unique relationship for $pL \geqslant 20$ torr.cm. At 158Mc/s Pateyuk[39] obtained results for neon in fair agreement with those measured here, whilst the breakdown fields of MacDonald and Betts[7] at 3000Mc/s are a factor 2 higher. Prowse and Clark observed relatively low fields, and indicated that their results could correspond to those to be expected from neon containing one part in 10^4 of argon. The cavities of MacDonald and Betts could be outgassed at a higher temperature than our vacuum system, and therefore it seems likely that they achieved a higher gas purity than that obtained here. It therefore would appear that the divergences could be due to the effects of gaseous impurities. In this respect it is



(a) Breakdown in hydrogen
p 32 torr
E 585 volt/cm.
total pulse duration: 500 microseconds.



(b) Breakdown in hydrogen
p 10 torr
E 230 volt/cm.
breakdown delay: 3.6 microseconds.



(c) Breakdown in neon.
p 174 torr
E 765 volt/cm.
breakdown delay: 1.6 microseconds.

Fig 9.4. Some typical breakdown oscillograms observed with hydrogen and neon.

of interest to note some results quoted in fig 9.2 which we obtained with this apparatus before the stopcock grease was eliminated. These show a similarity to the observations of Prowse and Clark whose gas handling equipment was also greasy and not baked.

Measurements of breakdown fields in nitrogen are presented in fig 9.3 where the values given by MacDonald et al[36] are also shown for comparison. At the high pressure ends of the scale, both the present and MacDonald's curves assume the same slope, but MacDonald's values are some 50% higher.

9.3. Breakdown with pulsed fields.

Some typical oscillograms are presented opposite, which depict breakdown under conditions of medium overvoltage. The top photograph, fig 9.4(a), shows a breakdown in hydrogen. The oscillator was switched on at t=0 and the irradiator was fired at the point indicated, when the gap voltage had reached a steady value. The voltage remained almost constant for a while before falling rapidly to a lower steady value, the maintaining voltage. The voltage then fell to zero when the oscillator was switched off. The second picture, fig 9.4(b), is of a similar discharge but the timebase speed was 100 times faster and was triggered only a few microseconds before the irradiator. Although in these photographs the irradiator marker cannot be clearly discerned,

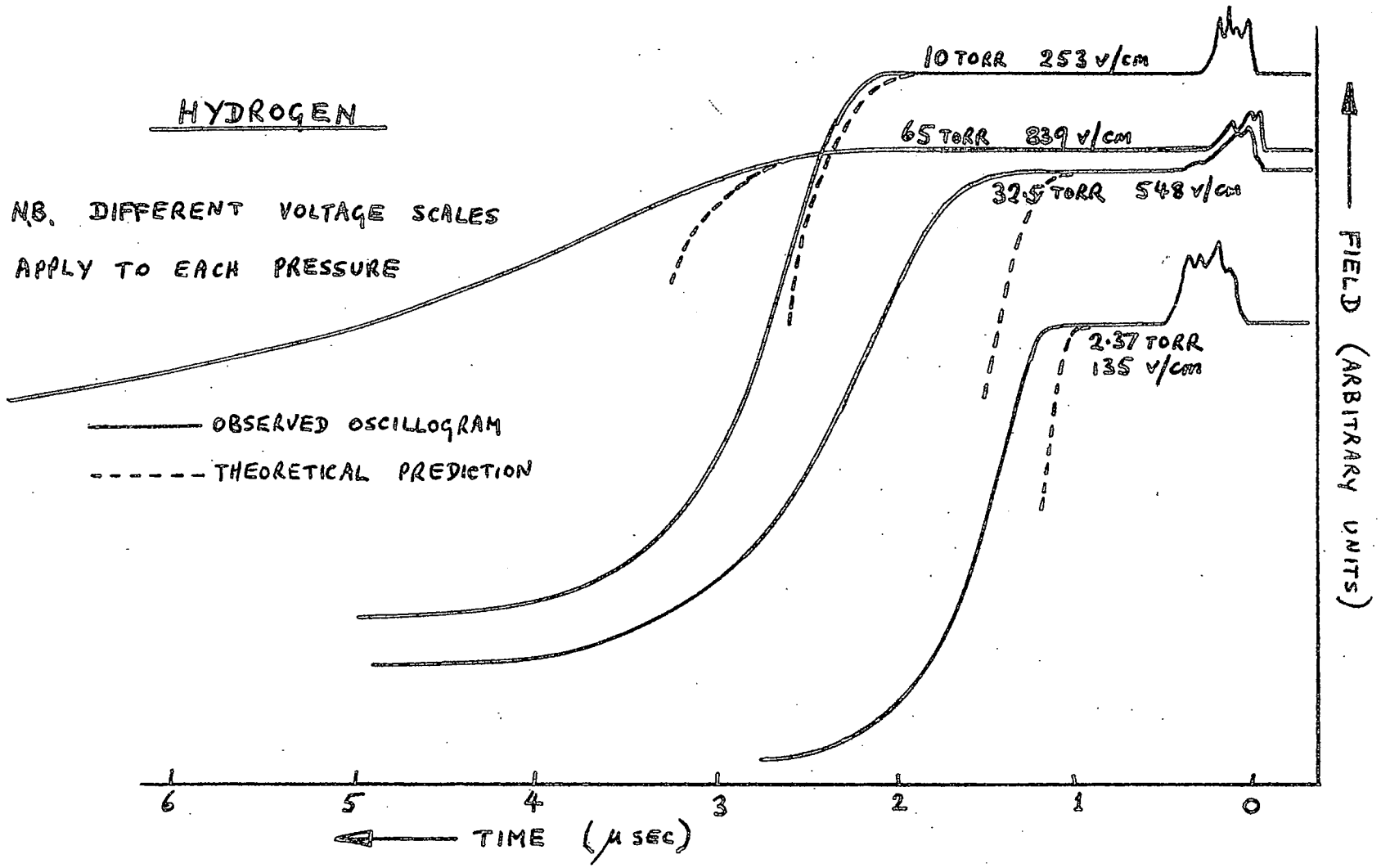


FIG 9.5 THE VARIATION OF BREAKDOWN OSCILLOGRAM WITH THE GAS PRESSURE

it was quite distinct in the original negatives upon which the measurements were made.

Oscillograms of discharges in nitrogen were very similar to those of hydrogen. Neon discharges were characterised by a very low maintaining voltage, which was, in fact, too small to measure with this apparatus, the deflection on the oscilloscope being less than the trace width. An oscillogram of a discharge in neon is shown in fig 9.4(c). This was obtained using the faster timebase again.

The transition from the non-conducting state to the establishment of a steady maintaining voltage often proceeded via an intermediate stage during which the gap voltage was very low. This "overshoot" of voltage can be clearly seen in fig 9.4(a). The rapid decrease of voltage from the initial state continued to a minimum value after which the voltage rose comparatively slowly, and approached asymptotically to the maintaining voltage.

9.4. Comparison of observed oscillograms with theory.

A selection of breakdown oscillograms is given in fig 9.5 for various pressures of hydrogen. It should be noted that the time axis represents zero voltage and that the voltage scale in this figure is different for each of the traces. Fig 9.6 shows how the oscillograms varied with overvoltage at constant pressure (10 torr). The full lines are tracings of experimental oscillograms and the dotted

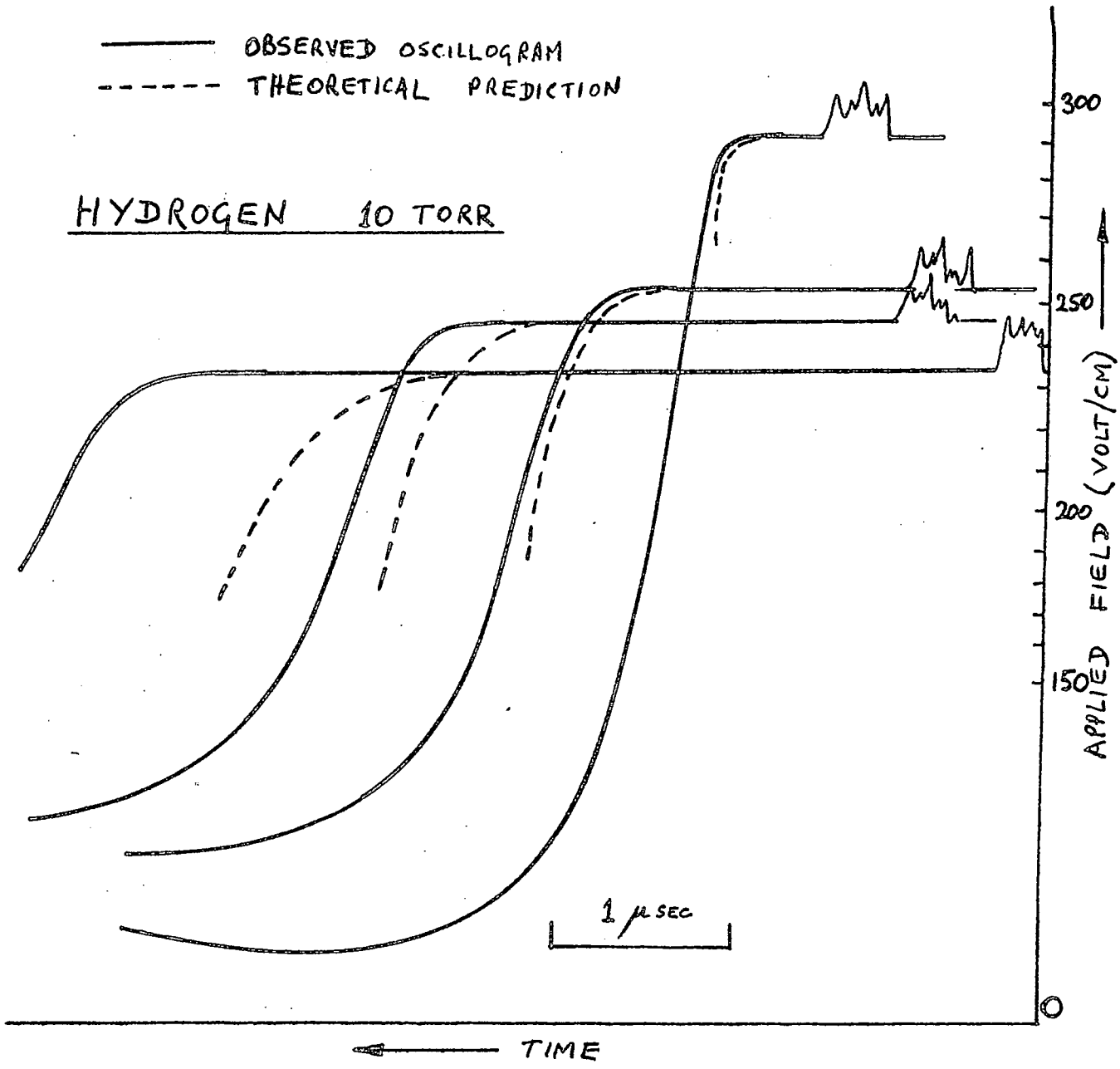


FIG 9.6 THE VARIATION OF THE BREAKDOWN OSCILLOGRAMS WITH THE APPLIED FIELD STRENGTH

curves were obtained by using the theory of the previous chapter. The ragged pulses superimposed on the traces at the right-hand ends are the irradiator markers, representing time zero.

It may be seen that the shapes of the theoretical and experimental oscillograms agree very well above the point of maximum slope on the experimental curves. Below this level the experimental curves become less steep as they tend towards the maintaining voltage. Agreement with theory would not be expected in this region because of the neglect of the variations of the net ionization rate, φ , with both gap voltage and electron concentration. The comparison here is essentially between the shapes of the theoretical and experimental oscillograms, rather than between the total breakdown delay times. This is partly because the statistical delay, which occurs at the beginning of every discharge, must be added to the formative time before agreement between the total delays could be expected. Further, the theoretical value which was assumed for the initial number of electrons may not have been correct. This number, N_o, was determined approximately by first estimating a value of mean absorption coefficient for the irradiator photons, for each pressure used. Then using fig 7.1, a "reasonable" estimate of N_o was made, between the limits $N_o = 100$ and the value given by the curve for $I_o = 10^7$. This "reasonable estimate could be

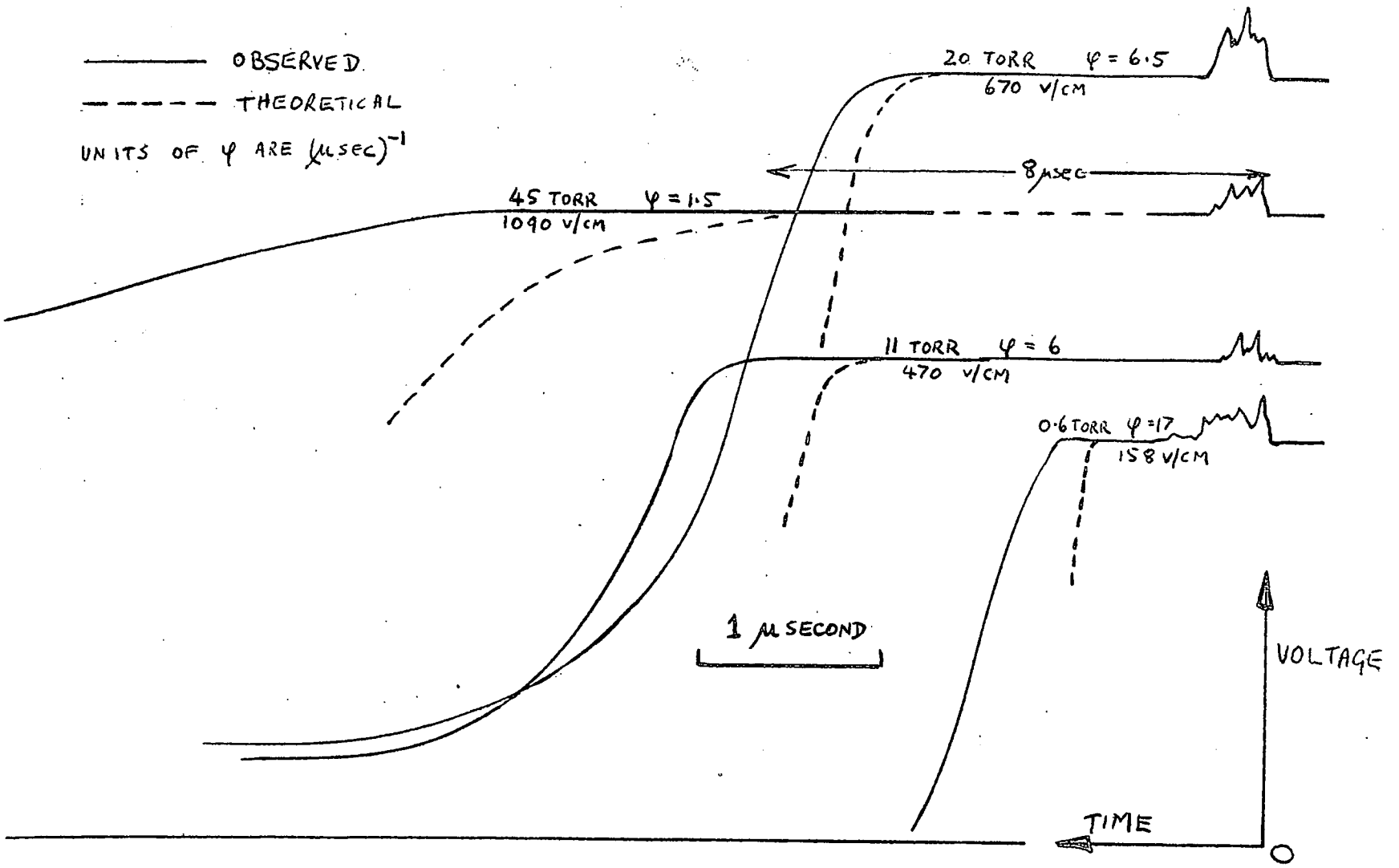


FIG 9.7 BREAKDOWN OSCILLOGRAMS IN NITROGEN: THE VARIATION WITH PRESSURE

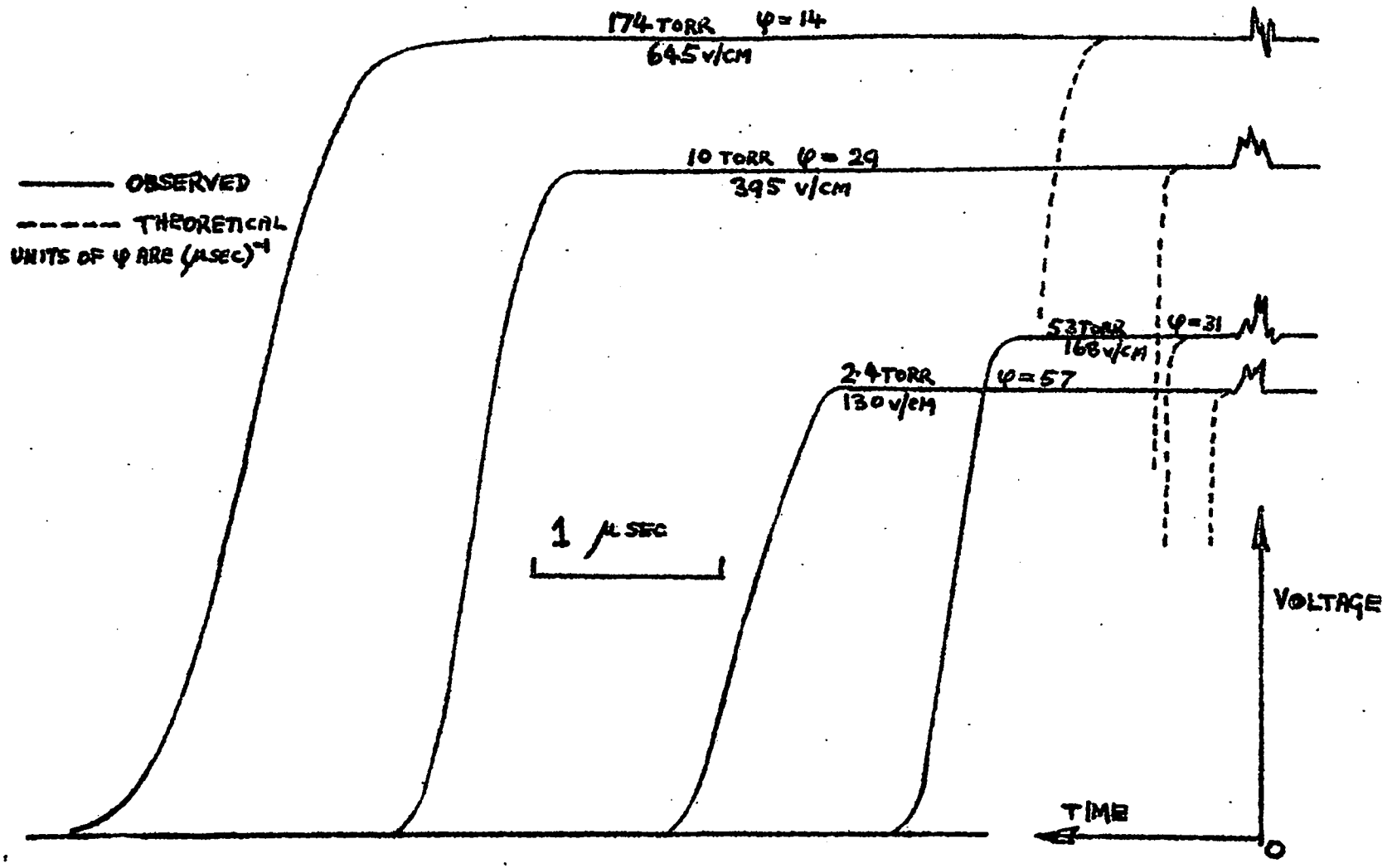


FIG 9.8 BREAKDOWN OSCILLOGRAMS FOR NEON : THE VARIATION WITH PRESSURE

in error by a factor probably not exceeding 100. To see the effect this magnitude of error would have on the oscillograms of fig 9.5, suppose for example the net ionization rate, φ , is 10 per microsecond. Assuming the growth of electron population is given by

$$n = n_o \exp(\varphi t).$$

then the time Δt taken for the population to grow by factor 100, which represents the uncertainty in the predicted delay time, would be given by

$$\Delta t = \pm \varphi^{-1} \ell n (n/n_o) \qquad \ldots\ldots 9.1.$$

$$= \pm 0.5 \text{ microsecond.}$$

This is greater than the error in the predicted delay for any of the curves of fig 9.5 and lies between 10 and 25% of the total delays.

Fig 9.7 shows a sample of the breakdown oscillograms obtained with nitrogen. One example is given of each pressure used with delays ranging from 1 μsec to 10 μsec. The maximum field that could be obtained (1150 volt/cm) did not give delays shorter than 8 μsec with the maximum pressure (45 torr). The values of φ appropriate to each oscillogram are given in the figure, and if these are substituted in equation 9.1 it can be seen that agreement is obtained to this accuracy between the theoretical and experimental formative delays.

In Fig 9.8 are some oscillograms obtained for breakdown in neon at various pressures. The breakdown delays as calculated

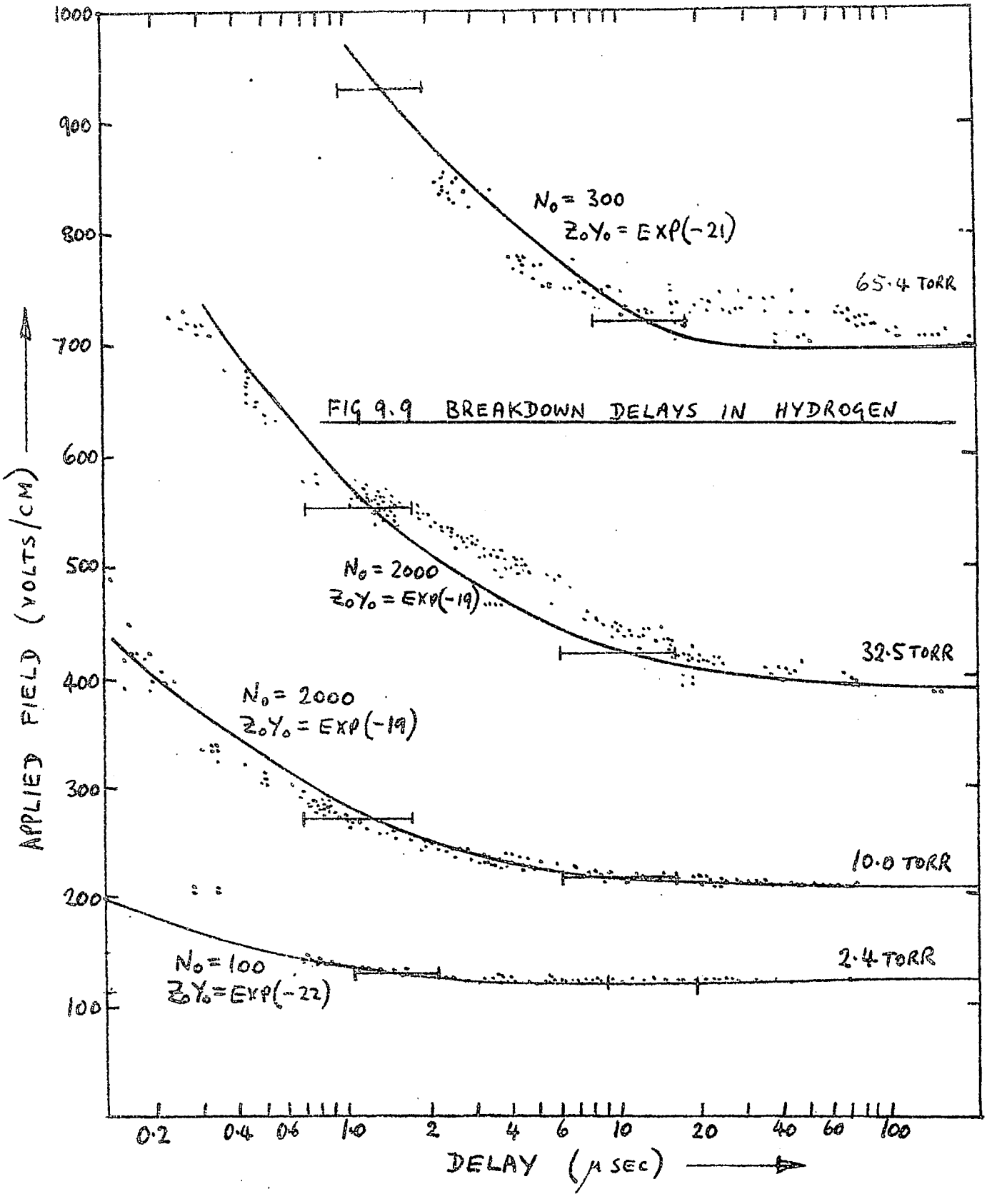
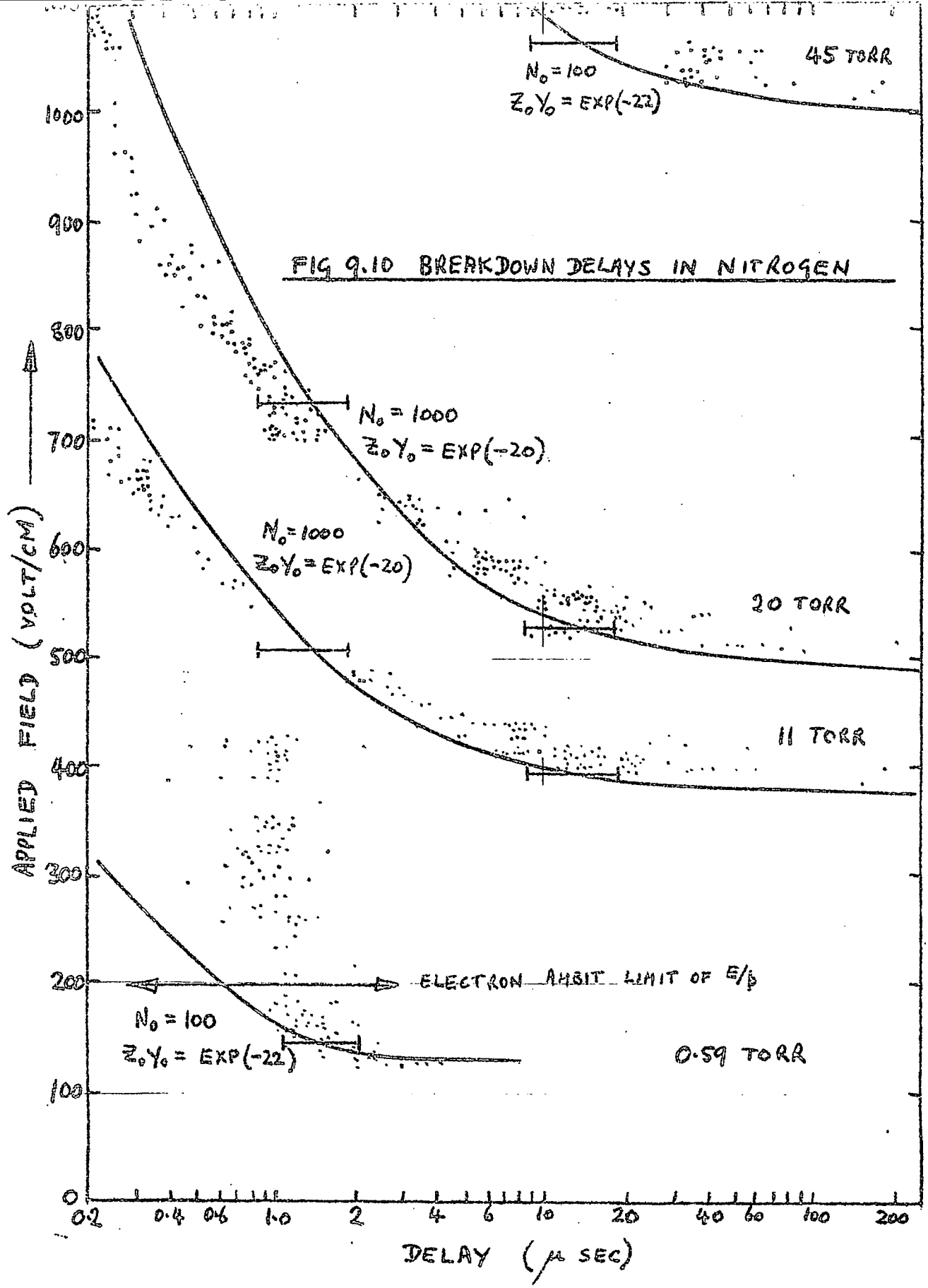


FIG 9.9 BREAKDOWN DELAYS IN HYDROGEN


45 TORR
$N_0 = 100$
$Z_0 Y_0 = EXP(-22)$
FIG 9.10 BREAKDOWN DELAYS IN NITROGEN
$N_0 = 1000$
$Z_0 Y_0 = EXP(-20)$
20 TORR
$N_0 = 1000$
$Z_0 Y_0 = EXP(-20)$
11 TORR
ELECTRON ABSIT LIMIT OF E/p
$N_0 = 100$
$Z_0 Y_0 = EXP(-22)$
0.59 TORR
APPLIED FIELD (VOLT/CM)
1000
900
800
700
600
500
400
300
200
100
0
0.2
0.4
0.6
1.0
2
4
6
10
20
40
60
100
200
DELAY (μ SEC)

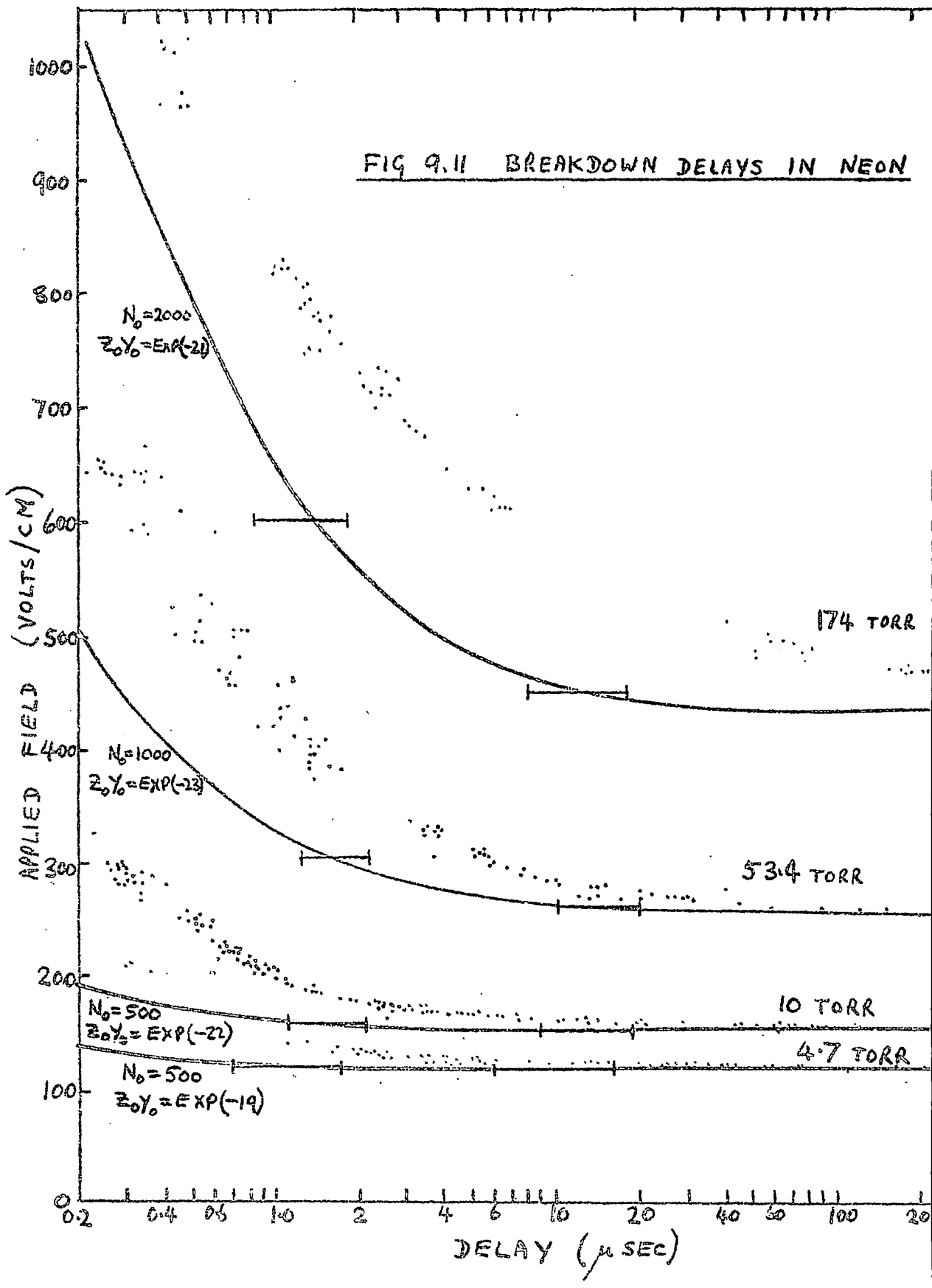


FIG 9.11 BREAKDOWN DELAYS IN NEON
APPLIED FIELD (VOLTS/CM)
DELAY (μ SEC)
1000
900
800
700
600
500
400
300
200
100
0
0·2
0·4
0·6
1·0
2
4
6
10
20
40
60
100
200
N_0=2000
Z_0Y_0=EXP(-21)
174 TORR
N_0=1000
Z_0Y_0=EXP(-23)
53·4 TORR
N_0=500
Z_0Y_0=EXP(-22)
10 TORR
N_0=500
Z_0Y_0=EXP(-19)
4·7 TORR

from the theory of Chapter 8 are too short at all pressures used, by factors ranging from 2.5 up to about 9. Such magnitudes of error are not accounted for by the simplifying assumptions made in the theory or in the estimation of N_o: in these calculations N_o was taken to be about 1000, but even putting $N_o=1$ would only increase the theoretical delays by a factor 1.6.

Figures 9.9, 9.10, and 9.11 show the theoretical and measured delays as a function of applied voltage for four pressures in each of the gases hydrogen, nitrogen and neon. Between 100 and 200 oscillograms were measured for each curve and plotted as points in these figures. With the exception of the results for the lowest pressure of nitrogen, all the curves bear a general resemblance to one another. This exception is readily explained in terms of the electron ambit becoming comparable to the spark gap width. The value of E/p for this condition has been calculated in Appendix 3 and is marked on the curve with a horizontal bar. A strikingly abrupt lower limit to the delays occurs for E values above this bar, where the increased losses of electrons cause relatively long delays. On the diagrams for hydrogen and nitrogen the agreement between observation and theory is generally quite good. It was assumed that between 100 and 1000 electrons initiated the discharges, but the modification to the theoretical curves caused by a hundredfold increase

or decrease in these numbers was calculated and is shown on the graphs by horizontal bars at value of E which give φ=10 and φ=1

9.5. Discussion

An examination of these results for hydrogen immediately shows the reason for formative delays not being observed by Prowse Rowbotham and Monk[9]: it is recalled that the resolution of their voltage measurement was about 2% and their minimum observable delay was 10 microseconds. The curves of fig 9.9 show that except for the highest pressures used, a 2% voltage increase above the minimum for breakdown gives a change in the delays from "infinity" to less than 10 microseconds. Further, at higher pressures, where delays might have been observed, there was insufficient voltage available at the test gap to obtain breakdown.

As was stated in the previous section it seems that the disagreement between theoretical and observed delays in neon is too large to be accounted for by errors in any of the variables used in the theory. Since the theory put forward so far for the case of neon was the same as that for hydrogen and nitrogen for which good agreement was obtained, it appears probable that breakdown in neon proceeded by a different mechanism. Below are listed some of the ways in which neon essentially differs from hydrogen and nitrogen:

(a) It is monatomic

(b) The resonance and ionization potentials are comparatively large (16.8 volts and 21.6 volts respectively)

(c) There is a high potential associated with a metastable state (16.6 volts)

(d) Being chemically inert it is comparatively difficult to obtain free from other inert gas impurities

The essential difference between monatomic and diatomic molecules is that whereas the monatomic molecule has three degrees of freedom which are all translational, the diatomic molecule has five, the extra two degrees of freedom being represented as rotational and vibrational states. These require for their excitation an energy much less than that required for atomic excitation. Thus electrons can lose small amounts of energy to diatomic molecules, but require relatively high energies to produce an inelastic collision with an atom. Therefore the electron energy distributions for monatomic and diatomic gases differ considerably. However, the methods by which delays are calculated here might be expected to eliminate serious errors due to this effect. One way in which the two types of gases could be expected to differ is in the time taken for the electron distribution to come to equilibrium. Von Engel[29] has given a formula for the time constant, τ, of the approach

to equilibrium: (this is based on simple theory, but may be used to indicate orders of magnitude)

$$\tau = 0.97\ (m/e)^{\frac{1}{2}}(1/k)^{\frac{3}{4}}(\ell/E)^{\frac{1}{2}}$$

where e/m is the ratio of electronic charge to mass,

k is the average fraction of its energy lost by an electron at a collision, and

ℓ is the electronic mean free path.

Using values for ℓ and k for the average electron energy appropriate to the value of E/p, we find that τ ranges between 2 and 5 nanoseconds for all the pressures and voltages used in all three gases. As τ is thus almost constant for all three gases and also very much less than the least observed delay (~200 nanoseconds) it seems unlikely that this effect could account for the disagreement between theory and observation for neon. (40)

(b) and (c) ensure that impurities have a catastrophic effect on the breakdown voltage, since most impurity atoms can become ionized by electrons of lower energy than that required to ionize a neon atom.

A particularly important impurity in neon is argon, which has an ionization potential (15.8 volts) slightly lower than the metastable potential of neon. As a result, metastable neon atoms can ionize argon atoms (Penning effect) (41) and breakdown occurs at a considerably smaller voltage than for pure neon. One part of argon in 10^5 of neon can cause a

50% reduction in the breakdown field[42] for neon in some circumstances.

The Penning effect can give a qualitative explanation of the anomalous delays in neon in the following manner: let us consider the balance which just suffices to give breakdown (infinite delay). Electrons are being lost continuously to the electrodes by diffusion and are being replaced by the following two stage process:

(a) $e + Ne \rightarrow e + Ne^{*}$

(b) $Ne^{*} + A \rightarrow A^{+} + Ne + e$

where Ne^{*} denotes the metastable state of the atom

In the infinitesimally slow approach to breakdown, argon ions can be neutralized at the electrodes and replaced by neutral atoms diffusing back into the spark gap. If now the gap is overvolted, the electron production rate increased due to the reactions

(c) $e + Ne \rightarrow Ne^{+} + e + e$

(d) $e + Ne^{*} \rightarrow Ne^{+} + e + e$

becoming faster, while diffusive losses increase relatively slowly. Further, in the shorter time which elapses before breakdown, the argon ions have insufficient time to be neutralised, so that the Penning effect is diminished. In the calculations of breakdown delays, the ionization rates were calculated from the breakdown condition for sustained fields. If now it is admitted that the Penning effect was

the mechanism causing breakdown at these field values, then it can be seen that a different process (such as (c) and/or (d)) could be operating for overvoltages. This would require greater electron energies than the Penning effect and would therefore give rise to the relatively long delays observed for a given overvoltage.

CHAPTER X

THE MAINTAINED DISCHARGE

10.1. The appearance of the discharge.

Most dramatic changes were observed in the discharges as the gas pressure was varied. Increasing the power input to the spark gap beyond that required for breakdown produced little effect apart from increasing the brightness slightly.

The colours of the discharges were very similar to those of the positive column of a d.c. glow discharge: a pale blue in hydrogen, and a pale pinkish blue in nitrogen, while in neon the colour gradually changed with increase of pressure, from orange to a brick-red.

At pressures less than 0.1 torr the glow in all the gases completely filled the discharge vessel except for the region between the electrodes. At about 0.2 torr the discharge took place between the electrodes, but also covered the sides of the electrodes and spread laterally so far as to touch the glass walls of the capsule. Increasing the pressure from 0.4 to 1.0 torr caused the reduction of the volume of the glow to the region immediately between the electrodes. At this stage two distinct regions could be discerned which produced most of the visible light. These regions were disc-shaped and symmetrically disposed, parallel to and close to each electrode. As the pressure

was further increased both the thickness and the area of these luminous discs decreased as they approached closer to the electrodes. The space between the bright discsglowed rather faintly, but between the electrodes and the luminous discs there was a space about 0.1mm thick which appeared to be completely non-luminous. Above about 20 torr a third luminous region could be observed in the centre of the gap, less bright than the other two, and more diffuse. In hydrogen and nitrogen this central glow divided into two more parallel disc-like regions at about 30 torr but in neon the one central glowing region remained up to the highest pressure used, (174 torr).

At pressures greater than 10 torr the edges of the bright discharge region were sharply defined. However, careful observation of discharges in nitrogen revealed a very faint blue glow surrounding the main discharge and extending somewhat beyond the radius of the electrodes.

A feature at pressures above about 10 torr was the movement of the discharge column. This effect appeared to be quite spontaneous and uncontrollable. With nitrogen and hydrogen the discharge column moved bodily around the electrodes, sometimes describing circles concentric with the spark gap axis, and at other times performing more-or-less random movements or remaining stationary. The circular movement would vary in speed up to about 10 revolutions per

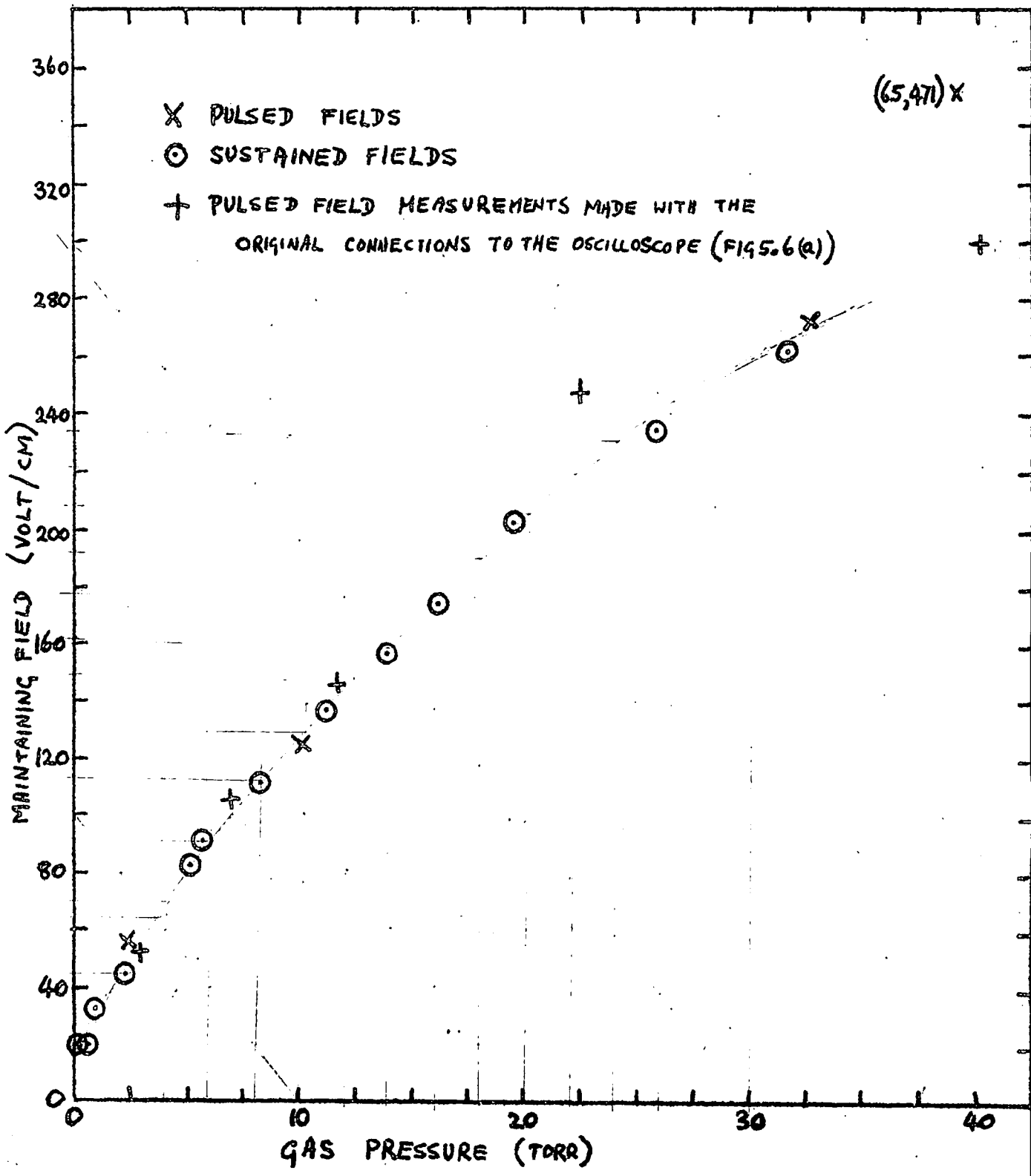


FIG 10.1 MAINTAINING FIELDS IN HYDROGEN DISCHARGES.

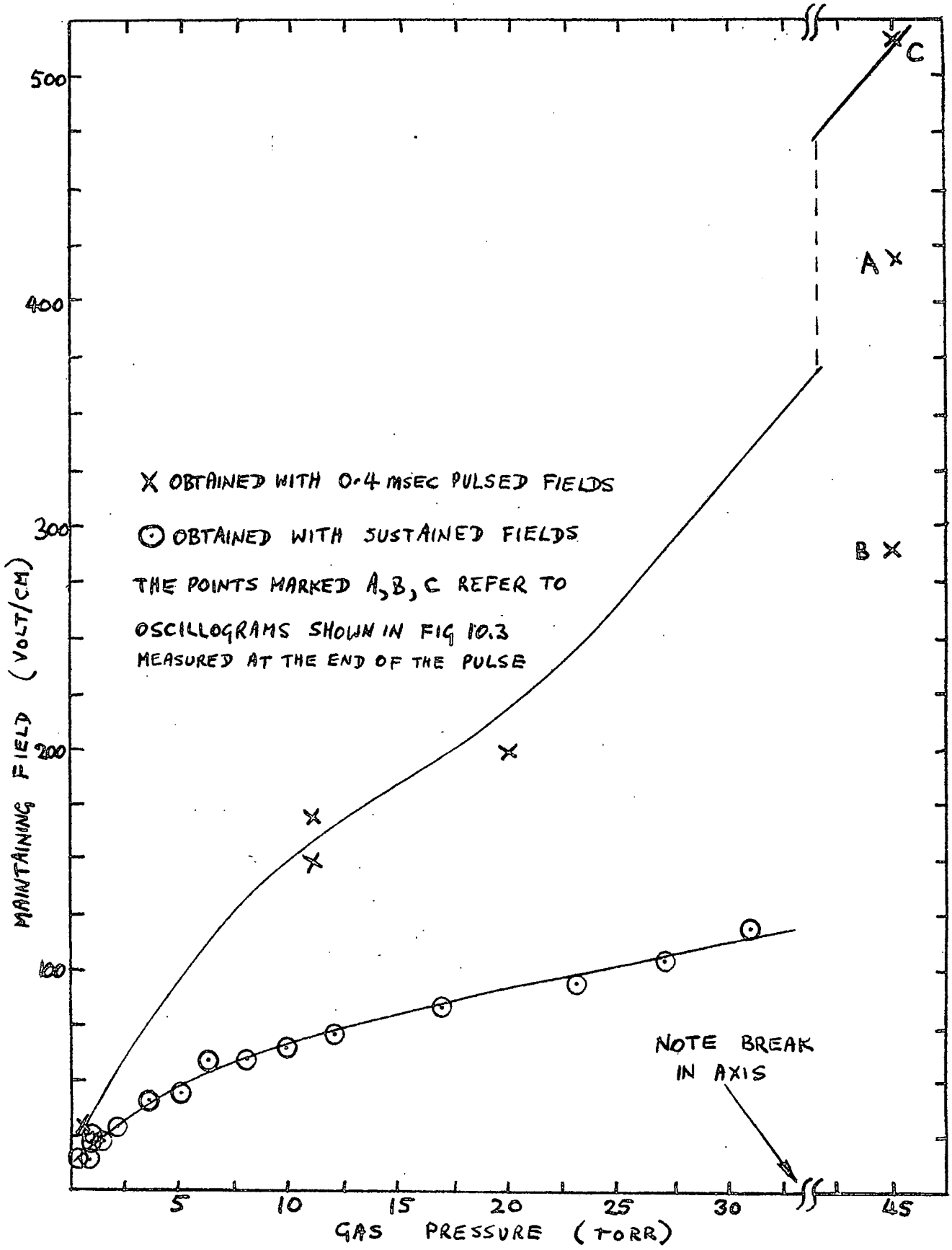


FIG 10.2 THE MAINTAINING FIELDS IN NITROGEN DISCHARGES

second and sometimes reversed its direction of rotation after several seconds. The discharge in neon displayed a somewhat different form of movement: immediately after breakdown the discharge was steady and similar to those of hydrogen and nitrogen, remaining so for about a second. Then the glow would divide longitudinally into streamers of diameter approximately 1mm which stretched between the electrodes. To the eye it appeared that a dozen or more streamers existed at once, all in a state of rapid agitation, but this could have been an illusion caused by persistence of vision: it is possible that a single narrow discharge in rapid random movement could have had the same appearance.

The motion of the discharge in all three gases appeared to be completely uninfluenced by the presence, just outside the spark chamber, of the pole of a powerful magnet.

10.2. The maintaining voltage.

We have seen already, in fig 9.4 an oscillogram which shows the time-resolved voltage at the electrodes as breakdown occurs in the spark gap. It is apparent that a steady maintaining voltage is reached after several microseconds have elapsed from the instant of breakdown. Graphs showing the observed maintaining fields, in hydrogen and nitrogen as a function of pressure, are shown in fig 10.1 and 10.2. No graph for the maintaining voltages in neon can be given because the voltages were too small to measure accurately. This means that up to the highest pressure used in neon,

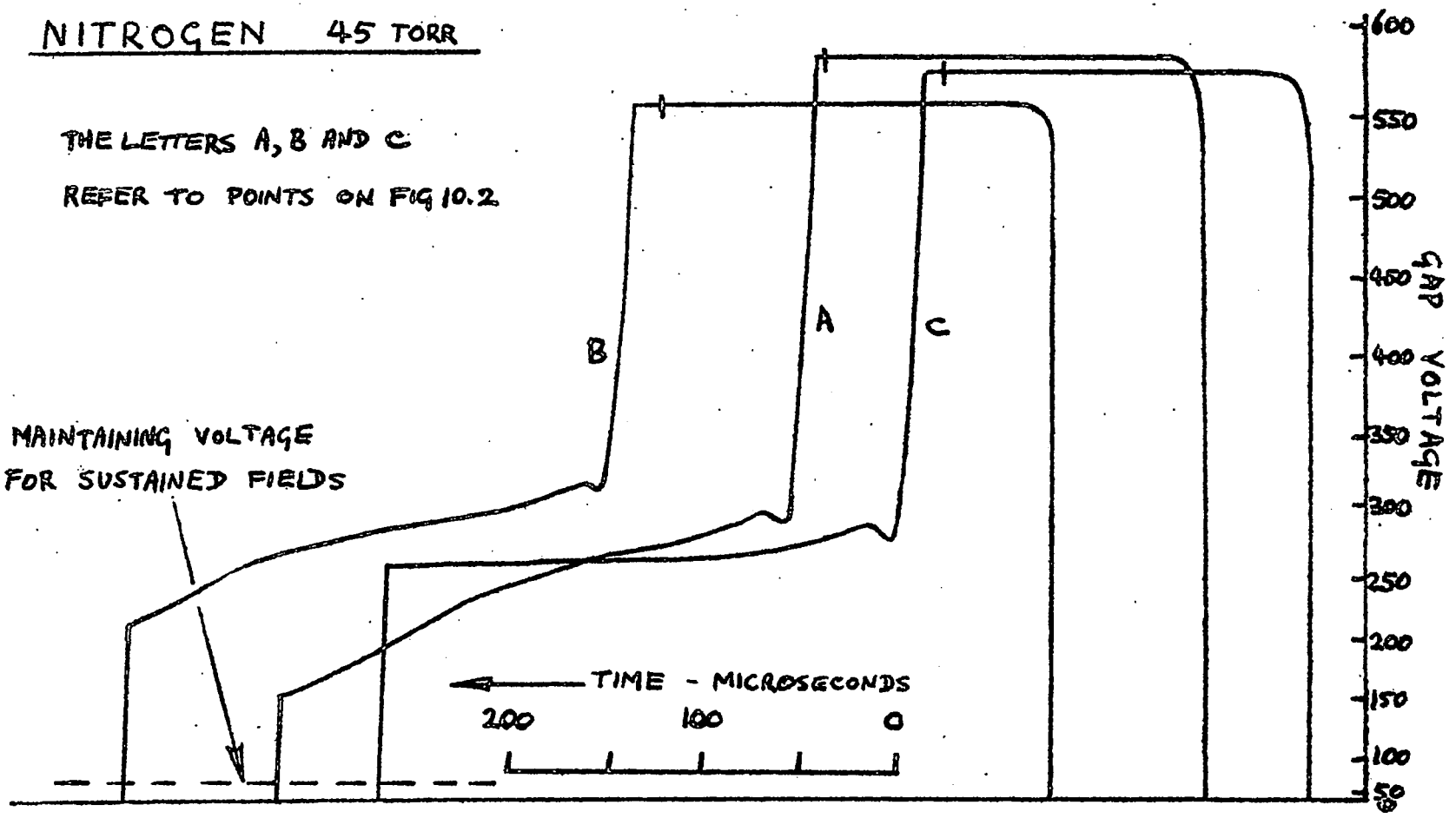


FIG 10.3 TRACINGS OF OSCILLOGRAMS FOR NITROGEN SHOWING THE SLOW DECREASE OF MAINTAINING VOLTAGE WITH TIME

174 torr, the maintaining field was less than about 20 volts/cm

The maintaining fields in hydrogen fig. 10.1 increase almost linearly with pressure over the region studied, although a slight decrease in the slope of the graph is apparent at the highest pressures. The corresponding graph for nitrogen, fig. 10.2, is very similar but in this case the maintaining voltages under sustained fields and pulsed fields were different, as shown by the two graphs. Some of the nitrogen oscillograms such as that shown in fig. 10.3 showed that after a steady maintaining voltage had been reached, a second decrease in voltage occurred towards the end of the pulse. This was a very much less rapid change of voltage than that occuring at breakdown, so that the final equilibrium voltage could not be observed within the 0.4 millisecond pulses used here. However, extrapolation of these oscillograms indicated that this equilibrium voltage could approximate to that measured under sustained fields. In most oscillograms no indication was given that this decrease of maintaining voltage would later occur, and its beginning seemed to be spontaneous and temporally unrelated to previous events in the discharge. A lowering of voltage indicates an increase of discharge current, as does the observation that when this "droop" of maintaining voltage occurred, the discharge could clearly be seen to be much brighter and more constricted than otherwise. The relative

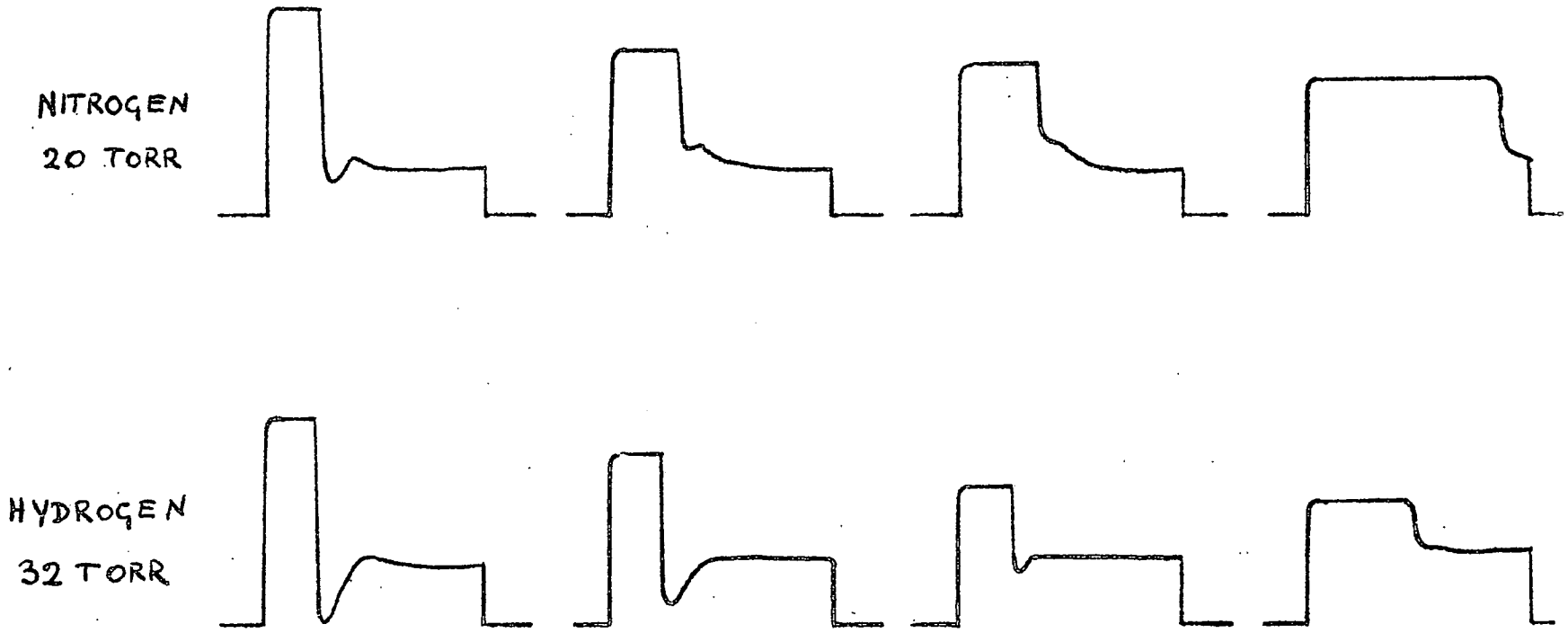
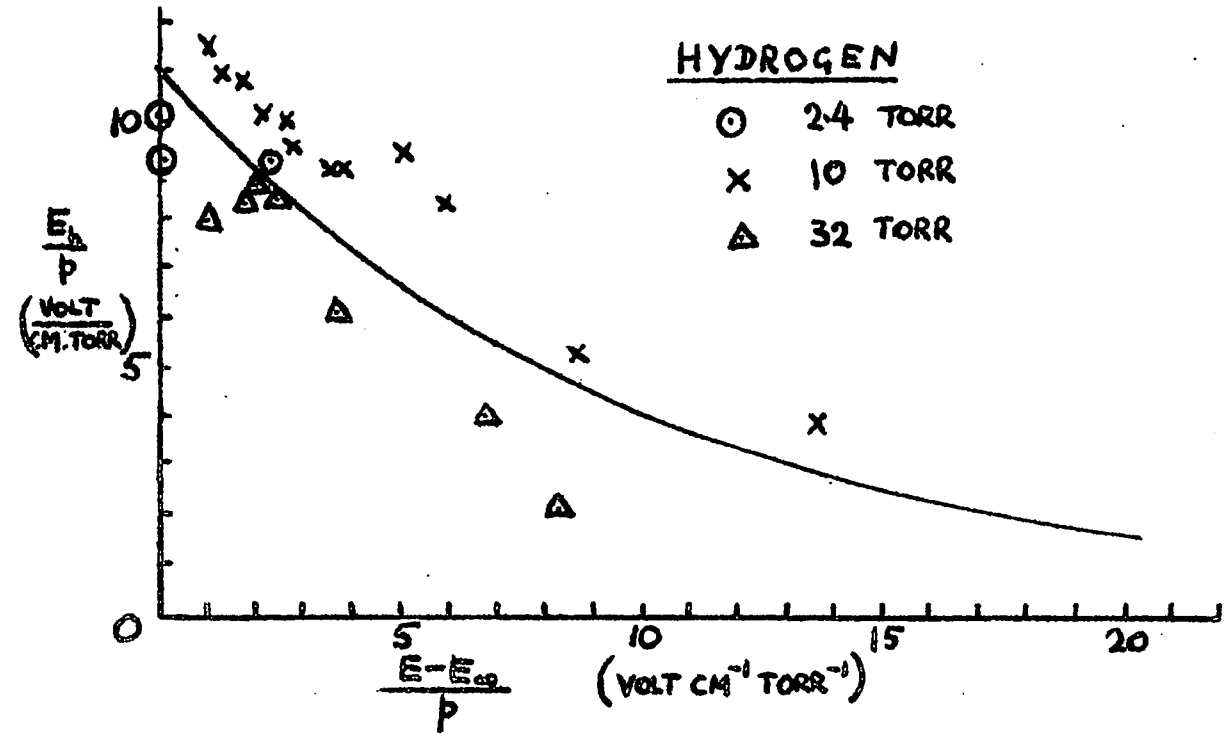


FIG 10.4 OSCILLOGRAMS OBTAINED WITH HYDROGEN AND NITROGEN, SHOWING THE VOLTAGE OVERSHOOT IMMEDIATELY AFTER BREAKDOWN.



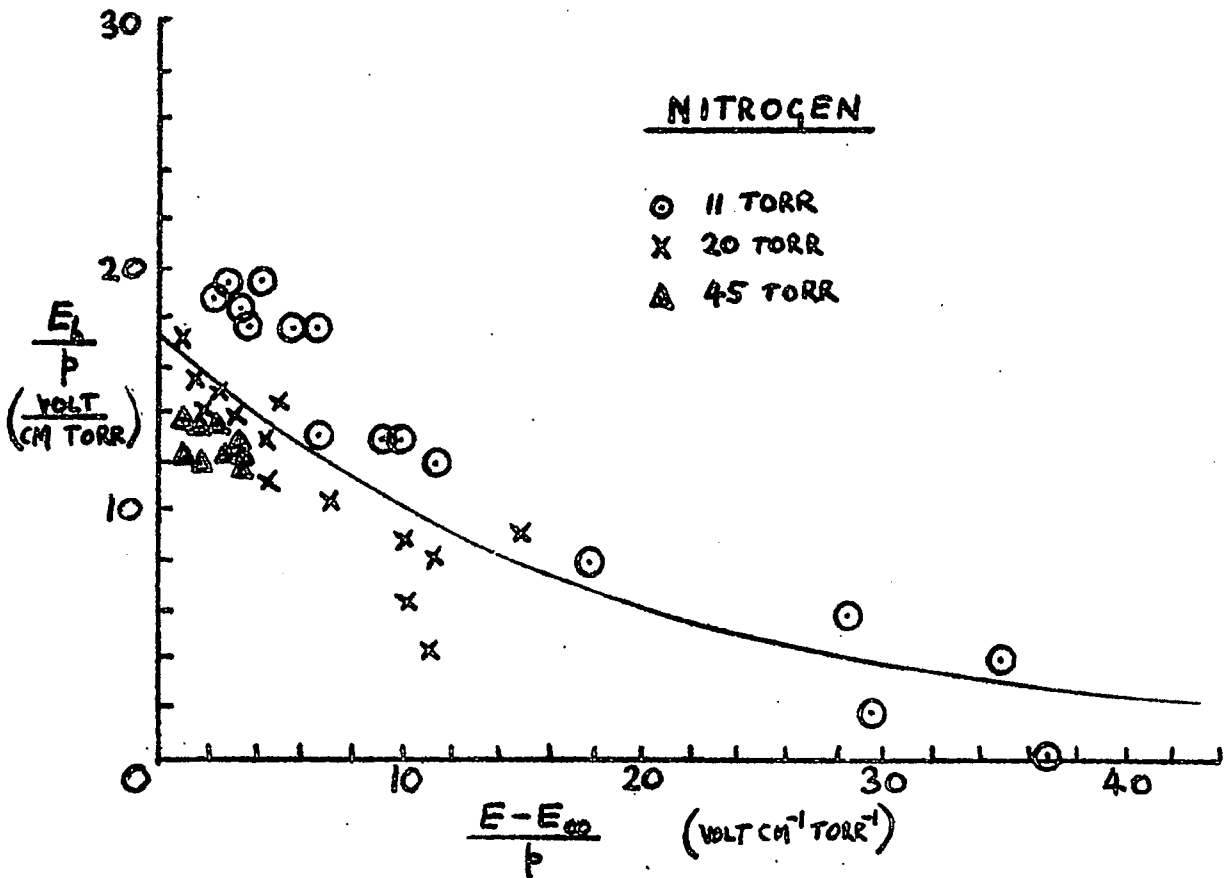


FIG 10.5 THE "OVERSHOOT FIELD" FOR HYDROGEN AND NITROGEN AS A FUNCTION OF THE OVERVOLTAGE

slowness of the change of maintaining field could indicate that it is caused by a modification to the electrode surface by electron or ion bombardment. This could imply that the discharge is at least partially maintained by emission of secondary particles from the electrode.

10.3. The voltage overshoot between breakdown and the steady maintaining voltage.

Oscillograms such as those sketched in fig 10.4 showed that under certain conditions the gap voltage did not tend towards the maintaining voltage immediately after breakdown. Instead the voltage often overshot, nearly reaching zero voltage in some cases, and then rose steadily towards the maintaining voltage. We will refer to the minimum field occurring during this transient state as the "overshoot field", E_h.

It was apparent that the overshoot field decreased with increasing overvoltage, and also that the magnitude of the overshoot became less significant at the higher pressures. In fig. 10.5 the quantity E_h/p is plotted against the overvoltage, for hydrogen and nitrogen. The overvoltage is expressed as $(E-E_\infty)/p$, where E denotes the applied pulse field and E_∞ the field for breakdown under sustained fields. It appears reasonable to fit to the observed points, curves of the form

$$E_h/p = A \exp\left[-(E-E_\infty)/Bp\right]$$

where A and B are empirical constants to be derived from experimental observations.

There was rather too much scatter in the observations to derive great confidence that the above form is correct, but it does comply with the general requirements, firstly that the curves should tend asymptotically towards the overvoltage axis, because the overvoltage may tend to infinity while E_h remains positive, and secondly, that at zero overvoltage the value of A is finite and such that

$$E/p \geqslant A > 0$$

In hydrogen, for the pressures 2.4, 10, and 32.5 torr the value of A was 11 ± 2 volt cm^{-1} $torr^{-1}$

In nitrogen, for the pressures 11, 20, and 45 torr the value of A was 17±4 volt cm^{-1} $torr^{-1}$

The values of B were 10 and 20 volt cm^{-1} $torr^{-1}$ for hydrogen and nitrogen respectively, within about 10%.

After the peak of the overshoot the field gradually approached the maintaining field in a variety of ways, as can be seen from the sketches in fig. 10.4. At high overvoltages and/or low pressures the voltage approached a steady maintaining value, in the manner, very approximately of the function $1-\exp(-t/T_1)$. In contrast, little or no overshoot was observed with low overvoltages and/or high pressures. In this case the rapid fall of voltage,

immediately after breakdown changed abruptly into a relatively slow decay of the approximate form $\exp(-t/T_2)$. At intermediate overvoltages and pressures a small overshoot was observed followed by an overshoot in the opposite direction, which, in turn was followed by an asymptotic approach to the maintaining voltage.

Table III gives the values of the time constants T_1 and T_2 observed in hydrogen and nitrogen as a function of pressure and overvoltage. In view of both the difficulty of estimating the asymptotes and the non-linearity of the oscillograms at low voltages, these time constants may be in error by as much as 30% in some cases, but serve to indicate the trends. Only two values of overvoltage are given, and are denoted by "high" or "low". In this context "high" means between 50% and 100%, while "low" means less than 10%.

Table III clearly shows that T_1 clearly increases with overvoltage in both hydrogen and nitrogen. T_2 on the other hand is not sensitive to overvoltage in hydrogen and tends to decrease as the overvoltage is increased in nitrogen. There seems to be little or no variation of T_1 with pressure. T_2 increases with pressure in hydrogen, but shows no trend in nitrogen. All the times noted in Table III are very long compared to any time-constant associated with the transmission line, the oscilloscope or the coupling between these items,

TABLE III

The overshoot time-constants as a function of overvoltage and gas pressure, in hydrogen and nitrogen.

	pressure (torr)	T_1 (μsec) (voltage) (rising)	T_2 (μsec) (voltage) (falling)	overvoltage
HYDROGEN	2.4	5	2	low
		5	2	high
	10	20	40	low
		40	40	high
	32.5	6	50	low
		50	50	high
	65	-	150	low
		-	100	high
NITROGEN	0.6	10	-	low
		10	-	high
	11	15	120	low
		40	50	high
	20	3	50	low
		10	30	medium
		30	20	high
	45	5	70	low
		10	50	high

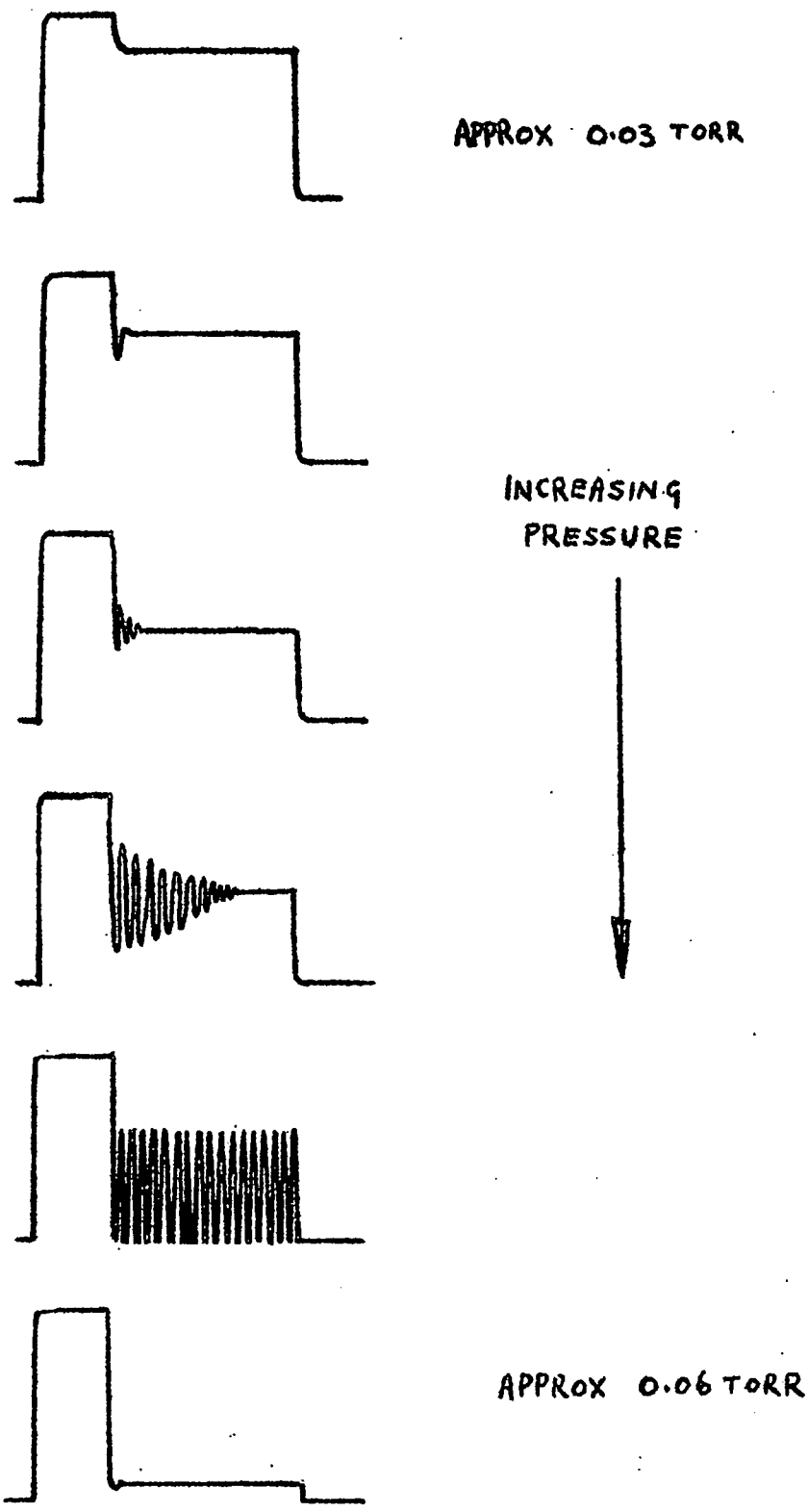


FIG 10.6(a) OSCILLATORY MAINTAINING VOLTAGES IN LOW-PRESSURE HYDROGEN DISCHARGES

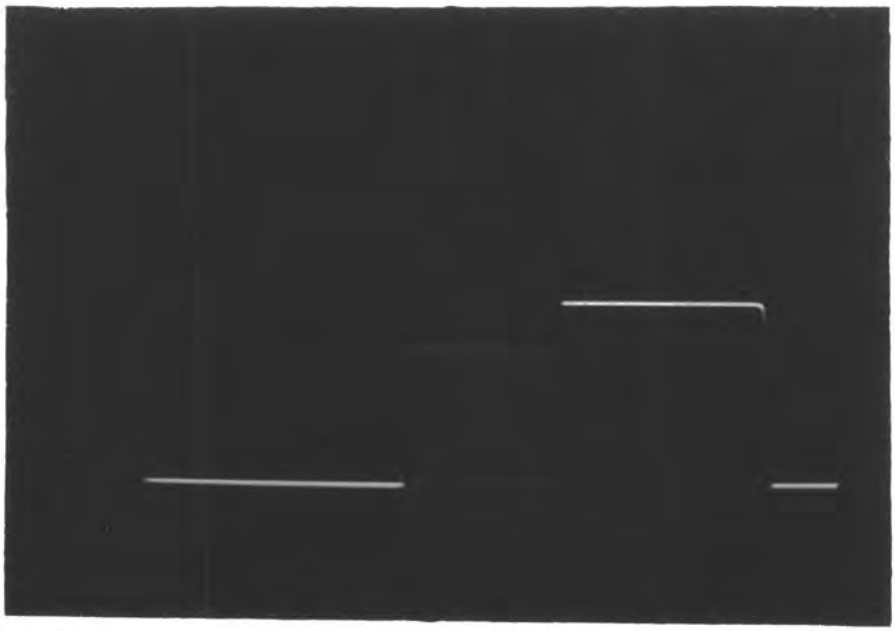
and must therefore be associated with the gas discharge.

10.4. The low pressure discharge

At pressures less than about 0.2 torr, the discharges were no longer confined to the region between the electrodes, but took place in the whole of the electrode capsule except this region. The glow was then very faint and could be seen only in a darkened room. When the pulsed high frequency field was applied to the gap at a pressure of the order 10^{-4} torr a breakdown could still be created, but no light from this could be seen even in a darkened room. The only evidence of a discharge was the discontinuity at breakdown in the voltage oscillogram, (which occurred in this case during the initial rise of voltage,) and a pulsating increase of pressure observable on the vacuum gauge (Penning gauge).

The maintaining voltages at pressures in the region of 10^{-4} torr were but a few percent less than the breakdown voltage, and decreased as the pressure increased towards 0.1 torr, when it was about 10% of the breakdown voltage.

In hydrogen a curious effect was observed at a pressure of 0.06 torr: the oscillograms showing this effect are sketched in the diagram opposite, fig. 10.6(a). As the pressure was allowed to gradually increase towards 0.06 torr the maintaining voltage became oscillatory with a period of several microseconds. At first just one cycle occurred, and with higher pressures took a damped oscillatory form until



(b) Applied pulse showing the lowered breakdown voltage after the initial breakdown.
Total timebase duration: 1 millisecond.

(c) As (b), with expanded timescale, showing form of oscillatory maintaining voltage.
Total timebase duration: 10 microseconds.

Fig. 10.6 (b) and (c) Oscillatory maintaining voltage oscillograms observed with hydrogen at 0.06 torr.

continuous undamped oscillations occurred at 0.06 torr. A slight increase in pressure above this level caused the complete disappearance of the oscillations, these being replaced by a low, steady maintaining voltage. The effect was reversible in that with decreasing pressure all these stages could be observed in reverse order. When the pressure was held constant at 0.06 torr the period was found to vary with the height of the applied voltage pulse, being 4 microseconds at 58 volt and 2 microseconds at 94 volts. The longest period which could be obtained was 10 microseconds, this being observed just above the breakdown threshold. The waveform of the oscillating maintaining voltage which can be seen in fig 10.6(b) was approximately a sawtooth, resembling that obtainable by connecting a capacitor across the terminals of a glow discharge tube. However, there cannot be a very close analogy in the mechanisms as it must be remembered that these oscillograms are of a rectified r.f. waveform. The following further points are of interest, concerning the oscillograms of fig 10.6(a) and (b).

(1) the voltage dropped to zero between successive cycles.

(2) the voltage fell from maximum to zero in less than 0.3 microseconds.

(3) the initial breakdown occurred at random times after the application of the pulse, because the

irradiator had no effect at these low pressures.

(4) when the oscillations were "damped" the final steady maintaining voltage was approximately the mean value of the oscillatory voltage.

(5) the voltage scale was compressed for low voltages in the oscillograms, and the rise of voltage should be recognised as being of the form (1-exp kt).

It is thought that the voltage dropped to such a low level after breakdown that the discharge was extinguished. This being so, the applied voltage would be expected to rise exponentially at a rate appropriate to the Q of the transmission line. The theory of resonant circuits shows that this rate should be such that the voltage reaches 63% the final voltage after Q/π cycles of the applied field. Q was measured by plotting a resonance curve, giving the value 1200, so that at 180 Mc/s, Q/π cycles occupy 2.1 microseconds. In fig 10.6(c) the 63% level, 37 volts, has been marked (taking into account the distortion at low voltages caused by the rectifier) and it can be seen from this figure that the time from zero to this level agrees very well with the above hypothesis. It should also be noted that this also explains the approximate inverse relation between the applied voltage and the period of the relaxation oscillations.

After the application of a pulse the initial breakdown

occurred after several tens of microseconds, whereas the relaxation oscillations occurred with a period of only a few microseconds and with an amplitude less than the applied voltage. It therefore appears that the breakdown threshold increases with time, remaining approximately the same for periods of a few microseconds but changing appreciably between pulses. The time constant of the rise of breakdown threshold would therefore be of the order of milliseconds. The probable explanation of the variation of the breakdown threshold is that electrons are deposited on the glass walls of the discharge chamber during each discharge, and these are neutralised by the arrival of the slower moving positive ions.

Although the intermittent discharge was only observed here with hydrogen it should be pointed out that no deliberate search for this effect was made with nitrogen or with neon.

Finally it should be noted that since the discharges considered in this section did not take place between the electrodes, the voltages quoted must be regarded as arbitrary units, not directly related to the effective electric field strength. Neither would it be proper, to assume that the existing fields would be spatially constant in this case.

10.5. The theory of the glow discharge

As electrons diffuse rapidly out of the discharge

region, they leave behind a positive space charge. This separation of charge creates an electrostatic field which can be so large as to reduce the outward flow of electrons, and accelerate the outward flow of the heavier positive ions, with the result that charges of both sign diffuse at the same rate.

If we express the space charge field as E_s, then the currents of positive and negative particles are given by the sum of diffusion and drift components in the following equations, respectively:

$$\rho_+ v_+ = -D_+ \nabla \rho_+ + \mu_+ E_s \rho_+ \qquad \ldots\ldots 10.1$$

$$\rho_- v_- = -D_- \nabla \rho_- - \mu_- E_s \rho_- \qquad \ldots\ldots 10.2$$

where ρ is the concentration of charge

v is the average particle flow velocity

D is the diffusion coefficient

μ is the mobility

and the suffices + and - refer to positive ions and electrons respectively.

In the steady state we may assume that

$$\left.\begin{aligned} v_+ &= v_- = v \\ \nabla\rho_+ &\approx \nabla\rho_- = \nabla\rho \\ \rho_+ &\approx \rho_- = \rho \end{aligned}\right\} \qquad \ldots\ldots 10.3.$$

Writing these assumptions into equations 10.1 and 10.2 and eliminating E_s we obtain

$$\rho v = -\left[\frac{D_- \mu_+ + D_+ \mu_-}{\mu_+ + \mu_-}\right] \nabla \rho \qquad \ldots\ldots 10.4.$$

Comparing this equation with the corresponding equation for free diffusion :

$$\rho v = - D \nabla \rho$$

we can see that the expression in parentheses in equation 10.4 represents an effective diffusion coefficient, which is called the ambi-polar diffusion coefficient, D_a.[26,29,43]

We have therefore $D_a \equiv \dfrac{D_- \mu_+ + D_+ \mu_-}{\mu_+ + \mu_-}$ ..... 10.5.

and since $\mu_- \gg \mu_+$

$$D_a \approx \mu_+ D_- / \mu_- \qquad \ldots\ldots 10.5a.$$

The operation of this ambi-polar coefficient in a discharge requires a certain minimum space charge. If the space charge is less than this minimum an intermediate coefficient must be used to represent the rate of loss of electrons. The value of this coefficient, denoted by D_s, is clearly such that

$$D_- > D_s > D_a$$

Allis and Rose[44] have derived the following expression for D_s in terms of D_a, D_-, the characteristic diffusion length L, and the gas conductivity σ:

$$D_s = D_a \left[\frac{D_- + L^2\sigma/\varepsilon_o}{D_a + L^2\sigma/\varepsilon_o} \right] \qquad \ldots\ldots 10.6.$$

where ε_o is the permittivity of free space.

Since $D_- \gg D_a$ the condition that $D_s \approx D_a$ is therefore:

$$D_- \ll L^2\sigma/\varepsilon_o \qquad \ldots\ldots 10.7.$$

It is now recalled that the voltage across the spark gap can be expressed in terms of the gas conductance, Y, by use of equation 8.1a. The "steady state" solution of that equation reads

$$V = (1+Z_o Y/\zeta)^{-1} \qquad \ldots\ldots 10.8$$

where V is the gap voltage relative to the breakdown voltage,

Z_o is the characteristic impedance of the transmission line and

ζ is a dimensionless quantity related to the quality factor of the line.

Since $Y = \pi R^2\sigma/d$ ..... 10.9.

where R, d are the gap radius and width respectively, we have, by substituting equations 10.8 and 10.9 into the inequality 10.7, the condition for ambi-polar diffusion

$$D_- \ll \frac{L^2\zeta d}{\pi\varepsilon_o Z_o R^2} \left[\frac{1-V}{V} \right] \qquad \ldots\ldots 10.7(a).$$

For the maintaining voltages in hydrogen and nitrogen, the value of V was approximately 0.5.

Putting in the other values:

$L = 1.25\ 10^{-3}$ metre

$\zeta = 0.01$

$d = 4.94 \times 10^{-3}$ metre

$\epsilon_0 = 8.85 \times 10^{-12}$ farad/metre

$Z_0 = 100$ ohm

$R = 5 \times 10^{-3}$ metre

the inequality 10.7(a) becomes,

$$D_- \ll 1.1 \times 10^7 \text{ cm}^2/\text{sec} \qquad \ldots\ldots 10.7(b)$$

Using the expression

$$D/\mu = (2/3) \times (\text{average electron energy})$$

and data from figs 8.4 and 8.5 we find

$$p D_- \approx 7 \times 10^5 \text{ cm}^2\text{torr/sec}$$

for both hydrogen and nitrogen, which satisfies inequality 10.7(b) for all pressures greater than about one torr.

For neon, $pD_- \lesssim 1 \times 10^7$ cm^{2}torr/sec.

In this case, however, the maintaining voltages were too low to measure; if we put V=0.1 as an upper limit, for the sake of an example, we have from equation 10.7(a) the inequality

$$D_- \ll 10^8 \text{ cm}^2/\text{sec} \qquad \ldots\ldots 10.7(c)$$

which therefore is also true for pressures greater than one torr.

It may be concluded therefore that the maintained discharges observed were controlled by ambi-polar diffusion.

An objection to the use of the electrode radius for R in equation 10.9 is valid. Strictly, R should be the radius of the discharge column. However, for pressures greater than a few torr, this was less than the electrode radius, and therefore use of the discharge radius would enhance the inequality 10.7(a) and the conclusion of ambi-polar diffusion.

Schneider[45] has modified equation 10.2 to include the electron-ambit losses for the case of a high frequency discharge. He writes

$$\rho_- v_- = -D\nabla\rho_- - \mu_- E_s \rho_- + \alpha x \rho_- \qquad \ldots\ldots 10.2(a)$$

where αx represents the probability that an electron will move out of the discharge under the influence of the applied field. x is the co-ordinate parallel to the applied field measured from the mid-gap plane.

Making the assumptions of equation 10.3, Schneider eliminates E_s between equations 10.1 and 10.2(a) and substitutes for ρv in the continuity equation

$$\text{div}\ \rho_- v_- = \nu \rho_- \qquad \ldots\ldots 10.10$$

where ν is the rate of generation of electrons, per electron.

The resulting second order differential equation can be solved by separation of the variables, to give a solution of the form

$$\rho = \rho_0\ A(x)\ B(r)$$

The function B is a Bessel function of the first kind and zero order, while A is found by substitution of a power series for x.

Thus $B(r) = J_o(2.405r/R)$

and $$A(x) = 1 + \sum_{n=1}^{n=\infty} \frac{[(2n-1)a-k][(2n-3)a-k] \ldots [a-k]x^{2n}}{2n}$$

..... 10.11.

where $k = \nu/D_a - (2.405/R)^2$

and $a = \mu_+\alpha/(\mu_+ + \mu_-)D_a$

For zero applied field $\alpha = 0$ and equation 10.11 reduces to

$$A(x) = \cos x\sqrt{k}$$

Finite values of the applied field cause some electrons to be lost to the electrodes by drift, so that α increases with the field, approaching the value k in the limit.

We may make a rough estimate of the relative rates of loss of electrons by diffusion and drift to the electrodes by comparing the distance an electron diffuses during a half period of the applied field, (in one dimension), with the distance moved due to the applied electric field, (the ambit). An expression for the ambit is derived in Appendix 3, while Einstein[46] has derived a formula for the r.m.s distance moved by a diffusing particle in time τ:

$$\sqrt{\overline{x^2}} = \sqrt{2D\tau}$$

In this case D is the ambi-polar coefficient D_a,

while τ is one half of the period of the applied field.

The classical expression for electron drift velocity is

$$v_{drift} = 2eE/3m\nu_c$$

Substituting in equation A3.1 for ν_c (since $\nu_c = 1/t_c$) and writing $\mu_- = v_{drift}/E$

and $E_o = \sqrt{2}E$

we have:

electron ambit $= 3\, v_{drift}/\sqrt{2}w$

therefore

$$\frac{\text{ambit}}{\text{diffusion dist}} = 3\, v_{drift}/w\sqrt{2} \div \sqrt{\pi D_a/w}$$

$$= 1.2\, v_{drift}/\sqrt{wD_a}$$

This may be written, since $w = 1.1\ 10^9\ sec^{-1}$

$$\frac{\text{ambit}}{\text{diffusion dist}} = 3.4 \times 10^{-5}(E/\sqrt{p})(\mu_- p)(D_- p\mu_+ p/\mu_- p)^{-\frac{1}{2}} \quad \ldots\ldots 10.12$$

into which the approximate values from Table IV may be inserted for the range $2 < p < 30$ torr.

Substitution of the values from Table IV into equation 10.12 shows that for hydrogen and nitrogen the ambits are respectively 5 and 4 times larger than the diffusion distance.

TABLE IV

Experimental values of mobility, diffusion coefficient and maintaining voltage.

	Hydrogen	Nitrogen	units	ref.
$\mu_- p$	$3.7\ 10^5$	$3.3\ 10^5$	torr $cm^2 volt^{-1} sec^{-1}$	34, 35
$\mu_+ p$	10000	2000	torr $cm^2 volt^{-1} sec^{-1}$	29
$D_- p$	$4\ 10^5$	$5.5\ 10^5$	torr $cm^2 sec^{-1}$	26, pp47 & 50
$E/\sqrt{p}$	30 - 50	18 - 20	volt $cm^{-1} torr^{-\frac{1}{2}}$	present paper

The value of E used in Table IV was the measured maintaining field. This is not strictly the correct value but is probably not in great error. A proper value would be that between the electrodes and the space charge potential.

It would appear, then, that the maintained discharge is drift controlled, and that we have a close approximation to the limiting condition mentioned above, in which a approaches the value k. Schneider has solved the transcendental equation 10.11 for the particular case a=0.99k. Applying the boundary condition A =0at $x = d/2$, he obtains the value $k = 39/d^2$. Under these conditions A(x) is approximately constant across the gap, and starts to fall sharply towards zero near the electrodes.

In the electron-depleted regions near the electrodes it

would be expected that there would be few excited molecules, as these would decay before diffusing far into this region. The region would therefore appear dark, in agreement with the experimental observation reported in section 10.1.

A possible explanation of the slow recovery of the gap voltage towards the maintaining voltage after the overshoot may be made as follows: considering first the gas immediately before breakdown, we see that the number of electrons in the gap is greater than that required for equilibrium under the existing value of field strength. This is because of the energy stored on the transmission line. Further, because an electron can retain a high energy even after many collisions with slow gas molecules, the initial rate of ionization can be maintained for a short time after the gap voltage has collapsed. Therefore there is suddenly an electron density in the gap which may be far greater than that required for equilibrium, with the reduced loss rate appropriate to ambi-polar diffusion. Therefore the voltage drops to a low value, and the electrons "cool down" and are lost from the discharge by ambi-polar diffusion. Then the voltage rises and the ionization rate increases till it balances the diffusion losses. While the applied field is very low the ionization rate is very low, so that the continuity equation reads

$$\frac{\partial \rho}{\partial t} = D_a \nabla^2 \rho$$

By analogy with the corresponding equation for free diffusion as given in Chapter 8, we can see that the time dependent solution to this is

$$\rho = \rho_o \exp(-D_a t/L^2)$$

where L is the characteristic diffusion length.

The time constant of this solution should represent the order of magnitude of the recovery time, τ. That is,

$$\tau \approx L^2/D_a$$

Using equation 10.5(a) and putting in values from Table IV and writing L = 0.125cm

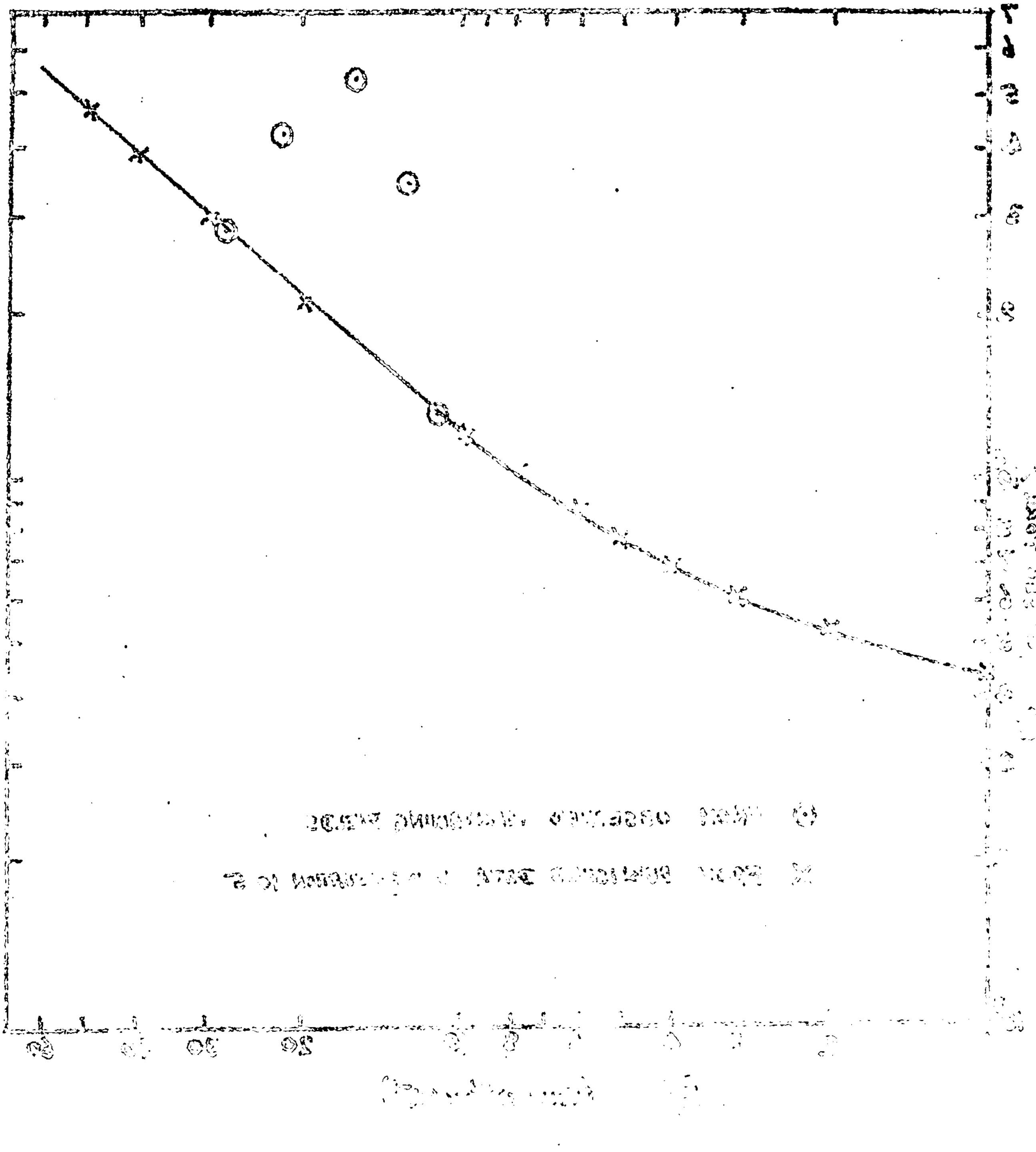
we have $\tau \approx 1.5p$ microseconds for hydrogen

and $\tau \approx 4.7p$ microseconds for nitrogen

It is interesting to compare these times with the experimental time constants T_1 given in Table III, section 10.3. It can be seen that agreement is obtained in the high overvoltage case for which the above theory is applicable, if reasonable allowance is made for the over-simplification of the theory and the uncertainties of the experimental observations.

10.6. Experimental determination of the ambi-polar diffusion coefficient.

We have seen in the previous section that the ambi-



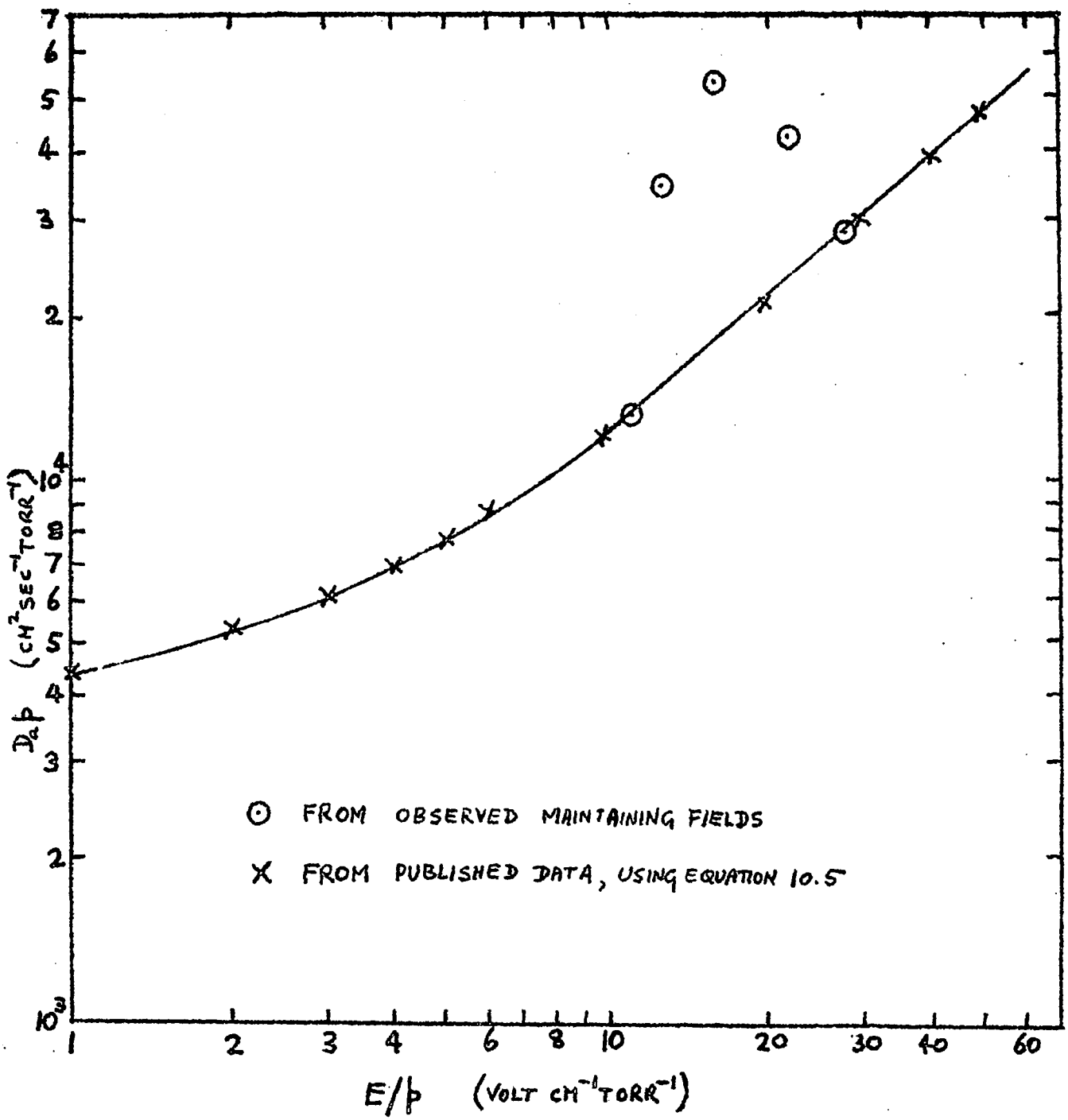


FIG 10.7 THE AMBI-POLAR DIFFUSION COEFFICIENT IN HYDROGEN

polar diffusion coefficient may be expressed in terms of the mobilities and free diffusion coefficients of particles of both signs. This coefficient is plotted in fig 10.7 as a function of E/p for hydrogen. The variation with E/p arises because although the quantities $\mu_+ p$, $\mu_- p$ and $D_+ p$ are almost independent of the field, $D_- p$ is considerably influenced by the field. D_- was obtained by using the expression

$$D/\mu = (2/3)(\text{average electron energy})$$

An empirical formula derived from fig 8.4 was used to express the average electron energy, $\bar{u}$ in volts:

$$\bar{u} = 0.115(E/p) + 0.45$$

and thus we obtain the relation

$$D_a p = 860(4.1 + E/p)$$

An alternative approach uses the observed values of maintaining field to obtain D_a. We know that in the equilibrium condition electrons are generated in the discharge and are lost by ambi-polar diffusion. (Even if electrons hit the electrodes under the influence of their drift velocity, they must approach the electron-depleted region adjacent to the electrodes by ambi-polar diffusive motion). We may therefore equate production and loss terms:

$$\nu_i = D_a/L^2$$

and hence $$D_a p = (pL)^2 (\alpha/p) v_{drift}$$

Now we have seen in equation 8.11 that α/p can be

expressed in terms of the breakdown field E_s. Therefore by substituting for α/p and putting $E_s/p_s = E/p$ we have

$$D_a p = (2/3)\ \bar{u} p E v_{drift}/E_s^2$$

$$= K(EL/E_s)^2$$

where K is defined in equation 8.14 and evaluated in fig 8.6.

E_s is known in terms of E/p from the breakdown fields given in fig 9.1 while E is the maintaining field. The values of E/p for the maintained discharges in hydrogen were high enough to correspond to experimental values of E_s/p_s for the range $1 \leqslant p_s \leqslant 20$ torr. Some representative values of $D_a p$ in this pressure range were calculated by this method and are plotted (as circles) on fig 10.7. It can be seen that agreement exists (as to order of magnitude, at least) between these and the values calculated by the previous method. However, it should be remembered that the space charge considerably modifies the field in the discharge, and therefore the proper interpretation of E may be not the maintaining field but some other effective field.

CHAPTER XI

CONCLUSIONS

A description has been given in the preceding chapters of modifications made to an apparatus which was designed originally by Prowse, Rowbotham and Monk for measurements of breakdown fields and formative delays in gases. An extension of the original series of measurements has also been described which was carried out under the following general conditions: gas pressures ranged from 2.4 to 65.4 torr in hydrogen, 0.59 to 45 torr in nitrogen, and 4.7 to 174 torr in neon. The frequency of the applied field was 180Mc/s and this, together with the width of the spark gap (0.494cm), and the gas pressures used, ensured that

(a) the mean free path and electron ambit were less than the gap width and

(b) the electron-molecule collision frequency was much greater than the radian frequency of the applied field.

The range of breakdown delays which could be measured extended from 0.2 microseconds up to about 400 microseconds, and the maximum r.f. field available was 1150 volts/cm. The modified apparatus thus proved to be capable of a greater range than the original, both in available fields

and in observable delays. Further, the stability of the applied field, and the quality of oscillograms were improved, resulting in considerably greater accuracy of the measurements of the field.

The ability to detect delays shorter than the 10 μsec which was the lower limit achieved by Prowse, Rowbotham and Monk, allowed the variation of formative delays with the applied voltage to be measured in both hydrogen and nitrogen. The resulting agreement with a theory based on an exponential increase in the electron population in the spark gap with time adds weight to the hypothesis that similar processes lead to breakdown under both threshold and overvolted conditions. The theory used the observed threshold fields for breakdown to calculate the net ionization rate as a function of overvoltage. This technique effectively normalised the graphs of breakdown delays as a function of the applied field to that value yielding an "infinite" delay.

An exception to the agreement between observation and theory of formative delays was found in neon where the

theoretical predictions were too small by a factor of about 10. This was tentatively explained in terms of a trace impurity of argon, which was considered to lower the threshold breakdown fields of neon by means of the Penning effect. In this effect an argon atom is ionized with a high probability, by collision with a metastable neon atom. As neon metastables are produced by electrons of energy lower than is required for direct ionization a condition thus soon arises in which all the argon but relatively little neon is ionized, and the process is retarded. An equilibrium state is then set up where the argon ions are neutralised by collision with the electrodes or are replaced by neutral atoms diffusing into the discharge from the surrounding gas. This state may, or may not, satisfy the electron density criterion for breakdown, according to the impurity concentration. At fields greater than threshold, ionization processes inherent to neon become dominant with the result that delays predicted from threshold breakdown fields are considerably shorter than those observed in practice. This hypothesis is open to the criticism that no proof of the existence of any argon impurity has been obtained. The inability of the apparatus to analyse the test gas must be admitted as a serious deficiency, which would have to be rectified if experiments of this type are to be continued. In view of the small amounts of impurity, (argon in particular), which

can appreciably reduce the breakdown voltage of neon it would seem to be dangerous to rely on assumptions of gaseous purity, merely on the basis of utilization of "clean" vacuum techniques.

The theoretical approaches made here are essentially of a simple kind, and indeed it is possible that the disagreement between theory and results for neon is due to an oversimplification in this case.

The neglect of the variation of ionization rate as the gap voltage falls, at breakdown, probably has little effect on the theoretical prediction, because of the high level (99% of the applied voltage) at which the delays were measured. The effect of space charge starts to become noticeable at about 10^{-7} amps in d.c. discharges. This current level in the hydrogen and nitrogen discharges is about that which gives a 1% drop in the applied voltage and it would therefore not be expected that the delay would be modified. In neon this current level is produced by relatively fewer electrons, and hence earlier, so that diffusion losses would be less. It might be expected therefore that the delays in neon would be shorter than predicted by theory if neglecting this effect had been important.

In general, a rigorous theory should use basic data such as the efficiencies and potentials for ionization and

excitation to calculate the distribution function of electron energies. References 5, 6 and 7 have shown how to calculate the breakdown fields under sustained fields, (i.e in a system independent of time) by this means and it should be, in principle, possible to obtain the time dependent solution for the case of an overvolted gap by a similar technique. However, the rigorous solution promises to be exceedingly difficult, and the technique used here for prediction of delays should, to a large extent, nullify errors arising from, for example, the use of average values of the drift velocity and electron energy, and the assumption of a possibly incorrect electron energy distribution (the Maxwellian distribution has been assumed).

The existence of maintaining voltages has been observed when the discharge was well established in hydrogen and nitrogen, and these have been measured as a function of pressure. For hydrogen the maintaining field was 0.67 of the breakdown field, over the pressure range 2.4 to 65 torr while for nitrogen the corresponding fraction was 0.17 over the pressure range 0.6 to 45 torr.

Calculations of the electron density in the discharge, from the maintaining voltage showed that electron diffusion losses were ambi-polar, or space-charge controlled, in the maintained discharge. In this condition the magnitude of the electron ambit relative to the distance an electron can

diffuse during a half cycle of the applied field was so high that the majority of electrons were driven into the electrodes by drift.

Values of E/p in the maintained discharge in hydrogen were sufficiently high to allow the ionization rate to be calculated from the breakdown field. Equating this to the rate of loss of electrons due to diffusion gave a diffusion coefficient which was equal to the ambi-polar coefficient – to within a factor 3 – as calculated from published data on electron mobilities and energies. This shows that the effective field in the discharge is approximately the same as that observed at the electrodes.

In neon, no maintaining voltage could be observed at pressures up to 174 torr, which puts an upper limit on the maintaining field of 20 r.m.s. volt/cm. These low maintaining fields in neon are a result of the relatively long mean-free-paths and hence high mean electron energy, compared with the other two gases used, at a given value of E/p.

Between breakdown and the establishment of a steady maintaining voltage a transient state was observed during which the gap voltage approached zero, to a degree depending on the overvoltage. The subsequent recovery to the steady maintaining voltage was approximately exponential in form, with a time constant which agreed tolerably well with that which might be expected if ambi-polar diffusion was controlling the loss of an excess electron population. The field

would be too low during the initial part of the recovery for there to be appreciable ionization at this stage, but as electrons were lost from the discharge, and the field built up, the ionization rate also built up and an equilibrium state resulted.

In view of the small number of gases investigated here it is clear that the scope of the experiments described needs to be extended before the conclusions can be verified. There are two principal ways in which the experimentation could be improved: firstly some advantage would accrue from knowing accurately the number of electrons produced in the gap by the irradiator, and secondly it would be better to measure the buildup of discharge current with time rather than the voltage. Measurement of the discharge current would appear to be difficult at the frequency used here however, both because of the relatively large capacitative currents flowing in the gap, and also because of stray fields which are picked up by the detector.

ACKNOWLEDGMENTS

I would like to express my sincerest thanks to my supervisor Dr. W. A. Prowse for his advice and guidance during the course of this work, and also to Professor G. D. Rochester for the laboratory facilities used. I wish to thank the Department of Scientific and Industrial Research for a maintenance grant during the years 1960 to 1963 and the British Electrical and Allied Industries Research Association for provision of a grant for apparatus during these years.

I would like to thank Mr. R. G. Earl with whom I worked in close co-operation on these experiments, and also my colleagues Dr. R. E. Long and Mr. J. Vincent for most helpful discussions.

I wish to acknowledge the help I have received from my colleague Miss M.G.E. Welton and from Mrs. E. Templeton of the Durham University Computing Laboratory, who corrected and operated my computer programmes, and from Mr. R. White who reproduced the photographs in this Thesis, and finally to my wife who has typed this work and whose patience has been of the utmost moral support.

APPENDIX I

Computer programme to evaluate equation 8.1a.

We require to sum the following series, which apart from a constant factor is equation 8.1a:

$$V = (1+g_o)(1+f^2g_1+f^4g_1g_2+f^6g_1g_2g_3+ \ldots\ldots) \qquad \ldots\ldots \text{A.1.1}$$

where $g_n = (1-\delta_n)/(1+\delta_n)$ for n=1,2,3, ... A.1.2

δ_n in the notation of Chapter 8 is given by

$$\delta_n = Z_oY_o \exp \varphi T(N-n)$$

This can be split into two factors to separate the constant and variable parts, and for 100 electrons initially in the gap we may calculate $Z_oY_o = \exp(-22)$, thus

$$\delta_n = (\exp(\varphi TN-22))(\exp(-\varphi Tn)) \qquad \ldots\ldots \text{A.1.3}$$

A calculation using the "steady state" solution (g=constant) indicates that by putting $N\varphi \approx 500$ a "droop" of about 1% will result. We here put $N\varphi = 416$ to allow a safe margin. This figure is chosen such that for φ=1,2,4,8,16, and 32, N is an integer, for convenience.

T is the delay of the transmission line for a return journey: the line being two wavelengths long at 180Mc/s therefore gives T=0.022 microseconds. (The calculations will be carried out in microseconds, with φ in $(\text{microseconds})^{-1}$.

It will be evident now that δ is very small and thus g is very nearly equal to unity. Also f=0.995, and therefore the series, equation A1.1, converges only very slowly, and

several hundred terms were necessary to obtain the final value of V. The computer programme is shown overleaf, in programme 1. This is written in the Algol notation as used in the Elliott 803 computer with 8-hole punched tape input and output.

Referring now to programme 1, lines 1 to 3 define the words used, in terms recognisable by the machine. The "real" numbers and "integers" are the working symbols, and "switch" is a label for an instruction which allows jumps backward and forward to be made in the programme.

We define the following quantities in machine symbols:

$$nb \equiv \varphi$$

$$vg \equiv N$$

$$vm \equiv \varphi T$$

$$vn \equiv \exp(\varphi TN-22) \equiv \delta_o$$

$$nl \equiv n$$

$$vr \equiv g_o$$

$$vs \equiv f^2 \frac{1-\delta_o \exp(-\varphi Tn)}{1+\delta_o \exp(-\varphi Tn)} \equiv f^2 \frac{1-\delta_n}{1+\delta_n} \equiv f^2 g_n$$

Referring now to programme 1 the following notes will clarify the procedure:

line 4: read in the initial value of φ from punched tape.

lines 5 and 6: calculate initial value of N and φT

line 7: print out the current value of φ followed by "psi"

lines 9 and 10: set the variables back to "one" after a cycle

```
1          begin real vm,vn,vp,vq,vr,vd,vs;
2          integer vg,nb,nl;
3          switch ss:= sa,sb,sc,sd,se,sf,input;
4  input:  read nb;
5  sc:     vg:= 416/nb;
6          vm:= nb*0.022;
7          print digits(3),nb,£ psi?;
8  sd:     vn:= exp(vg*vm-22);
9          vp:= vq:=1.0;
10         nl:= 1;
11         vr:= (1-vn)/(1+vn);
12         if  vr< 0.1 then goto se;
13 sa:     vs:= 0.99*(1-vn*exp(-vm*nl))/(1+vn*exp(-vm*nl));
14         vp:= vp*vs;
15         vq:= vq+vp;
16         if vp ≤ 0.01 then goto sb;
17         nl:= nl+1;
18         if nl ≠ 250 then goto sa;
19 sb:     vd:= vq*(1+vr);
20         print digits (3),nl, prefix (£ ?), digits(3),vg,
                aligned(3,2), vq, aligned(3,2),vd;
21         vg:= vg+10;
22         if vd > 50 then goto sd;
23 se:     nb:= nb*2;
24         print ££L ??;
25         goto if nb=64 then sf else sc;
26 sf:     end of programme;
```

PROGRAMME I

of the programme loop.

line 11: calculates current value of g_o

line 12: if g_o is less than 0.1 go to line 23 to calculate next value of φ.

line 13: calculates the term $f^2 g_n$, initially with n=1

line 14: the next term in the series is calculated by multiplying the preceding term by $f^2 g_n$.

line 15: the series (so far, of n terms) is summed by adding the term calculated in line 14 to the existing sum formed on previous cycles of this loop.

line 16: the series is terminated if the term to be added is less than 0.1

line 17: n is increased ready for recycling the loop sa

line 18: loop sa is traversed again up to a maximum of 250 times, thus adding one term on each cycle.

line 19: the sum from line 18 is multiplied by $(1+g_o)$ to obtain the final answer for that particular value of N.

line 20: n, N, the sum vq, and product vd are printed out, in line.

lines 21 and 22: N is increased by 10 and loop sd traversed until vd becomes less than 50.

line 23 and 25: φ is increased by factor 2. When this reaches 64 the programme terminates.

line 24: A vertical feed is provided on the output to separate results for different values of φ.

APPENDIX II

Recurrence formula and modified computer programme.

The programme which has been given in Appendix I is inefficient with computer time because the series of equation 8.1(a) has to be calculated for each value of N. In this Appendix it will be shown how the solution for a given value of N can be calculated from the solution for the previous value of N. Thus the sum has to be completely carried out only for the initial value of N.

The series to be summed is, again:

$$V = (1+g_0)(1+f^2g_1+f^4g_1g_2+ \ldots)$$

Now we wish to include N in the notation, so remembering that

$$g_n=(1-\delta_n)/(1+\delta_n)$$

and $$\delta_n=\mathbb{Z}_0Y_0\exp\varphi T(N-n)$$

we may define $y_{N-n} \equiv f^2g_n$

and $$y_{N-p}^{N-q} \equiv y_{N-p}\cdot y_{N-p+1}\cdot y_{N-p+2} \cdots y_{N-q}$$

Further, we denote by V_∞^N the sum of an infinite number of terms with N as the parameter. Thus:

$$V_\infty^N=(1+y_N/f^2)(1+y_{N-1}^{N-1}+ y_{N-1}^{N-2}+\ldots.+y_{N-1}^{N-9}+ y_{N-1}^{N-10}+ y_{N-1}^{N-11}+ \ldots)$$

In a similar fashion, for the next value of V we may replace N with N+10:

$$V_\infty^{N+10}= (1+y_{N+10}/f^2)(1+y_{N+9}^{N+9}+\ldots y_{N+9}^{N+1}+ y_{N+9}^{N+0}+ y_{N+9}^{N-1}+ \ldots)$$

Therefore:

$$V_\infty^{N+10} = (1+y_{N+10}/f^2)(1+..+y_{N+9}^{N+1}+y_{N+9}^{N+0}+y_{N+9}^{N+0}y_{N-1}^{N-1}+y_{N+9}^{N+0}y_{N-1}^{N-2}+...)$$

$$= (1+y_{N+10}/f^2)(1+...y_{N+9}^{N+0}(1+y_{N-1}^{N-1}+y_{N-1}^{N-2}+...))$$

$$= (1+y_{N+10}/f^2)(1+y_{N+9}^{N+9}+..+y_{N+9}^{N+1}+y_{N+9}^{N+0}V_\infty^N/(1+y_N/f^2)) \quad ...A2.1$$

Similarly we have

$$V_\infty^{N+5} = (1+y_{N+5}/f^2)(1+y_{N+4}^{N+4}+..+y_{N+4}^{N+1}+y_{N+4}^{N+0}V_\infty^N/(1+y_N/f^2)) \quad ...A2.2$$

Referring to programme 2 we see that with only minor modifications, lines 4 to 19 are the same as the beginning of programme 1.

In line 4 a new symbol, nd, appears. This is the exponent of Z_oY_o which is typically about-20, and is initially read in from the input data tape.

φ, represented by nb, this time has its initial value, unity, written into the programme, in line 5.

Line 6 calculates a suitable initial value for N.

The series is summed to 500 terms in this case, as can be seen in line 17.

Line 19 prints out the current value of N and φ.

Line 20 sets the interval between successive values of N equal to 10, although if φ is greater than 15, the interval is changed to 5, in line 22.

Lines 23 to 25 set initial values for the suffices: nc is initially the 9 in equation A2.1. vc and vb are the sum and first term of the series, respectively.

PROGRAMME II

```
1     begin real va,vb,vc,vd,ve,vg,vh,vj,vl,vm,vn,vp,vq,vr,vs;
2           integer nb,nc,nd,ne,nh,nl;
3           switch ss:=sa,sb,sc,sd,se,sf,sg,sh,sj,input;

4     input: read nd;
5     sj:    nb:=1;
6     se:    vg:=(nd-10)*32/nb;
7            vm:=nb*0.022;
8            vn:=exp(vg*vm-nd);
9            vp:=vq:=1;
10           nl:=1;
11           vr:=(1-vn)/(1+vn);
12    sh:    vl:=vn*exp(-vm*nl);
13           vs:=0.99*(1-vl)/(1+vl);
14           vp:=vp*vs;
15           vq:=vq+vp;
16           nl:=nl+1;
17           if nl≠500 then goto sh;
18           vd:=vq*(1+vr);
19           print digits (3),nd,prefix (&& ??),nb,&psi?;
20           ne:=10;
21           if nb < 15 then goto sg;
22           ne:=5;
23    sg:    nc:=ne-1;
24           vc:=vb:=1;
25    sa:    nh:=nc;
26    sb:    ve:=exp(nb*0.022*(vg+nh)-nd);
27           va:=0.99*(1-ve)/(1+ve);
28           if nh=0 then goto sc;
29           if nh=ne then goto sd;
30           vb:=vb*va;
31           vc:=vc+vb;
32           nc:=nc-1;
33           if nc≠0 then goto sa;
34           nh:=0;
35           goto sb;
36    sc:    vh:=vc+vd*vb*va/(1+va/0.99);
37           nh:=ne;
38           goto sb;
39    sd:    vj:=vd/198.69;
40           print aligned (3,0),vg,prefix (&& ??),aligned (1,4),vj
41           vd:=(1+va/0.99)*vh;
42           vg:=vg+ne;
43           if vd > 30 then goto sg;
44           nb:=nb*2;
45           print &&L ?;
46           if nb < 33 then goto se;
47           nd:=nd+1;
48    goto   if nd=25 then sf else sj;
49     sf:   end of program;
```

Lines 26 to 33 calculate the sum of the first ten terms of the series in equation A2.1 (or first five terms of equation A2.2 where appropriate). When nh is zero the programme jumps to line 36 which calculates the last term in the series.

Line 37 sets nh to 10 (or 5) and y_{N+10} is formed in lines 26 to 29. The programme then jumps to line 39.

Line 39 normalises the result to give unity for small N: in this case g equals unity, and the series becomes

$$V = (1-0.99^{500})/(1-0.99) = 198.69$$

Line 40 prints the result V_{∞}^{N} with the corresponding value of N.

Line 41 calculates V_{∞}^{N+10} which is the starting point for the next set of calculations with N replaced by N+10, and N+10 by N+20.

Line 42 increases N by ten or five according to whether or not φ is greater than 15.

Line 43 stops the sums for a given value of φ when the result becomes less than 30, (c.f. an initial value of about 200).

Line 44 increases φ by a factor 2, and line 45 starts a new set of calculations with this value of φ.

Line 46 prevents calculations being carried out if φ is greater than 32.

Line 47 starts the programme all over again with a new

value of nd, the negative of the exponent of Z_0Y_0, up to a limit of 24 set by line 48.

APPENDIX III

The experimental conditions for a diffusion-controlled ultra-high-frequency discharge.

Before a gaseous discharge is termed "ultra-high-frequency" the following two experimental conditions must hold :

(1) the electrode separation must be greater than the mean free path of electrons, and

(2) the electrode separation must be greater than the distance which the electrons drift under the influence of one half-cycle of the applied field.

Clearly, the condition (1) is a function of the gas, the pressure and the electrode geometry, while (2) includes, in addition, the frequency and amplitude of the field.

1. Mean free path condition

The mean free path (ℓ) may be determined from the experimentally observed values of the collision probability (P_c) for electrons. In terms of this

$$\ell = 1/p_o P_c$$

where p_o is the gas pressure at a temperature 0^oC.

p_o can be found from the actual gas pressure, p, by use of the relation $p=p_o T/273$

where T is the gas temperature in ^oK.

P_c varies with electron energy, and therefore it is

necessary to use a representative energy. Examination of the published data for H_2 and N_2 shows that P_c increases with energy up to a maximum in the region of 4 volts and then decreases to half this maximum value at an energy about equal to the ionization potential. The mean electron energy is 4 volts for E/p equal to 30 and 100 volts/cm/torr for H_2 and N_2 respectively so that for the E/p values used here, the mean free path was about twice as long for electrons of ionizing energy as for the electron of average energy. In neon the value of P_c is approximately independent of the electron energy. As it is the electrons in the high energy "tail" of the distribution function which critically determine the breakdown voltage, it is important that the

	V_i(volt)	P_c($cm^{-1}torr^{-1}$)	p_o' (torr)
H_2	15.6	24	0.083
N_2	15.5	40	0.05
Ne	21.56	12	0.17

TABLE V

The collision probability of electrons at the ionization energy, and the least pressure which satisfies the mean-free-path criterion.

mean free path of these should be less than the gap width. The values for P_c [47] are therefore quoted here in Table V for the ionization energy of a molecule, together with the values for the lower pressure limit p_o' calculated from the condition (1) expressed as

$$d = 1/P_c p_o'$$

where d is the electrode separation (0.494cm).

2. Electron Ambit Condition.

During the intervals (t_c on average) between collisions with gas molecules, electrons are given an acceleration (eE_o/m) sin wt by the electric field. This gives rise to a velocity after t seconds of

$$v = (eE_o t_c/m) \sin wt$$

Here E_o is the peak value of the electric field

w is the angular frequency of the field

e and m are the electronic charge and mass.

The distance moved by an electron in the direction of the field during a quarter period is therefore on average

$$x = (eE_o t_c/m) \int_0^{\pi/2} \sin wt$$

$$= eE_o t_c/mw \qquad \text{.....} \qquad \text{A3.1}$$

which, in the limit imposed by condition (2) becomes equal to half the gap width.

therefore

$$d/2 = eE_o t_c/mw$$

It was shown in Chapter IX that the electron concentration varies sinusoidally across the gap, being zero at each electrode and maximum at mid-gap. Therefore more electron-molecule collisions occur near mid-gap than near the electrodes, and therefore the effective gap width is greater than the actual by a factor $2/\pi$, which is the average value of the sine function between 0 and π. Including this factor, the condition then becomes

$$d/\pi \geqslant eE_o t_c/mw$$

Taking into account the fact that the electrodes are not of infinite extent, we now substitute, in place of d/π, the diffusion length L, as defined in equation 8.6. Writing $t_c = \ell/v_r$ where v_r is the random electron velocity, and $\ell = 1/P_c p_o$ we have

$$L \geqslant eE_o/mwv_r P_c p_o \quad \ldots\ldots \quad A.3.1a$$

or

$$L \geqslant (E/p_o P_c w)(e/mu)^{\frac{1}{2}} \quad \ldots\ldots \quad A.3.1b$$

where E is the r.m.s. field $(=E_o/\sqrt{2})$ and u is the electron energy given by

$$\tfrac{1}{2} mv_r^2 = eu$$

Brode[47] has given P_c as a function of $u^{\frac{1}{2}}$. From his graph for H_2 it can be seen that $P_c u^{\frac{1}{2}}$ is approximately constant with an average value 90 $volt^{\frac{1}{2}}$ cm^{-1} $torr^{-1}$ over the range $1 < u < 70$ volts.

Substituting this value of $P_c u^{\frac{1}{2}}$ in equation A3.1b we have,

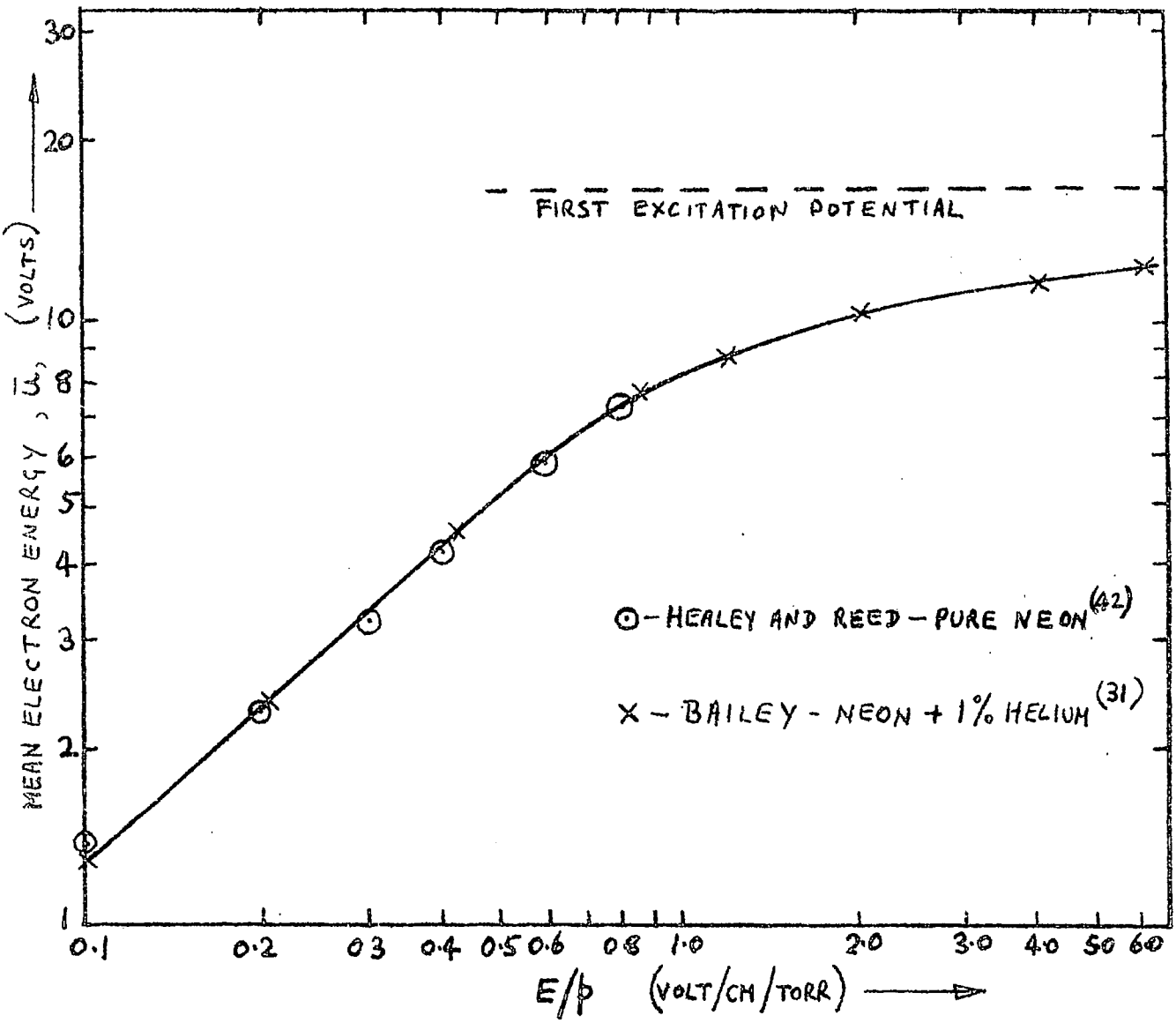


Fig. A3.1 MEAN ELECTRON ENERGY IN NEON

if $w = 3.6\pi 10^8$ sec,$^{-1}$, $L = 0.125$cm,

$$E/p_o \leq 302 \text{ r.m.s. volts/cm/torr.}$$

In N_2 the value of $P_c u^{\frac{1}{2}}$ is not constant, but for the purpose of obtaining a working limit we may set $P_c u^{\frac{1}{2}} = 100$ volt$^{\frac{1}{2}}$cm^{-1}torr^{-1}. Over the range $1.5 < u < 16$ volts the maximum deviation from this value is 40 volt$^{\frac{1}{2}}$ cm^{-1}.torr^{-1}.

We therefore obtain for nitrogen the condition

$$E/p_o \leq 340 \text{ r.m.s. volts/cm/torr.}$$

In Ne, P_c is almost constant, being 10 ± 3 cm^{-1} torr^{-1} over the range of u from 1 to 100 volts. Therefore the value of $P_c u^{\frac{1}{2}}$ increases steadily with electron energy. Reference to fig. 8.4. shows that the average electron energy increased rapidly with E/p, becoming 8 volts at E/p= 1 volt cm^{-1} torr^{-1}. However, at higher values of E/p (for which there are no published results) the average electron energy must tend to a limit not exceeding the ionization potential, and probably not greater than the first excitation level, 16.5 volts.

Fig. 8.4 has been redrawn for neon as fig A3.1 opposite, on a log-log scale, and extrapolated in the region of high E/p in such a way as to tend asymptotically towards the limit $\bar{u} = 16.5$ volts. As a guide to the extrapolation, Bailey's data[31] for neon containing 1% of helium was used.

Using the values $P_c = 10$ cm^{-1} torr^{-1},

$w = 1.1 \times 10^9$ radian/sec

$e/m = 1.76 \times 10^{11}$ coulomb/kg

$L = 1.25 \times 10^{-3}$metre

we obtain from equation A3.1b the relation

$$(E/p)u^{-\frac{1}{2}} \leq 33.6 \text{ r.m.s. volt}^{\frac{1}{2}}/\text{cm/torr} \quad \ldots\ldots \text{A3.2}$$

Table VI gives the value of $E/p/u^{\frac{1}{2}}$ as a function of E/p, as derived from fig A3.1.

E/p (rms volts/cm/torr)	0.1	1	10	100
$(E/p)u^{-\frac{1}{2}}$ (rms volts$^{\frac{1}{2}}$/cm/torr)	0.08	0.35	2.5	25

TABLE VI

The values of $E/p/u^{\frac{1}{2}}$ for neon as a function of E/p.

It is clear from Table VI that the condition A3.2 holds within a factor which is adequate to allow for errors in extrapolation. This can be seen to be especially true when it is remembered that the left-hand side of equation A3.2 only varies with $u^{\frac{1}{2}}$..

REFERENCES

1. Townsend J.S. Nature 62 (1900) 340
   Phil.Mag. 1 (1901) 198

2. Dutton J, Haydon S.C. and Llewellyn-Jones F. Proc.Roy. Soc. A213 (1952) 203

3. Fucks W. Appl. Sci.Res. B5 (1955) 109

4. Varnerin L.J. and Brown S.C. Phys.Rev. 79 (1950) 946

5. Macdonald A.D. and Brown S.C. Phys.Rev. 76 (1949) 1634

6. Reder F.H. and Brown S.C. Phys.Rev. 95 (1954) 685

7. Macdonald A.D. and Betts D.D. Canad.J.Phys. 30 (1952)565

8. Prowse W.A, Rowbotham J.R, and Monk P.G. E.R.A. Report L/T 386

9. Prowse W.A, Rowbotham J.R. and Monk P.G. Proc.Phys. Soc. 79 (1962) 158.

10. Prowse W.A. and Bainbridge G.R. Canad. J.Phys. 35 (1957) 324.

11. Long R.E. "An experimental study of the development of gaseous ionization at u.h.f." Ph.D. Thesis, Durham 1962.

12. Huxley L.G.H. "A survey of the principles and practice of waveguides" C.U.P. (1947).

13. Nichols M.J. "Measurements on the growth of gaseous ionization at u.h.f". PhD Thesis, Durham 1960.

14. Bainbridge G.R. and Prowse W.A. Canad.J.Phys. 34(1956) 1038.

15. Chance B. "Waveforms" M.I.T. Radiation Laboratory Series, vol 19.

16. Alpert D. J.App.Phys. 24 (1953) 860.

17. Jones C.V. ERA Rept L/T 334 (1956).

18. Jaffe A.A, Craggs J.D, and Balakrishnan C. Proc.Phys. Soc. B62(1949) 39.

19. Raethe H. Z.Phys. 110 (1938) 611.

20. Christoph W. Ann.Physik, 30 (1937) 446.

21. Greiner E. Z. Phys. 81 (1933) 543.

22. Weissler G.L. Proc. Roy. Soc. A220 (1953) 71.

23. Weissler G.L. Handbuch der Physik XXI p304 Springer-Berlin.

24. Reich H.J. "Microwave Theory and Techniques" van Nostrand (1954).

25. Mitchell A.C.G and Zemanski M.W. "Resonance Radiation and Excited Atoms" C.U.P. (1934).

26. Brown S.C. "Basic data of plasma physics" J.Wiley & Sons Inc (1959).

27. Slater J.C. Reviews of Modern Physics 18 (1946) 441.

28. Loeb L.B. Basic Processes of Gaseous Electronics P310, Univ. of Calif.Press (1955).

29. Von Engel A. "Ionized Gases" Clarendon Press (1955).

30. Townsend J.S and Bailey V.A. Phil.Mag. 42(1921)873.

31. Bailey V.A. Phil.Mag. 47 (1924) 379.

32. Crompton R.W and Sutton D.J. Proc.Roy. Soc. (London) A215 (1952) 467.

33. Brown S.C. and MacDonald A.D. Phys.Rev 75(1949)411.
Bowls. W.E. Phys. Rev. 53 (1938) 293.

34. Bradbury N.E. and Neilson R.A. Phys.Rev. 49(1936)388

35. Neilson R.A. Phys.Rev. 50 (1936) 950

36. MacDonald A.D, Gaskell D.U and Gitterman H.N. Phys.Rev. 130 (1963) 1841.

37. Githens S. Phys.Rev. 57 (1940) 822.

38. Prowse W.A. and Clark J.L. Proc.Phys.Soc.72 (1958) 625.

39. Pateyuk G.M. J.E.T.P. 3 (1956-7) 14.

40. Healey R.H. and Reed J.W. "The Behaviour of Slow Electrons in Gases". Amalgamated Wireless (Australasia) Ltd (1941).

41. Penning F.M. Physica Haag 8 (1928) 137.
Proc.Acad.Sci.Amst. 32 (1929) 341
Physica Haag 10 (1930) 47.

42. Von Engel A. Handbuch der Physik XXI p554 Springer-Berlin.

43. Townsend J.S. C.R. Acad. Sci.186 (1928) 55.

44. Allis W.P. and Rose D.J. Phys.Rev.93 (1954) 84.

45. Schneider F. Zeits. fur Angewandte Physik 6 (1954) 456.

46. Einstein A. Ann. der Physik 17 (1905) 4, 549
Zeits. f. Elektroch. 14 (1908) 235

47. Brode R.B. Revs. Mod. Phys. 5 (1933) 257.