A Theory of Localized Electronic Breakdown in Insulating Films

By N. KLEIN

Faculty of Electrical Engineering,
Technicon–Israel Institute of Technology, Haifa, Israel

ABSTRACT

A review of measured breakdown properties in insulating films and of electronic theories of breakdown shows both theoretical and experimental inconsistencies. A theory is proposed in which electronic breakdown is caused by local chance events, such as a succession of avalanches at one spot. Successive avalanches sustain the growth of space charges, the local cathode field, and the avalanche rate. When the cathode field becomes large enough to make continuation of avalanching a certainty, instability with current runaway arises, causing breakdown. According to theory, breakdowns occur over a range of fields, their chance increasing very strongly with field; the breakdowns occur randomly in space and in time; the time to instability on a breakdown event decreases as some exponential function of increasing field; the breakdown field can both increase and decrease with temperature; it may be electrode dependent, and it decreases first rapidly and then more slowly with increasing film thickness. Observations made in a number of insulating films support the proposed theory.

CONTENTS

§ 1. INTRODUCTION

NON-THERMAL electrical breakdown in solids has been the subject of investigations for many years. Considerable difficulties were encountered not only in the theoretical interpretation, but also in experimental work.

Shorting breakdowns and high test voltages made it very difficult to obtain reproducible results. The high voltages are avoided in thin films, but the testing of thin specimens encountered considerable problems. Many of the observations proved so erratic as a result of shorting breakdowns at weak spots that they were not suitable for comparison with theoretical models of the breakdown processes.

The situation has improved in recent years with the development of testing methods in which very large numbers of specimens are tested (Fritzsche 1967), or in which many tests are carried out on one specimen, when breakdowns are non-shorting (Inuishi and Powers 1957, Siddall 1959, Klein *et al.* 1965), or breakdowns are non-destructive (Riehl *et al.* 1969). These methods usually permit the identification of breakdown events due to weak spots and make it possible to determine a bulk breakdown value. The methods were mainly applied to tests on thin insulators and the experimental results obtained seem to represent basic breakdown properties.

The aim of this paper is to find an interpretation for the breakdown events in thin films, when the breakdown is not initiated thermally, by Joule heating. To recognize from the outset the requirements on a satisfactory interpretation, the paper will start with a summary of the breakdown observations in thin insulators, much of it obtained recently. The existing theories of breakdown in solids will be reviewed and examined for their applicability to thin film breakdown. The examination will reveal difficulties and a new statistical theory will be proposed. In this theory, initiation of breakdown will be ascribed to a localized chance event, such as a succession of avalanches, which are promoted by the interaction of electron injection and of space charges in the film. Support for the theory will be found in good qualitative and certain quantitative agreements with observations.

§ 2. Observations on the Properties of Non-thermal Breakdown

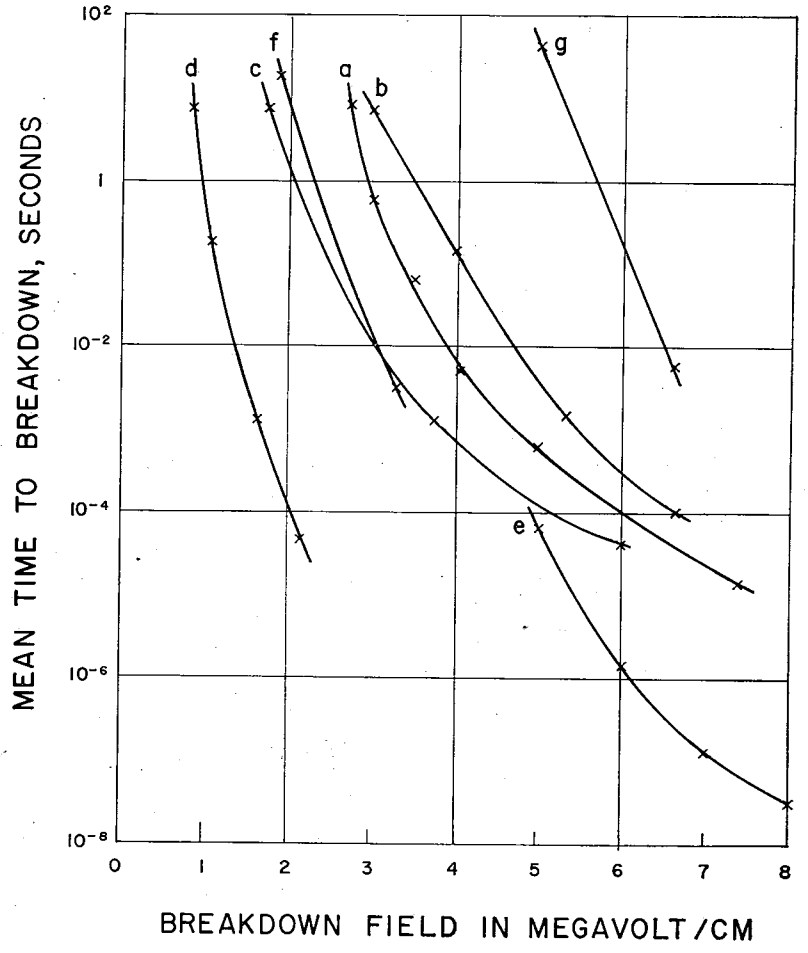
The breakdown properties which are reviewed here were obtained on planar specimens with vapour deposited electrodes or with movable spherical electrodes. Results of experiments, as shown in fig. 1 (*b*),–4 were obtained on specimens cleared of weak spots with non-shorting breakdowns. The temperature rise of the specimen due to leakage currents was found to be insignificant for the results shown in fig. 1 (*b*),–4. This excludes thermal mechanisms as the primary cause of breakdown.

Dependence of time to breakdown on the breakdown field, temperature, thickness and electrode material were observed, certain statistical properties and also data on the mechanism of breakdown, detailed below.

(1) The dependence of τ_b, the mean time for the application of voltage to the completion of breakdown to the breakdown field F, is presented in fig. 1 (Plessner 1948, Kawamura *et al.* 1954, Smith and Budenstein

1969, Klein 1969 *b*). Here fig. 1 (*a*) includes test results, some of which may have been obtained at weak spots, while the results of fig. 1 (*b*) were obtained on specimens cleared of weak spots. The insulators tested, with the exception of mica, were either polycrystalline, or amorphous. Figure 1 shows that breakdown may occur over a wide range of fields

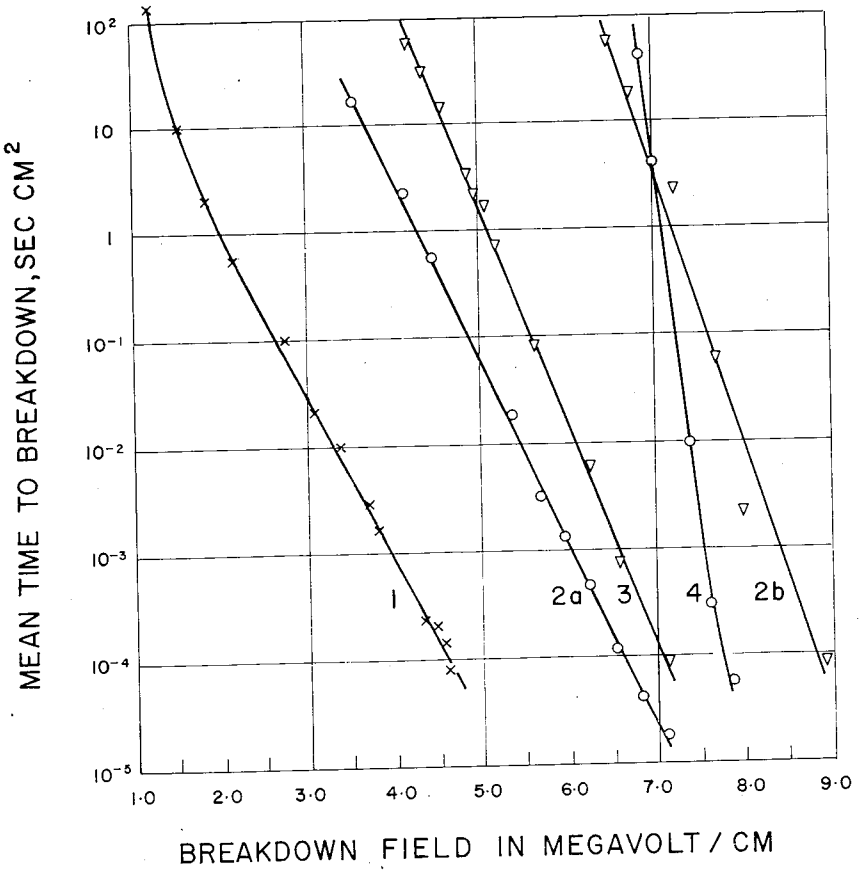
Fig. 1



(*a*)

(*a*) Mean time to breakdown τ_b as function of breakdown field in:
(*a*) 2700 Å thick CaF_2
(*b*) 1900 Å thick NaF
(*c*) 530 Å thick KBr
(*d*) 10 800 Å thick KBr
} (from Plessner 1948)
(*e*) 25 000 to 35 000 Å thick mica (from Kawamura *et al.* 1954)
(*f*) 2400 Å thick CaF_2 at 28°C
(*g*) 2400 Å thick CaF_2 at −176°C (from Smith and Budenstein 1969)
Ramp voltage tests, except for (*e*), obtained with step voltages.

Fig. 1 (*continued*)



(*b*)

(*b*) Mean time to breakdown τ_b as function of breakdown field on step voltage tests in:

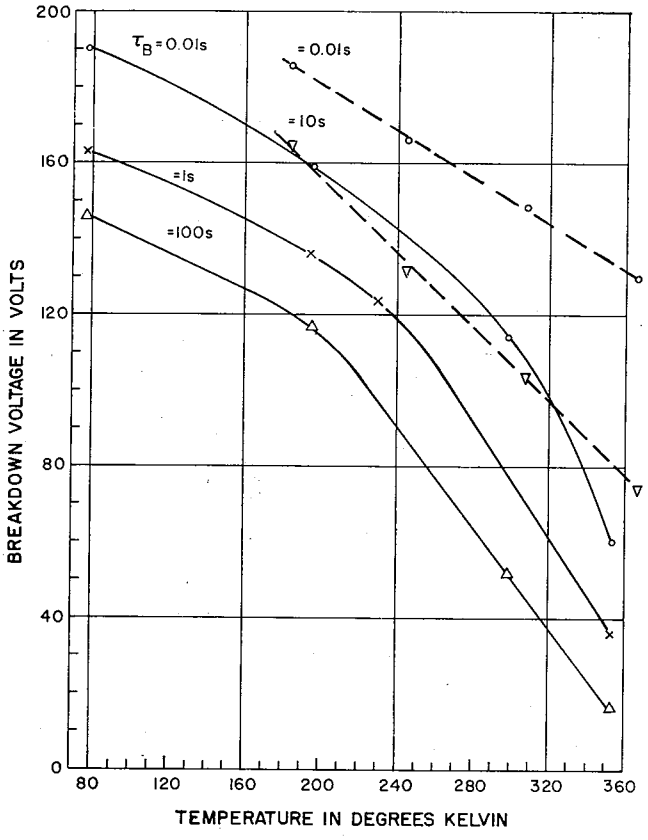
(1) Hf–HfO_2–Au, Hf neg., oxide 3250 Å thick;
(2) Al–Al_2O_3–Au, oxide 1700 Å thick: (*a*) Al neg; (*b*) Au neg.
(3) Ta (some Mb)–Ta_2O_5–Au, Ta neg., oxide 2300 Å thick;
(4) Si–SiO_2–Al, Si pos., *n*-type, 0·01Ω–cm, oxide 1500 Å thick. Room temperature (from Klein 1969 b).

and that the average value of τ_b decreases quasi-exponentially with increasing field. The curves for CaF_2, KBr and HfO_2 in fig. 1 (*a*) and (*b*) indicate that there may be a low-field limit for the occurrence of breakdown. In vapour grown silicon oxide at high fields for breakdown for a decrease in τ_b from 5×10^{-6} to 5×10^{-8} sec no increase in F was observed (Klein and Burstein 1969). In contrast to the results in fig. 1, no field dependence for τ_b was found in polystyrene films not cleared of weak spots (Plessner 1948).

(2) Dependence of F on electrode material, or polarity, is shown for Al_2O_3 by the curves 2*a* and 2*b* in fig. 1 (*b*). F is quite different, depending whether the gold or the aluminium electrode is negative (Klein 1969 *b*). Small dependence of F on electrode material was found in silicon dioxide (Chou and Eldridge).

(3) Figure 2 shows that F decreases with increasing ambient temperature T_0 in HfO_2 (Klein 1969 b) and Ta_2O_5 (Klein *et al.* 1970). Similar dependence was also found in a number of halide films and CeO_2 (Smith and Budenstein 1969, Budenstein *et al.* 1969), although an increase in F with T_0 was observed in CaF_2 and LiF below $-150°C$ (Smith and Budenstein 1969, Budenstein *et al.* 1969). In the range of electronic breakdowns in vapour grown silicon oxide (Klein and Burstein 1969), and thermally grown silicon dioxide (Chou and Eldridge) F was found to depend but little on T_0.

Fig. 2

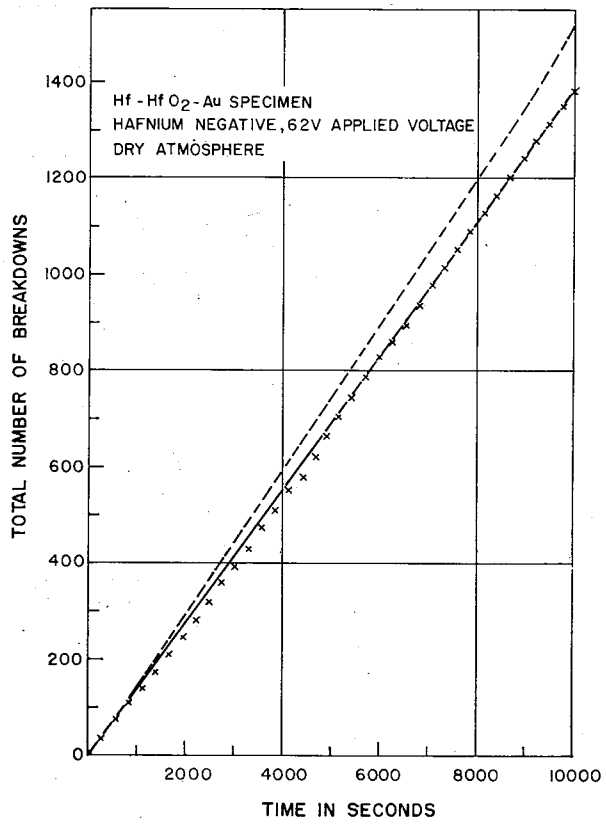


Breakdown voltage versus temperature with mean time to breakdown as parameter. Step voltage tests. Full lines Hf–HfO_2–Au, Hf neg., oxide 3250 Å thick (from Klein 1969 b); broken lines Ta(1·5%Mb)—Ta_2O_5–Au, Ta neg. oxide 2200 Å (from Klein *et al.* 1970).

(4) The breakdown field was usually found to decrease with increasing thickness w (Austen and Whitehead 1940, Plessner 1948, Lomer 1950), and for instance recent tests (Agarwal and Srivastava 1971) on barium stearate films showed that for $w < 250$ Å the breakdown field was proportional to $w^{-1.01}$ and for $w > 250$ Å the breakdown field was proportional to $w^{-0.59}$.

(5) A further important property is observed in many substances when uniform specimens showing nonshorting breakdowns are tested. The breakdowns occur at random locations and each breakdown affects only a very small area. This area was found to be probably much less than 10^{-8} cm^2 in 0·2 cm^2 thermally grown silicon dioxide specimens (Klein 1966) and less than 10^{-10} cm^2 in about 10^{-4} cm^2 mica specimens (Davidson and Yoffe 1968).

Fig. 3



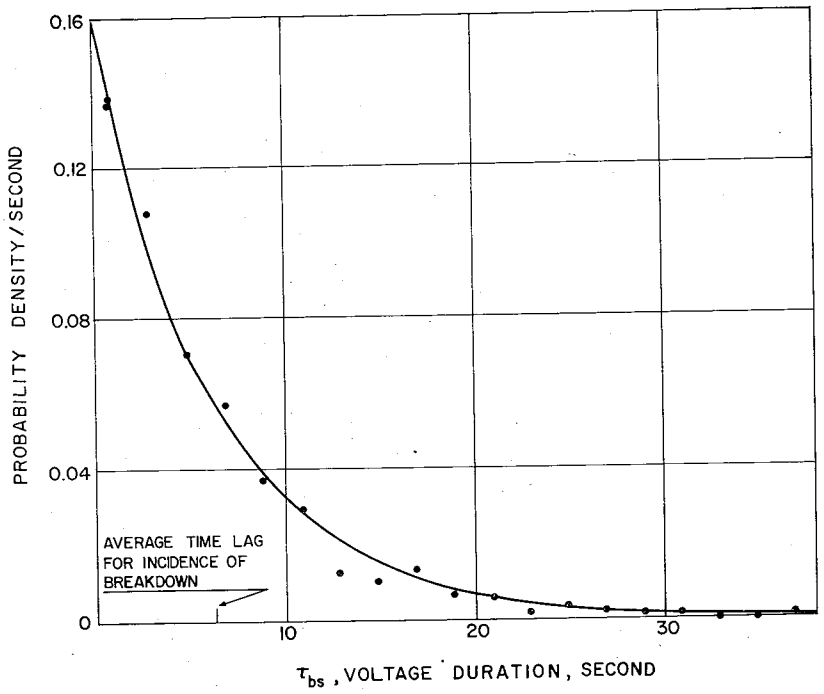
Total number of breakdowns versus time on the application of 62 v to a 3250 Å thick, 0·1 cm^2 HfO_2 specimen. Crosses indicate batches of about 40 breakdowns (from Klein 1969 a). Broken line corrected for decrease in specimen area with increasing number of breakdowns.

When a constant voltage is applied to a specimen, breakdown events repeat at an average breakdown rate $R_b = 1/\tau_b$. R_b may, in certain cases, remain practically constant over more than 1000 events, as illustrated for a HfO_2 specimen in fig. 3 (Klein 1969 a). In such an experiment there is a large scatter in τ_{bs}, the time between consecutive breakdowns. Plotting the probability density of events occurring after time τ_{bs} versus τ_{bs} in fig. 4, an exponential decrease is found for the probability density (Klein 1971), indicating a Poisson distribution for the incidence of breakdown. τ_{bs} has the nature of a statistical time lag and the breakdown events appear to be random, not only spatially but also in time.

(6) Some experiments were carried out to identify the mechanism which initiates breakdown. On breakdown events in alkali halides, Cooper and Elliot (1966) observed light emission starting at the cathode a few tens of nanoseconds prior to voltage collapse. The light emission was ascribed to recombination processes of carriers produced by impact ionization, indicating that breakdown starts by avalanching.

Van Geel *et al.* (1957) observed the anodic oxidation of aluminium and tantalum at fields 5–7 Mv/cm. The oxidation is accompanied by localized breakdowns and emission of electroluminescent light. The intensity of light emission increases exponentially with thickness of the

Fig. 4



Probability density of breakdown occurring after voltage duration τ_{bs} versus τ_{bs} in the experiment of fig. 3 (from Klein 1971).

oxide layer. This was ascribed to recombination of carriers produced by impact ionization indicating avalanching in the range of breakdown.

Williams investigated the effect of fields close to breakdown on free carriers injected into an insulator conduction band by the technique of internal photo-emission. He observed multiplication of the carriers in such medium bandgap substances as ZnSe (Williams 1967). In contrast, multiplication was not found in CdS (Williams 1962).

§ 3. Electronic Interpretations of Breakdown

Although the mechanism of non-thermal breakdown in solids has been ascribed to ionic processes, to mechanical rupture, and to decomposition of the substance by the high field, the most frequent interpretations are electronic. Processes such as double injection, insulator to metal transition, field emission, impact ionization, electrode injection, space charge effects, or their combination, have been proposed as the cause of electronic breakdown. In some unusual cases, breakdown was ascribed to electro-acoustic and electromechanical effects.

Double injection and insulator to metal transition processes have been used to interpret some events occurring at fields below 10^5 v/cm. At the much higher breakdown fields in insulators, breakdown theories have been based mainly on impact ionization and/or field emission in interaction with electrode injection and space charge effects.

The insulator breakdown theories will now be briefly reviewed, to see whether they provide adequate models for breakdown in thin films.

(1) The original theories of von Hippel and Fröhlich (reviewed in Franz 1956, Stratton 1961, O'Dwyer 1964), ascribe the breakdown in insulators to electron avalanches produced by impact ionization. These older theories established the concept of the intrinsic breakdown field, which is independent of the duration of voltage application, the electrode material and the insulator thickness, in disagreement with the thin film observations. It was postulated that intrinsic breakdown occurs at that field at which the energy gain of the conduction electrons from the field cannot be balanced by phonon collisions. It was assumed that the consequence of this lack of balance was impact ionization with subsequent current runaway and thermal destruction.

(2) The example of the stable d.c. operation of Zener diodes in the avalanching range shows that the occurrence of impact ionization need not entail current runaway and destruction. In subsequent development of the early theories the breakdown field was determined not by the onset of impact ionization, but by the conditions for destruction (Davydov and Shmushkewitch 1940, Heller 1951, Franz 1952, Yamashita and Watanabe 1952, 1954). Assuming current increase by impact ionization, or by field emission from the valence band, time dependent current and heat balance equations of the specimen were derived. The equations were solved for the conditions of breakdown. This was assumed to arise when the Joule

heat due to the increased carrier density caused a critical temperature rise accompanied by melting, or even evaporation in the insulator. Calculations resulted in a quasi-exponential decrease of time to breakdown with increasing field.

Forlani and Minnaja (1964) extended the model by considering the origin of the charge carriers. They assumed that charge carriers are injected from the cathode into the insulator by field emission, and are then multiplied by impact ionization. They too determined the breakdown field from the destructive effect of the Joule heat. This theory predicted dependence of the breakdown field also on the electrode.

These theories were based on many simplifying assumptions, an important one being the neglect of space charges. Space charges, however, were bound to arise, because the positive charges left behind by the ionizing electrons were assumed to be immobile relative to the ionizing electrons. These positive charges decrease the field and the ionization rate at the anode and in consequence, the large scale avalanching required by the theories for the breakdown could not develop (O'Dwyer 1967).

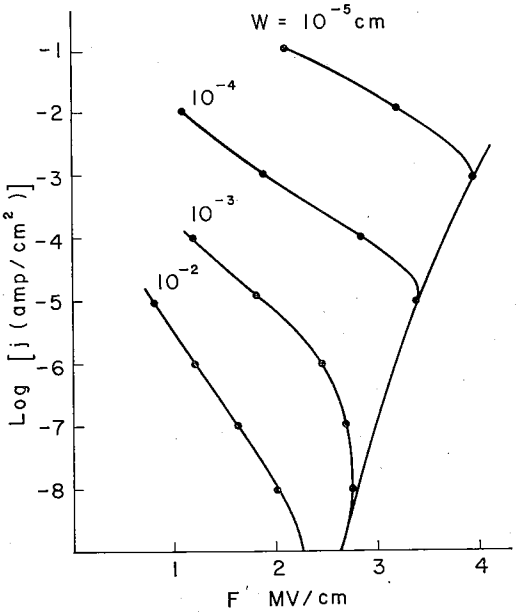
(3) To take account of the space charge effects, Poissons' equation is added to the current continuity, and when necessary to the heat transport equation and the set of equations is solved for the condition of current runaway and breakdown. The solution of these equations is usually difficult, and only a few d.c. solutions are available for cases in which the runaway is not influenced by thermal effects. Thus Egawa (1966) and Misawa (1966) determined breakdown conditions for avalanching p–n junctions and O'Dwyer (1969) for insulators, when the runaway is caused by electronic instability.

In O'Dwyer's model, highly mobile electrons injected into the insulator at the cathode by field or thermionic emission produce avalanching. The very slow positive charges created by impact ionization enhance the cathode field and thus the electron injection, and when this positive feedback effect is sufficiently large, instability arises. The steady-state solution of the set of equations for this model is illustrated for specimens of several thicknesses by calculated current density versus field characteristics in fig. 5. The highest field in these curves is the d.c. breakdown field, at which instability and current runaway may arise. As destruction arises at much higher current densities and lower fields, the breakdown field in this model is not determined by the condition for destruction, but by that for current instability.

Transient solutions for voltage pulses are not available for the set of equations. Baessler *et al.* (1969) developed an approximate theory and predicted a quasi-exponential decrease in time to breakdown τ_b with increasing field.

(4) While O'Dwyer and Baessler's theories are in qualitative agreement with the thin film observations, new difficulties are introduced. These and the earlier theories are continuum theories with uniform charge

Fig. 5



Calculated steady state F–j characteristics of specimens with field emission at the cathode and avalanching in the insulator, causing electronic instability. The parameter is the specimen thickness (from O'Dwyer 1969). α is given by eqn. (3).

injection and avalanching in the whole specimen. Injection, however, is not uniform, but by single electronic charges producing single avalanches and the question arises whether these avalanches fill the volume of the thin film insulator. This may be assessed, because the volume fraction f occupied by the positive charge clusters of the avalanches is roughly $j_c A_v t_t/e$, j_c being the injected current density, A_v the avalanche cross section, t_t the transit time of the positive charges and e the electronic charge. Interpretation of the test results on oxides will show that in the range of breakdown observations the volume fraction is very small and, for instance, in the silicon dioxide experiments f may have varied from 10^{-5} to 10^{-3}. The avalanches which are produced in these cases are localized and spaced widely from each other. When breakdown is the consequence of a single avalanche, continuum theories do not give an adequate description of the events.

A theory for localized breakdown was given earlier by the single electron avalanche breakdown theory of Fröhlich (1940) and Seitz (1949). According to this theory on breakdown, an injected electron produces an avalanche of a magnitude, sufficient to cause at least local melting. While this theory predicts electrode and time dependence of the breakdown field and random localization of the event, it disregards, as do the earlier

continuum theories, the effect of positive charges remaining in the insulator and the occurrence of instability prior to destruction. If account is taken of these effects, the growth of an individual electron avalanche is found to be quenched by the positive charge left in the insulator. The avalanche remains small and can produce small temperature rises but not breakdown. Thus, the single electron avalanche breakdown model is not adequate for the interpretation of the events either.

(5) If a single avalanche is insufficient, a rapid succession of avanlanches may produce breakdown as suggested recently by Watson *et al.* (1970). Such processes have been shown to lead to breakdown in gases and in the following a theory will be developed describing how a succession of avalanches can produce breakdown in thin films.

Of course at very high, hypothetical fields the injected current density may be so large that rapid successive avalanching and breakdown is a certainty. The essential point that will arise here is that breakdowns are observed locally as chance events at much lower fields. Breakdowns occur because there is a finite probability, strongly increasing with field, that a succession of avalanches or some other fluctuation phenomena cause local instability and current runaway. The statistical breakdown theory that follows will be based on the mechanism of a local succession of avalanches. It will be shown that breakdown occurs over a wide range of fields and that the mean time to the incidence of breakdown is a strong function of the applied field.

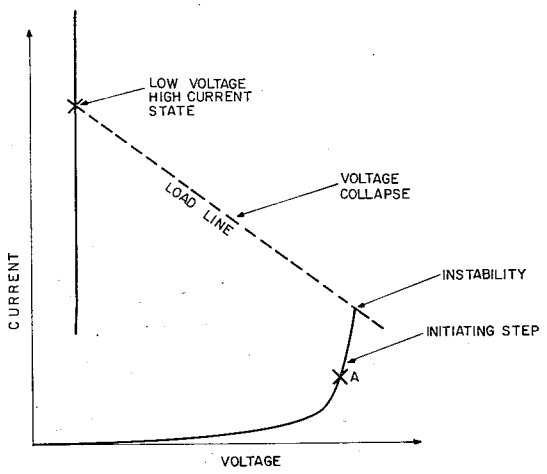
§ 4. A Model of Localized Electronic Breakdown

The development of the breakdown model will be started by some phenomenological considerations. While breakdown events may be initiated and developed in different substances by a variety of processes, phenomenologically the events have much in common (Klein 1969 b). This is apparent, when voltage versus time oscillograms of breakdown events and the steady state V–I characteristics of various specimens are compared. Such characteristics, shown schematically in figs. 5 and 6 may either be continuous or consist of separate high voltage—low current and high current—low voltage branches. The breakdown event can be studied with the help of these characteristics.

Breakdown develops in a sequence of stages (Klein *et al.* 1965, Berglund and Klein 1971): (1) the initiating stage increasing the electrical conductance leading to (2) instability causing current runaway and resulting in (3) voltage collapse across the specimen, with local melting and vaporization. If one employs a very thin electrode the metal vaporizes over the breakdown spot, producing a non-shorting hole, so the specimen reverts after the destructive stage to the insulating state. On the other hand, when the breakdown is shorting and the series resistor very large, the specimen may settle down during a fourth stage to a low voltage—high

current state. The third and fourth stage in a breakdown event can be readily interpreted. The destruction during the voltage collapse is caused mainly by the discharge of the electrostatic energy stored in the specimen through the breakdown spot and this was confirmed by the agreement found in the magnitude of measured and calculated volumes vapourized on non-shorting breakdowns (Klein 1966, Klein and Burstein 1969). In the case of shorting events the material changes lead to the formation of a filament of high conductance between the electrodes and to a high current—low voltage state. As the dominant processes may change from stage to stage, it is important to discuss the individual stages separately. The questions which are the subject of this paper relate to the mechanisms of the initiating and instability stages.

Fig. 6



Schematic steady state, discontinuous V–I characteristics of specimen in which switching or breakdown occurs.

The statistical model for the interpretation of localized breakdown events in thin films will be developed according to the following outline: An electron injected into the conduction band of an insulator is the initial event. The electron causes impact ionization and produces a finite avalanche of free electrons, leaving behind nearly immobile positive charges. The electrons are swept out from the insulating film, typically in picoseconds, while the positive charges drift to the cathode slowly in micro- to milliseconds. The positive charges left in the insulator prevent the formation of large, destructive avalanches and enhance the field at the cathode. In most cases the positive charge clusters leave the insulator without further effects. However, the enhancement of the field at the cathode greatly increases the local injection rate of electrons, so there is a finite probability that an injected electron will hit the tiny charge cluster during its transit through the insulator. When this happens,

there is further avalanching with its attendant field enhancement and increase in the chance for successor avalanches. When this build up results in certainty of electron injection during the transit time of positive charges, local instability with current runaway arises. This causes collapse of the voltage and destructive breakdown at a small spot.

The elaboration of this model involves the study of the interactions of a number of processes, each of which may be complicated and may originate from several causes. Consideration of the details of the individual processes would make the treatment very difficult and cumbersome. As we are interested in presenting and discussing only the salient features of the localized electronic breakdown model, a number of simplifications will be made, which we will attempt to justify as the discussion proceeds.

4.1. *The Injection of Charge Carriers*

The electrons initiating the avalanches can be injected into the conduction band of the insulator by a number of processes, such as thermionic emission with Schottky lowering of the barrier, thermionic-field emission, field emission from the cathode, hopping from the electrode to the conduction band and possibly by photoionization by recombination light. In the continuation use will be made of the simplest forms of the relations for the current density j_{ct} by thermionic and j_{cf} by field emission and

$$j_{ct} = AT^2 \exp\left(-\frac{E - bF^{1/2}}{kT}\right), \quad \ldots\ldots\ldots \quad (1)$$

and

$$j_{cf} = CF^2 \exp(-D/F). \quad \ldots\ldots\ldots \quad (2)$$

Here

$$A = 4\pi m^* ek^2/h^3; \quad b = (e^3/4\pi\epsilon_0\epsilon)^{1/2}; \quad \ldots\ldots \quad (1\,a)$$

$$C = e^3/8\pi hE; \quad D = 4(2m^*)^{1/2}E^{3/2}/3e\hbar; \quad \ldots\ldots \quad (2\,a)$$

with T the temperature, E the energy difference between the Fermi level of the cathode and the conduction band edge of the insulating film when no field is applied, F the electric field, k the Boltzmann constant, m^* the effective mass of the electron, e the electronic charge, h the Planck constant, ϵ_0 the dielectric permittivity of free space and ϵ the relative permittivity of the insulating film.

4.2. *The Avalanche*

The injected electrons are assumed to cause impact ionization either by interband transitions, or by impurity ionization. Very little is known by direct experimental evidence of the nature of these ionization processes. In contrast to the common semiconductors, data on the field dependence of the ionization coefficients of insulators are hardly available. It has

usually been assumed (O'Dwyer 1964), for calculation of breakdown fields, that the ionization coefficient α obeys the relation

$$\alpha = (B/F) \exp(-H/F), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

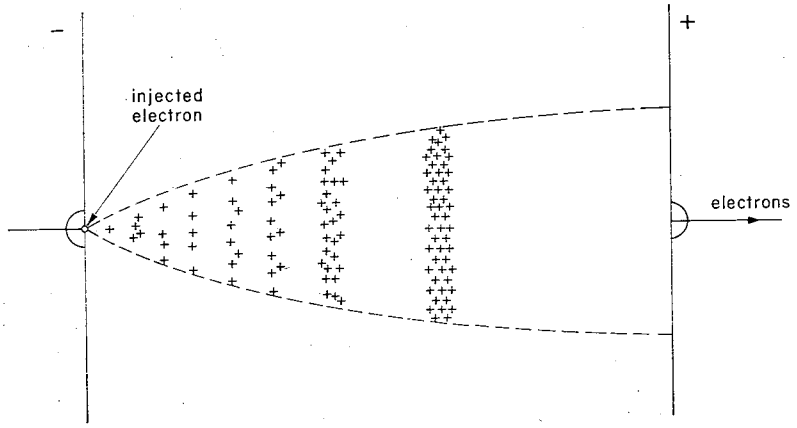
where B and H are constants of the insulator. Since the observed breakdown fields extend over a wide range in most insulators tested (fig. 1), eqn. (3) will be applied to 'low fields' only, while for 'high fields' a relation proposed earlier for semiconductors (Shockley 1961) is found more suitable:

$$\alpha = F/V_i, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

where V_i is the ionization voltage. We will continue to develop the theory with eqn. (4) but will return later to eqn. (3), to show its significance in the interpretation of breakdown events at the low field limit for breakdown.

An electron injected at the cathode produces an avalanche as shown schematically in fig. 7. This figure is assumed to present the situation immediately after the electrons have been swept out at the anode with the positive charges essentially still at the places where impact ionization took place. If fluctuations in collision distances are disregarded, the positive charges appear schematically in layers separated by mean free paths for ionization by electrons. As a result of the increasing opposing field of the positive charges the mean free path for ionization increases towards the anode. The last charge layer is at a distance a_v from the cathode, a_v denoting the mean avalanche length. The effective field acting on the electrons in the film is so low beyond a_v that no more impact ionization occurs, when a_v is not much smaller than the insulator thickness w.

Fig. 7



Schematic picture of an avalanche produced by a single electron, showing the positive charges left in the insulator after the electrons were swept out at the anode.

Denoting by i the number of ionizing collisions of the injected electron in its course through the insulator, by V_i the ionizing potential, and by F the applied field, the avalanche length is:

$$a_v = \kappa i V_i / F. \quad \ldots\ldots\ldots \quad (5)$$

Here $\kappa > 1$ is a factor taking account of the average increase of the mean free paths for ionization towards the anode. The electrons diffuse sideways in their progress towards the anode and the cross section of the avalanche cone A_v at the last ionizing collisions (Seitz 1949) is $4\pi l_p a_v/3$, where l_p is the mean free path for phonon collisions. Replacing for a_v from eqn. (5)

$$A_v = 4\pi\kappa i l_p V_i / 3F. \quad \ldots\ldots\ldots \quad (6)$$

The number of electronic charges N generated by impact ionization is 2^i. The field of the positive charges immediately beyond the last ionization opposing the applied field is somewhat less than $2^i e/\epsilon_0 \epsilon A_v$. Ionization stops after the ith collision, because the opposing field becomes a considerable fraction of the applied field. The positive charge is then related to the applied field by

$$\sigma F = e 2^i / \epsilon_0 \epsilon A_v, \quad \ldots\ldots\ldots \quad (7)$$

where σ is a coefficient and $0 < \sigma < 1$. Using eqns. (6) and (7) a relation is obtained for N the number of charges of one polarity generated in the avalanche

$$N = 2^i = \frac{4\pi\epsilon_0}{3e} \sigma\kappa i \epsilon l_p V_i. \quad \ldots\ldots\ldots \quad (8)$$

N, the size of the largest avalanche that can develop, is independent of the field, while by eqns. (4) and (5) the avalanche dimensions a_v and A_v are inversely proportional to the field. When $a_v > w$ full size avalanches do not develop and the avalanche size decreases with field.

Since breakdowns in some oxides are observed over a wide range of fields (fig. 1), impact ionization may be assumed to start at relatively low fields. The coefficient σ expressing the influence of the opposing field of the positive charges may therefore reach values not much less than unity. Assuming values for σ between 0·7 and 0·9 the stepwise development of avalanches from ionization to ionization was calculated. From such calculations, the factor κ in eqn. (5) for the avalanche length was found to vary roughly from 1·4 to 2.

As typical variations in σ and κ appear to be small, the avalanche size by eqn. (8) is mainly determined by material properties, expressed by the product $\epsilon l_p V_i$. With $4\pi\epsilon_0/3e = 2{\cdot}31 \times 10^8$ v^{-1} m^{-1}, $\sigma = 0{\cdot}85$ and $\kappa = 1{\cdot}8$, several assumedly limiting cases were calculated for the avalanche sizes. A low value for N and i resulted, when $\epsilon = 4{\cdot}5$, $l_p = 5 \times 10^{-10}$ m and

$V_i = 5$ v and a high value, when $\epsilon = 40$, $l_p = 2 \times 10^{-9}$ m and $V_i = 12{\cdot}2$ v. The results are presented in table 1.

Table 1. Some single electron avalanche data

	N	i	a_v,cm	A_v,cm^2	θ,°c	τ_{th},sec
Low values	16	4	$3{\cdot}6 \times 10^{-6}$	$7{\cdot}5 \times 10^{-13}$	0·17	5×10^{-11}
High values	4096	12	$6{\cdot}6 \times 10^{-5}$	$5{\cdot}6 \times 10^{-11}$	54	4×10^{-9}

Typical low and high values for a_v and A_v obtained with eqns. (5) and (6) are also given in table 1. For the calculations, the earlier material properties and $F = 10^7$ v/cm were assumed for the case of low values for N and $F = 4 \times 10^6$ v/cm for the case of high values.

To assess the local temperature rise θ caused by the avalanche, it will be assumed that the total charge passes the insulator so fast that heat conduction is negligible. Then $\theta \simeq$ New F/cwA_v where c is the specific heat of the insulator per unit volume. Replacing for A_v from eqn. (6)

$$\theta = \frac{\sigma \epsilon_0 \epsilon F^2}{c}. \qquad \ldots\ldots\ldots \quad (9)$$

The time needed for the dissipation of the heat pulse caused by the avalanche can be estimated. Assuming that the heat is conducted radially in the insulator, θ at the centerline of the avalanche decreases by 80% after a time $\tau_{th} \simeq r_v^2/\kappa_{th}$ where r_v is $(A_v/\pi)^{1/2}$ and κ_{th} is the thermal diffusivity of the insulator (Carslaw and Jäger 1959).

Presumably low and high values of θ and τ_{th} were calculated and given in table 1, for θ, when $\epsilon = 4{\cdot}5$ and $F = 10^6$ v/cm and when $\epsilon = 40$ and $F = 6 \times 10^6$ v/cm. The values of τ_{th} were obtained with the A_v values in table 1 and $\kappa_{th} = 5 \times 10^{-3}$ cm^2/sec.

The values of θ and τ_{th} are both overestimated. Those of θ are overestimated because the electronic pulse does not cross the whole insulator, and those of τ_{th} because heat conduction takes place not only radially, but also towards the electrodes.

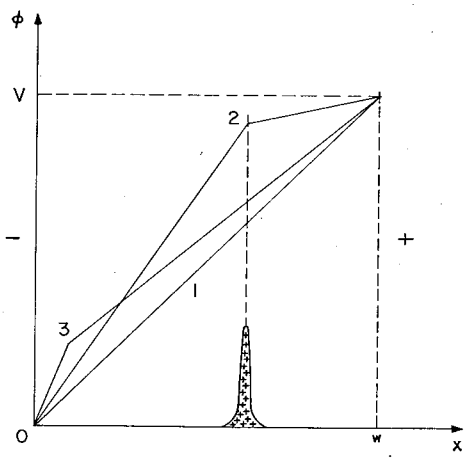
The picture which emerges from these simple considerations is that of very small avalanches, which are far from being destructive and may cause only small local temperature rise.

4.3. *The Enhancement of the Cathode Field*

It is possible to calculate accurately the field distortion caused by the positive charges remaining in the insulator (Davies *et al.* 1964, Davies and Evans 1967). Most of the charges are generated where the avalanche cone has grown to its fullest. It will be sufficient for present purposes to assume that the positive charges are all located where the last ionizations

took place. This is illustrated in fig. 8, where the electrostatic potential ϕ in the insulator is plotted versus the spatial coordinate x perpendicular to the electrodes. The positive charge is initially located at $x=a_v$. The field is uniform prior to the avalanche $F=V/w$ with V the voltage across the specimen.

Fig. 8



The electrostatic potential Φ as function x perpendicular to the electrodes ,where a non destructive avalanche was produced. Positive signs denote the idealized positive charge cluster. Curve 1, Φ prior to impact ionization; curve 2, after the electrons have been swept out; curve 3, at some x during the drift of the positive charge.

If the positive charge of the avalanche would be uniformly distributed in the plane at $x=a_v$, the charge density would be $\sigma\epsilon_0\epsilon F$. If such a charge density would occur over a large area compared with the thickness w, the field at the cathode would be $F[1+\sigma(w-a_v)/w]$. The avalanche charge is not uniformally distributed and extends only over a small area A_v, hence σ is replaced by a smaller factor γ_0 and

$$F_{co}=F\left(1+\gamma_0\frac{w-a_v}{w}\right). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

F_{co} is the field at the cathode, enhanced by the positive avalanche charge, immediately after the departure of the electrons from the insulator. The potential distribution at this time is represented by graph 2 in fig. 8, indicating the small field remaining between $x=a_v$ and the anode.

While the positive charge cluster drifts towards the cathode, both the cathode and the anode fields increase (graph 3 in fig. 8) and the cathode field F_c is

$$F_c=F\left(1+\gamma\frac{w-x}{w}\right). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

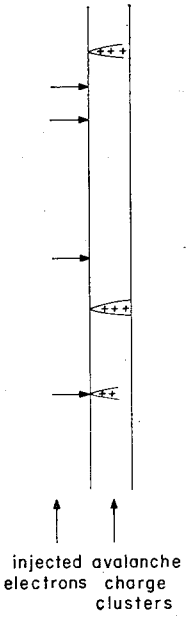
Here γ is the factor, which was γ_0 in eqn. (10) at the start of the drift. The factor γ increases during the drift for geometrical reasons, but decreases by the effects of diffusion and local field on the charge cluster. The field at the cathode is limited to twice the applied field, when γ approaches unity and the field at the anode increases towards the applied field during the drift.

While the cathode field F_c will not become twice the average field F at a positive charge cluster, F_c is on the average expected to be larger than F by several tens of per cent during the drift of the cluster.

4.4. *Instability by a Local Succession of Avalanches*

The magnitude of the cathode field enhancement F_c/F is very important for the continuation of the event towards localized breakdown. It is shown in Appendix that if breakdowns were not localized but involved the whole insulator, then by O'Dwyers' (1969) continuum theory, instability with current runaway can arise even at a few per cent cathode field enhancement. When breakdowns are localized current injection is non uniform. Where injection occurs the cathode field enhancements are much larger than they would be at instability according to the continuum theory. The insulator appears therefore to be potentially unstable at a positive charge cluster during its transit.

Fig. 9



Among the electrons injected into the insulator film only those hitting a positive charge cluster can promote breakdown.

The meaning of the potential instability at the charge cluster becomes clearer, when the development of the chance event, of a succession of avalanches leading to breakdown is analysed.

When by chance, as shown in fig. 9, an injected electron hits the charge cluster during transit, a second avalanche is produced in the enhanced field at the cathode. According to eqn. (5), the second avalanche is shorter than the first avalanche. The remaining positive charge of the first and that of the second avalanche further enhance the cathode field. This further increases the probability of electron injection at the cluster and the production of a third avalanche and so on. We investigated various avalanche sequences, changing intervals between successive avalanches, and found that the cathode field increases with the number of successive avalanches. The field increase was most pronounced by the production of a subsequent avalanche, at a time when the charge cluster has already drifted close to the cathode.

Increase of the cathode field by the succession of avalanches can lead to instability. Instability arises when the avalanches build up the cathode field to such an extent as to make it certain that an injected electron will hit the charge cluster during its transit. Once instability has been reached, the local avalanches succeed each other at a rapidly increasing rate, resulting in current runaway and breakdown.

In this description of the instability process we have not yet mentioned recombinations. Such processes are likely to accompany the event and their effect is to slow down the development to instability.

4.5. *The Mean Time to Instability*

The localized breakdowns occur at an average rate R_b. This is proportional to the rate of electron injection causing primary avalanches, and to the probability P that upon a primary avalanche a succession of avalanches will occur, resulting in local instability. The inverse of the rate of breakdown R_b is the mean time to instability for breakdown τ_s and

$$\tau_s = e/j_c(F)A_s P, \qquad (12)$$

with A_s the specimen area. Disregarding fluctuations in injection and in the subsequent processes, relations will now be established for τ_s, when instability is the outcome of a succession of one, or two, or n avalanches locally.

In the extreme case, at very high fields, the first avalanche is sufficient for instability, the succession of avalanches is certain, $P=1$, and every injected electron causes breakdown. In this case at least one electron hits the charge cluster during its transit and the relation for instability is

$$j_c(\beta_1 F)A_{v1}t_{t1}/e \geqslant 1. \qquad (13)$$

Here $\beta_1 = F_c/F$ is the effective field enhancement factor at the cathode, accounting for changes in the coefficient γ of eqn. (11), and in F_c/F during the transit of the charge cluster and for possible recombination processes.

A_{v1} and t_{t1} are the effective cross section and the transit time for the positive charge cluster.

Denoting by F_l the lowest cathode field which causes instability, one avalanche instability occurs for average applied fields

$$F \geqslant F_l/\beta_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

after the mean time to instability

$$\tau_s = e/j_c(F)A_s. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

The lowest applied field for one avalanche instability is denoted by $F_1 = F_l/\beta_1$.

At applied fields lower than F_1 the rate of breakdowns decreases with field rapidly below the rate of positive charge clusters produced by the first injected electrons. Instability can arise as a two-avalanche process; the cluster must be hit by a second chance electron (fig. 9) which, by avalanching builds up F_c above F_l, leading to current runaway and breakdown. At still lower fields, in a three-avalanche process, the cluster must be hit by a second and a third chance electron to obtain $F_c > F_l$ and breakdown, and in an n-avalanche process the avalanches of n successively injected electrons build F_c up locally above F_l for breakdown.

In the continuation, the notation of eqns. (12)–(15) will be used, but for the pth avalanche the index p will replace the index 1. The two-avalanche process extends over the range of fields

$$F_l/\beta_2 \leqslant F \leqslant F_l/\beta_1, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (16)$$

when β_2, the field enhancement factor after the second avalanche, is larger than β_1.

The smallest applied field in this range $F_2 = F/\beta_2$ is determined by

$$j_c(\beta_2 F)A_{v2}t_{t2}/e = 1. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

The probability P that a destructive event is initiated is the probability P_2 that a second electron hits the cluster of the first avalanche and

$$P_2 = j_c(\beta_1 F)A_{v1}t_{t1}/e \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (18)$$

and with eqn. (12) the average time to initiate breakdown per unit area of specimen

$$\tau_s = e^2/j_c(F)A_s j_c(\beta_1 F)A_{v1}t_{t1}. \quad . \quad . \quad . \quad . \quad . \quad (19)$$

A succession of n avalanches produces instability over the range of fields

$$F_l/\beta_n \leqslant F \leqslant F_l/\beta_{n-1} \quad \text{with} \quad \beta_n > \beta_{n-1}. \quad . \quad . \quad . \quad . \quad (20)$$

The smallest applied field in this range $F_n = F_l/\beta_n$ is determined by

$$j_c(\beta_n F)A_{vn}t_{tn}/e = 1. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (21)$$

The probability P that destructive breakdown follows is the probability that $n-1$ electrons hit the cluster successively. The probability that a $(p+1)$th electron hits the cluster is $j_c(\beta_p F)A_{vp}t_{tp}/e$, hence

$$P=\prod_{p=1}^{p=n-1} j_c(\beta_p F)A_{vp}t_{tp}/e \quad \text{with} \quad P_n<P_{n-1} \qquad (22)$$

and the mean time to initiate breakdown per unit area of specimen

$$\tau_s=e^n/j_c(F)A_s\prod_{p=1}^{p=n-1} j_c(\beta_p F)A_{vp}t_{tp}. \qquad (23)$$

It is of interest to note that the part of insulator volume occupied by avalanche charge clusters $f=j_c(F)A_{v1}t_{t1}/e$ is smaller than any probability that a $(p+1)$th electron hits the cluster $j_c(\beta_p F)A_{vp}t_{tp}/e$. Thus on multi avalanche instabilities only a fraction of the insulator is occupied by charge clusters, as was assumed earlier in § 3.

With increase in the number of avalanches required for instability, the probabilities P_p decrease and the mean times to instability τ_s increase rapidly. A low field limit for breakdown F_{mn} may be stipulated arbitrarily for a very high value of τ_s. At F_{mn} the rates of electron injection or impact ionization or both may become very low and the magnitude of F_{mn} depends on the thickness of the insulation. In an experiment with decreasing fields the development of full size avalanches of length a_v stops, when $a_v>w$, the insulator thickness. This, according to eqn. (5), arises at fields $F<F_a$, where

$$F_a=\kappa i V_i/w. \qquad (24)$$

The positive charge cluster, the cathode field enhancement β_1 and P drop then considerably and for a small field decrease determined by fluctuations. τ_s may increase to such an extent as to reach the low field limit. The attainment of the low field limit is enhanced, when the relation for the ionization coefficient changes from eqn. (4) to (3), i.e. $\alpha\propto\exp(-H/F)$.

The determination of specific relations for $\tau_s=f(F)$ with eqns. (15), (19) and (23) is difficult, because not only j_c, but also β_p, A_{vp} and t_{tp} depend on F in an involved manner. Considerable approximations are needed to obtain tractable relations for τ_s. Let us consider the parameters β_p, A_{vp} and t_{tp} first.

The first avalanche cross section from eqn. (6) is $A_{v1}=4\pi\kappa i l_p V_i/3F$ and the effective transit time t_{t1} of the charge cluster from eqn. (5) is $a_{v1}/\mu_p\beta_1 F=\kappa i V_i/\mu_p\beta_1 F^2$. Here μ_p is the mobility of the positive charges. On repeated avalanching both A_{vp} and t_{vp} decrease due to the increase in the field enhancement factors β_p. At the same time A_{vp} increases both by diffusion and because of the scatter in the location where the injected electron hits the charge cluster. These opposing effects influencing the product $A_{vp}t_{tp}$ will be neglected. This may be justified, especially when

Table 2. Breakdown relations for carrier injection by thermionic or field emission at the cathode

	Thermionic emission	Field emission
Lowest field for an n-avalanche process	$F_n = \left[\frac{kT}{b\beta_n^{1/2}} \ln \frac{\beta_1 e\mu_p F_n^3 \exp(E/kT)}{AKT^2}\right]^2$	$F_n = \frac{D}{\beta_n \ln (CK\beta_n^2/e\mu_p\beta_1 F_n)}$
Mean time to instability:		
One-avalanche process	$\tau_s = \frac{\exp(E/kT)}{A_s AT^2/e} \exp(-bF^{1/2}/kT)$	$\tau_s = \frac{e}{A_s CF^2} \exp(D/F)$
Two-avalanche process	$= \frac{\mu_p\beta_1 F^3}{A_s K}\left[\frac{\exp(E/kT)}{AT^2/e}\right]^2 \exp - \left[\frac{bF^{1/2}}{kT}(1+\beta_1^{1/2})\right]$	$= \frac{e^2\mu_p}{A_s KC^2\beta_1 F} \exp\left[\frac{D}{F}\left(1+\frac{1}{\beta_1}\right)\right]$
n-avalanche process	$= \frac{1}{A_s}\left(\frac{\mu_p\beta_1 F^3}{K}\right)^{n-1}\left[\frac{\exp(E/kT)}{AT^2/e}\right]^n \times \exp - \left[\frac{bF^{1/2}}{kT}(1+\beta_1^{1/2}+\beta_2^{1/2}+\ldots+\beta_{n-1}^{1/2})\right]$	$= \frac{1}{A_s}\left(\frac{e\mu_p\beta_1 F}{KC}\right)^{n-1} \times \frac{\exp\left[\frac{D}{F}\left(1+\frac{1}{\beta_1}+\frac{1}{\beta_2}+\ldots+\frac{1}{\beta_{n-1}}\right)\right]}{CF^2(\beta_1\beta_2\ldots\beta_{n-1})^2/e}$
n-avalanche process	with μ_p given by eqn. (26) $\tau_s = \frac{1}{A_s}\left(\frac{\mu_{po}\beta_1 F^3}{K}\right)^{n-1}\left(\frac{\exp[(E-E_\mu)/kT]}{AT^2/e}\right)^{n-1} \times \frac{\exp(E/kT)}{AT^2/e} \times \exp - \left\{\frac{bF^{1/2}}{kT}[1+(\beta_1^{1/2}+\beta_2^{1/2}+\ldots+\beta_{n-1}^{1/2})(1-S)]\right\}$ (38)	with μ_p given by eqn. (27) $\tau_s = \frac{1}{A_s}\left(\frac{e\mu_{po}\beta_1 F e^{-\frac{E_\mu}{kT}}}{KC}\right)^{n-1} \times \frac{\exp\left\{\frac{D}{F}\left[1+(1-R)\left(\frac{1}{\beta_1}+\frac{1}{\beta_2}+\ldots+\frac{1}{\beta_{n-1}}\right)\right]\right\}}{CF^2(\beta_1\beta_2\ldots\beta_{n-1})^2/e}$ (39)

μ_p is strongly dependent on field and temperature. The expression $A_{vp}t_{tp}$ is then approximately

$$A_{vp}t_{tp}=\frac{4\pi l_p(\kappa i V_i)^2}{3\beta_1\mu_p F^3}=\frac{K}{\beta_1\mu_p F^3} \quad \text{with} \quad K=4\pi l_p(\kappa i V_i)^2/3. \quad . \quad (25)$$

Replacing from eqn. (25) for $A_{vp}t_{tp}$ and from eqns. (1) and (2) for j_c into eqns. (13)–(23) breakdown relations were derived for the two cases of carrier injection at the cathode by thermionic or by field emission. The resulting relations for τ_s and the lowest field F_n for an n avalanche instability process are tabulated in table 2.

The mobility is constant in the derived relations. If, however, it is assumed that the positive charges are removed from the insulator by electronic hopping processes, μ_p may be some exponential function of both field and temperature (Jonscher 1967, 1971; Hill 1971). To assess the effect of μ_p on τ_s approximate functions for μ_p were chosen to reduce the mathematical complexity in the relations for τ_s. When thermionic injection prevails at the cathode μ_p is approximately given by

$$\mu_p=\mu_{p0}\exp[-(E_\mu-bSF^{1/2})/kT], \quad . \quad . \quad . \quad . \quad (26)$$

where E_μ is an activation energy and b and S are constants. In the case of field emission at the cathode

$$\mu_p=\mu_{p0}\exp\left(-\frac{E_\mu}{kT}-\frac{DR}{F}\right) \quad . \quad . \quad . \quad . \quad . \quad (27)$$

with R a constant.

The last two equations in table 2 (eqns. (38), (39)) were obtained by replacing μ_p with its appropriate, equivalent (eqns. (26) or (27)). These relationships show that the influence of μ_p on τ_s is expressed by the decrease in the absolute values of the slopes of the breakdown characteristics $\log\tau_s=f(F^{1/2})$ or $f(1/F)$ by the factors $(1-S)$ and $(1-R)$, respectively. The energy barriers for hopping (Jonscher 1967, 1971, Hill 1971) are usually smaller than those for electron injection into the insulator so $(1-S)$ and $(1-R)$ may often be not very different from unity.

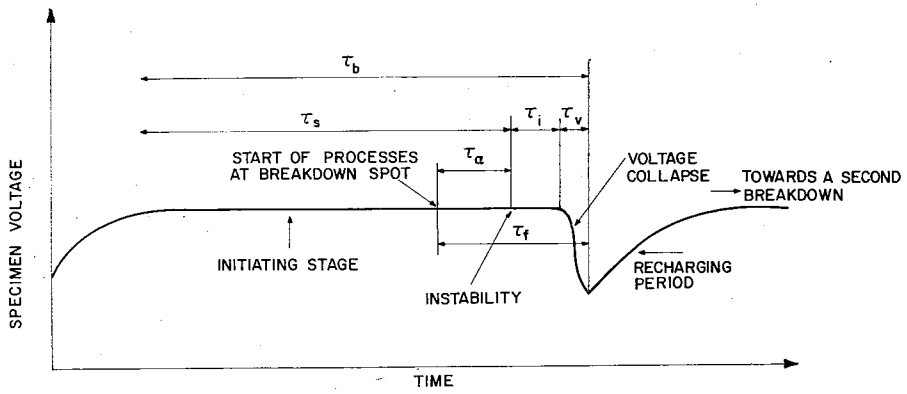
According to the results in table 2, the slopes of the $\tau_s=f(F)$ relations are discontinuous at the boundaries of breakdown ranges where the instability process changes from $p-1$ to p successive avalanches. The discontinuities are connected with the neglect of statistical fluctuations in the sequence of processes composing the breakdown event. Such fluctuations accompany the event from the injection of the first electron until the current runaway process. If the fluctuations were accounted for, the transition in slopes between adjacent ranges would be continuous.

4.6. *The Completion of the Breakdown Event*

Up to this point in the analysis the instability in the localized breakdown model has been ascribed solely to electronic effects. The influence of temperature rises has been disregarded. In reality the temperature rise

during the avalanching contributes considerably to the growth of the electrical conductivity and to the current runaway (Klein and Burstein 1969), which causes the voltage collapse. Impact ionization may stop during the voltage collapse, but the breakdown event continues to destruction because due to temperature rise the breakdown spot becomes thermally unstable. On non shorting breakdowns the event continues only until the voltage decreases to a minimum value. The conducting filament, which may then be an arc between the electrodes, extinguishes and the specimen recharges as illustrated in fig. 10 (Klein 1966).

Fig. 10



Sketch of specimen voltage versus time, with one breakdown event. τ_b, τ_s and τ_f are indicated with τ_b and τ_s in this case assumed to begin approximately when the specimen voltage has reached full value.

The breakdown event described here involves only a very small volume of the insulator film. In such cases, the d.c. V–I characteristic of the specimen is discontinuous (Berglund and Klein 1971), as illustrated by fig. 6. Prior to a shorting breakdown, the current in the bulk of the insulator is represented by the high voltage–low current branch of the characteristic. After the breakdown a conductive filament is in parallel with the bulk of the insulator and the low voltage–high current branch of the characteristic represents mainly the current in the filament. When the breakdown is non-shorting, a conductive filament is only temporary until its vaporization and after the event the specimen is again represented by the high voltage–low current branch.

The temporal development of the breakdown event is illustrated in fig. 10, where the specimen voltage is plotted versus time and where the durations of the breakdown stages are marked. We will discuss the durations of the stages for the whole specimen and also for the breakdown spot.

The time to breakdown τ_b of the specimen, the period from the application of the full voltage to the completion of the event with voltage

collapse, comprises the durations of the initiating, instability and voltage collapse stages (fig. 10), τ_s, τ_i and τ_v, respectively, and

$$\tau_b = \tau_s + \tau_i + \tau_v. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (28)$$

During the initiating stage τ_s many harmless avalanches occur in the whole specimen until at one spot the chance event of a succession of avalanches leads to an instability. The period τ_s given by eqns. (15), (19) and (23) is a statistical time lag. The event continues at the breakdown spot with current runaway. After τ_s, the runaway comprises two periods: In the first τ_i the voltage change is insignificant; in the second τ_v the voltage collapses. The differentiation into two periods may seem arbitrary at first sight. However, the processes in each period are different. During τ_i the runaway is electronic with avalanching, but the runaway becomes thermal during τ_v. The differences will become clearer through the calculation of τ_i and τ_v.

Avalanching continues during the period τ_i, and we will assume that the avalanches are due to the successive injection of u electrons after the n electrons, which produced instability. The $n+1$th electron is injected on the average after a time $t_{tn} = e/j_c(\beta_n F)A_{vn}$, given by eqn. (21). Similar expressions apply for electrons injected successively and

$$\tau_i = \sum_{q=n}^{q=n+u-1} \frac{e}{j_c(\beta_q F)A_{vq}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (29)$$

Successive electrons are supposed to enhance the cathode field rapidly, so that $j_c(\beta_n F) \ll j_c(\beta_{n+1} F) \ll j_c(\beta_{n+2} F) \ll \ldots$, and approximately

$$\tau_i \simeq e/j_c(\beta_n F)A_{vn}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (29\,a)$$

τ_i is thus determined mainly by the interval which begins when the cathode field rises locally to F_l and ends when the next electron hits the charge cluster.

The period τ_v is determined by the duration of the discharge of the electrostatic energy stored in the specimen through the breakdown spot and is approximately given by

$$\tau_v = C_s R_d. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (30)$$

Here C_s is the specimen capacitance and R_d the effective resistance of the discharge path. Values for τ_v varying from nano to milliseconds have been observed.

At the breakdown spot itself the event requires a formative time τ_f (fig. 10), which consists of three periods:

$$\tau_f = \tau_a + \tau_i + \tau_v. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (31)$$

τ_a is the period which begins with the injection of a first electron and ends when the succession of n injected chance electrons has raised the cathode field to F_l for instability. The period τ_a overlaps with part of the statistical time lag τ_s. Since the positive charge cluster at the breakdown

spot must be hit by an injected electron before the charge leaves the insulator, an upper limit for τ_a is approximately the sum of $n-1$ transit times and

$$\tau_a = \sum_{p=1}^{p=n-1} a_{vp}/\mu_p \beta_p F. \quad \ldots\ldots\ldots \quad (32)$$

When the large majority of the positive charge clusters leave the insulating film without developing into breakdown, $\tau_a \ll \tau_s$.

At low and medium fields for breakdown, when $\tau_s \gg \tau_f$ or $\tau_s \simeq \tau_b$, individual breakdown events usually form at separate times. The breakdown rate measured on specimens with non shorting breakdowns is then $R_b = 1/\tau_s \simeq 1/\tau_b$, because breakdown arises at every spot which becomes unstable. At high fields for breakdown, when τ_s and τ_f can be of the same order of magnitude, or $\tau_s < \tau_f$, breakdown events develop concurrently. Ideally the breakdown rate should still be $1/\tau_s$. Voltage collapse, however, requires recharging periods between subsequent breakdowns (fig. 10). Space charge clusters forming breakdowns may dissipate during these periods; in consequence R_b diminishes and so does the significance of measurements at high fields.

§ 5. Discussion

5.1. *The Mechanism of Breakdown and its Properties*

A complex probabilistic mechanism has been proposed here for breakdown. The incidence of breakdown depends upon the succession of strongly field dependent chance events and upon the interaction of a number of processes, each of which can have a varied nature. The processes important for the initiation of breakdown are the increase of the space charge near the cathode and the increase in the electron injection rate into the conduction band of the insulator. These processes promote the development of breakdown, but there are several others which have an opposing effect, such as recombination and the expansion of charge clusters by diffusion and local field during the drift towards the cathode. These dispersive effects cause a decrease in the cathode field expressed as a decrease in the field enhancement factors β_p.

A further effect opposing the development of breakdown events is the decrease of the transit time t_{tp} of the charge clusters of successive avalanches drifting in the increasing field. This is pronounced since the mobility at the positive charges μ_p increases with field and temperature. Owing to the decrease of t_{tp} the chance of an injected electron hitting the charge cluster decreases, and so does the probability for the continuation of the breakdown process. This effect is expressed mainly by the factors $(1-R)$ and $(1-S)$ in the relations of table 2.

The mechanism for breakdown which has been developed here is based upon electron injection. In principle an analogous model can be elaborated also for hole injection.

It is interesting to note that there are some similarities between our model for the breakdown of thin insulators and observations on gas breakdown (Raether 1964). Successions of avalanches have been observed when gas breakdown is started by a single electron avalanche. In gases, as in insulators, the positive charge mobility is much smaller than that of electrons. The Townsend condition for gas breakdown with self-sustained avalanching is the production of an additional avalanche by an electron during the drift of the positive charges towards the cathode. This, according to eqn. (13), is also the condition for thin film breakdown.

Another point of interest arises on inspection of the breakdown relations in table 2. The relations show that the main parameters determining the time to instability τ_s are the energy barrier E controlling injection, the mobility μ_p and the field enhancement factor β_p determined by charge accumulation at the cathode. In contrast, the factors expressing impact ionization explicitly, i and V_i, appear through the coefficient K as pre-exponential factors and seem to have only a minor influence on τ_s. It appears that the formation of positive space charge is essential for the development of breakdown, but the process by which it is generated may be of lesser importance and may not be restricted to avalanching. Processes, such as emptying of traps near the cathode by the field, or trapping of injected positive charges at the cathode and of negative charges at the anode (Mott 1969, Schmidlin 1970) may also generate the space charges. Such processes would offer interpretations alternative to avalanching for the initiation of breakdown, if they could be shown to lead to randomly occurring localized breakdowns.

The discussion of theory will be continued on the assumption that breakdown is initiated by a succession of avalanches and we proceed to consider the individual breakdown properties. The first of these is the random occurrence of breakdown events which are initiated in a very small volume. The initiation in a very small volume is supported by the observation of very small volumes of breakdown damage in mica and silicon dioxide mentioned earlier (Klein 1966, Davidson and Yoffe 1968). The random occurrence of breakdown implies that, on a test of long duration at constant voltage, a Poisson distribution will be found for the breakdown events in space and in time around a mean time to instability τ_s. This was observed in detail on a hafnium dioxide specimen (figs. 3 and 4). We will return to the effect of space charges on such tests later.

Other properties to be discussed will be the dependence of breakdown on the field, the electrodes, space charges, the temperature, and the thickness and the results of theory will be compared with the thin film observations. Difficulties will be encountered in the calculations of breakdown because these require data on carrier injection and multiplication, on the creation of space charges and on their drift. Most of the data are not known and the comparisons will be mainly semi-quantitative in nature.

5.2. *The Influence of Field*

According to table 2, the time to instability τ_s is an exponential function of $F^{1/2}$, when the injection is thermionic and an exponential function of $1/F$ for injection by field emission. Presenting breakdown characteristics as semilogarithmic plots of τ_s either versus $F^{1/2}$ or $1/F$, whichever applicable, an increase in the absolute value of the slope of the characteristics is predicted as the field decreases and hence the number of successive avalanches required for instability increases. The change in slope is especially marked in the vicinity of the low field limit for breakdown, where τ_s tends to infinity. As mentioned earlier, a low field limit is indicated by the data in several insulators (figs. 1 (*a*), (*b*)).

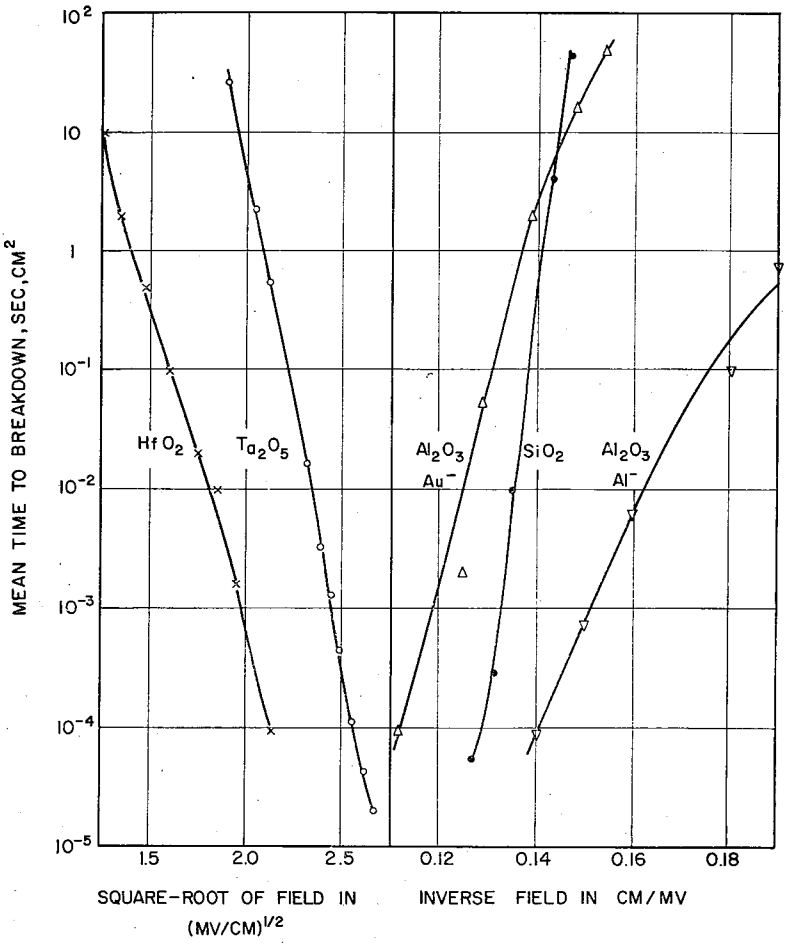
The relations of table 2 show that when breakdown is initiated by electron injection from the cathode, the absolute values of the slope of the breakdown characteristics ought to be larger than that of the semilogarithmic characteristics of injected current versus $1/F$ or $F^{1/2}$. We will investigate whether this prediction is confirmed by the observed thin oxide breakdown characteristics.

To this end the nature of the carrier injection process must be determined first. Assuming that the injection is directly from the cathode, calculations with eqns. (1) and (2) can indicate the nature of the injection process. For example, calculations with $m^*/m = 1$, $\epsilon = 6$ and E ranging from 0·9 to 1·2 ev show that equality of thermionic and field emission currents is obtained at fields ranging from 2·25 to 2·4 Mv/cm. Thermionic emission dominates for lower energy barriers and field emission for higher barriers. It is often found that current injection from the cathode is not described by the extreme cases of eqns. (1) and (2), but by the combined effect of thermionic-field emission. Furthermore the temperature increases during the development of a breakdown event and the nature of injection may change.

The nature of the injection process has been ascertained only for vapour grown silicon dioxide. There it was found to be field emission (Lenzlinger and Snow 1969). Values for the barrier energies E at aluminium oxide also indicate field emission as the primary cause of injection, though possibly enhanced by thermal effects. The value of E is not known for the oxides of hafnium and tantalum. As τ_s rapidly decreases with increasing temperature (fig. 2) thermionic emission will be assumed as the dominant injection mechanism for the latter two oxides.

The breakdown characteristics of fig. 1 (*b*) were replotted to fit the assumed injection mechanism. For the oxides of tantalum and hafnium in fig. 11 the mean time to breakdown was plotted semilogarithmically versus the square root of field and for the oxides of aluminium and silicon versus the inverse field. The slopes $|G_t|$ and $|G_f|$ of these curves for the thermionic and field emission injection respectively were determined at high breakdown rates, where the curves are approximately linear; the results are shown in table 3.

Fig. 11



(*a*) (*b*)

Mean time to breakdown: (*a*) versus square root of field in HfO_2 and Ta_2O_5; (*b*) versus inverse field SiO_2 and Al_2O_5. Data of fig. 1 (*b*).

Table 3. The slopes of the breakdown characteristics of fig. 11

Specimen	HfO_2	Ta_2O_5	Al_2O_3 (Au electr. neg.)	Al_2O_3 (Al electr. neg.)	SiO_2
Slopes $\lvert G_t\rvert$, $\lvert G_f\rvert$	$1{\cdot}5\times10^{-2}$	$2{\cdot}0\times10^{-2}$	$3{\cdot}7\times10^{8}$	$2{\cdot}1\times10^{8}$	$7{\cdot}8\times10^{8}$
	in $cm^{1/2}/v^{1/2}$		in v/cm		
Factors L_t, L_f	2·1	3·1			3·2

For the oxides of hafnium and tantalum the slope from eqn. (38) in table 2 is approximately:

$$|G_t| = \frac{b}{kT}[1+(\beta_1^{1/2}+\beta_2^{1/2}+\ldots+\beta_{n-1}^{1/2})(1-S)] = \frac{b}{kT}L_t, \quad . \quad (33)$$

where L_t is the slope enhancement factor, when injection is by thermionic emission. The magnitude of L_t permits to estimate roughly the number of successive avalanches causing instability, if it is assumed that the values for β_p vary roughly between 1·2 and 1·5 and $(1-S)$ is somewhat smaller than unity. At room temperature, with eqn. (1 *a*).

$$b/kT = 1{\cdot}48\times10^{-2}/\epsilon^{1/2}\ \mathrm{cm}^{1/2}\ \mathrm{v}^{-1/2},$$

where ϵ is a high frequency permittivity of the insulator (Jonscher 1967). Using for $\epsilon^{1/2}$ the value of the refractive index at wavelengths close to 4000 Å, $\epsilon^{1/2}=2{\cdot}05$ for HfO_2 (Sviridova and Svikovskaya 1967), and 2·26 for Ta_2O_5 (Young 1961), one obtains from eqn. (33), $L_t=2{\cdot}1$ for HfO_2 and 3·1 for Ta_2O_5. The values of L_t indicate that in the field range considered, instability occurred in hafnium dioxide after a succession of two or three avalanches, and in tantalum pentoxide after there or four successive avalanches.

Breakdown in the silicon dioxide specimen can be discussed in more detail, because E, m^*/m and $j_c=f(F)$ is known (Lenzlinger and Snow 1969). This permits to present in fig. 12, $R_b \simeq 1/\tau_b$, the probability for destructive breakdown P and j_c versus the inverse field. The R_b curve was replotted from fig. 1 (*b*) and the probability P was calculated with eqn. (12), $P=eR_b/j_cA_s$. The very small values of P, increasing from 10^{-13} to 10^{-9}, indicate that the breakdown observations are the outcome of several successive avalanches prior to instability.

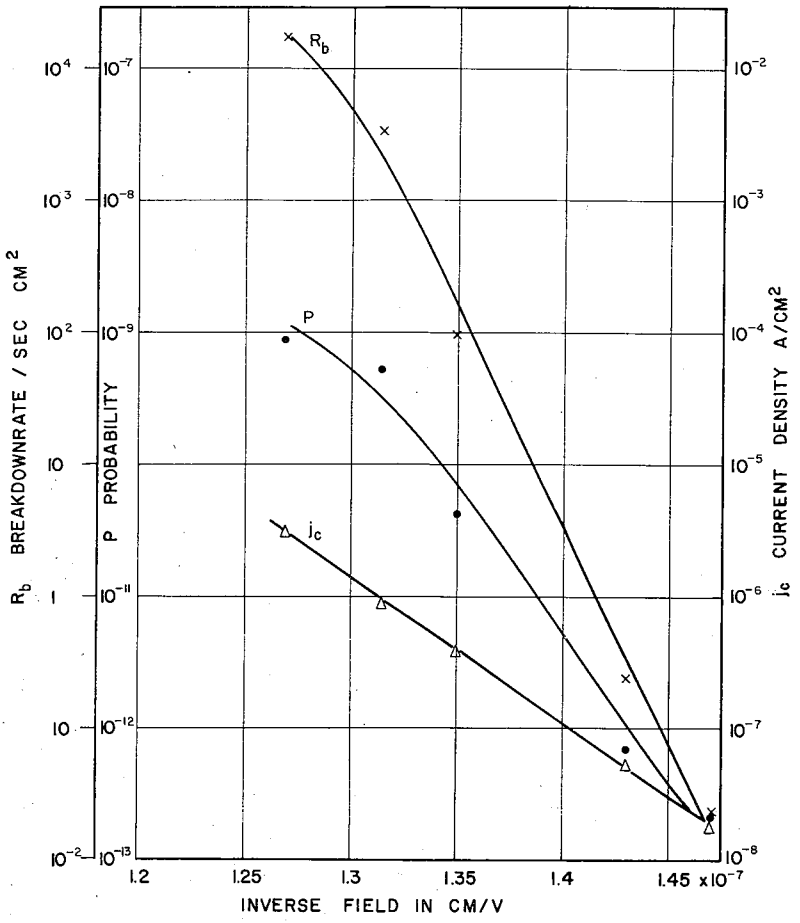
The approximate slope of the breakdown characteristic of SiO_2 is given by the field emission relation eqn. 39 in table 2 and

$$|G_f| = D\left[1+(1-R)\left(\frac{1}{\beta_1}+\frac{1}{\beta_2}+\ldots+\frac{1}{\beta_{n-1}}\right)\right] = DL_f \quad . \quad (34)$$

with L_f the slope enhancement factor, when injection is by field emission. L_f permits the estimate of the number of successive avalanches to instability in the same manner as was found with L_t. Calculating with eqn. (2 *a*), with $E=3{\cdot}2$ ev at the aluminium oxide interface and $m^*/m=0{\cdot}39$, one obtains $D=2{\cdot}4\times10^8$ v/cm and $L_f=3{\cdot}2$ in the linear range of the breakdown characteristic. This result again indicates that instabilities were due to a succession of several avalanches.

Caution should be observed in the interpretation of the slope G_f of silicon dioxide with eqn. (34). In silicon dioxide the range of fields over which breakdowns are observed is much narrower than in other insulators. As carrier injection is already significant well below the breakdown range, the event is limited by bulk processes rather than carrier injection.

Fig. 12



Injected current density, probability of destructive breakdown and breakdown-rate versus inverse field for the SiO_2 specimen of fig. 1 (*b*).

Accordingly the data for SiO_2 appear to represent mainly the 'low field' end of the breakdown characteristic. If the bulk process is impact ionization, its coefficient, α is very small and given by an exponential relation such as eqn. (3). The avalanche sizes N_e are very much smaller than the full size avalanche N given by eqn. (8) and breakdowns are likely to develop only on the rare chance event of a full size avalanche. The probability that this will occur is $\exp(-N/N_e)$ (Wijsman 1949), and this factor multiplies the expressions derived earlier for P and τ_s. This change in the relationship makes τ_s much more dependent on avalanching and correspondingly less dependent on carrier injection.

It may be queried at this point whether the silicon dioxide results can be explained, assuming that the positive charges are completely immobile.

In such a case every charge cluster continues to be hit by electrons producing avalanches at an increasing rate. Instability arises then by the one-avalanche process with every injected electron causing breakdowns at a rate $R_b = j_c/e$. This is not the case in SiO_2, since by fig. 12 R_b is 10^9 to 10^{13} times smaller than j_c/e and the instabilities were interpreted as multi-avalanche processes. A finite mobility has therefore to be assigned to the positive charge carriers. If based on the magnitude of L_t it is assumed that instability arises by a three-to-four-avalanche process, then the probabilities P_p that an electron hits a cluster during its transit can be roughly estimated. The results of fig. 12 for the breakdown probability indicate that the P_p values fall between 10^{-5} to 10^{-2}. This shows that at any time only a very small fraction of the insulator is occupied by positive charge clusters. The magnitudes of P_p make it possible to estimate roughly with eqn. (22) the transit times and the mobilities of the positive charges. The mobilities are found to fall somewhere between the values of 10^{-9} to 10^{-5} cm^2/v, sec, plausible for hopping processes. $R_b \ll j_c/e$ occurs also in the other oxides in which the slope enhancement factor $L_t > 1$ and only a fraction of the insulator is occupied by charge clusters.

According to eqns. (16) and (20), the number of successive avalanches for instability may be constant in field ranges in which the ratio of the highest to lowest breakdown is as large as 3 : 2. Breakdown observations were made over much larger ranges, the ratio being for instance nearly 4 : 1 in hafnium dioxide. In such cases, as mentioned earlier, the absolute value of the slope of the breakdown characteristics is expected to increase with decreasing field. Such an increase in slope is discernible for several alkali halides and for mica in fig. 1 (*a*). The increase is rapid in the vicinity of lowest fields for breakdown. A decrease in slope with decreasing field was found too in part of the breakdown characteristics, as shown for Al_2O_3 in fig. 11. The reason for the behaviour in Al_2O_3 is not clear. It may be connected with space charge effects.

Summing up this subsection, it was found that the high field slopes of the breakdown characteristics of the oxides of hafnium, tantalum and silicon were, as expected by theory, larger than the slopes of the current characteristics, indicating that only a few successive avalanches are needed for instability.

5.3. *The Influence of Electrode Material*

When breakdown is initiated by electrons injected from the cathode by field or thermionic emission (table 2), the breakdown field for a given value of τ_s and the slope of the breakdown characteristic $|G|$ are mostly an increasing function of the interface barrier E which depends on the electrode material. In contrast, when the initiating electrons reach the conduction band from the cathode by hopping processes or carrier excitation due to recombination light, the breakdown field and the slope of the characteristic should show little or no dependence on the interface barrier and electrode material.

Consider the electrode dependence of the aluminium oxide measurements (fig. 1 (*b*)). The interface barriers can be evaluated with the relation $E = W - \chi$, W being the work function of the electrode and χ the electron affinity of the oxide. χ was found 1·58 ev (Pollack and Morris 1964) from tunnelling experiments through very thin films. With $W = 4{\cdot}8$ and 3·8 ev for gold and aluminium respectively, the interface barriers are $E_{\mathrm{Au}} \simeq 3{\cdot}2$ ev and $E_{\mathrm{Al}} \simeq 2{\cdot}2$ ev. These values of E indicate that current injection is mainly by field emission and offer a qualitative explanation for observations of the larger breakdown fields and the larger absolute values of the slope of breakdown characteristics when the gold electrode was negative.

According to table 2, assuming equal slope enhancement factors for the two polarities, the ratio of the slopes of the characteristics $G_{\mathrm{f,Au-}}/G_{\mathrm{f,Al-}}$, equals $D_{\mathrm{Au-}}/D_{\mathrm{Al-}}$. It follows then from eqn. (2 *a*) that

$$G_{\mathrm{f,Au-}}/G_{\mathrm{f,Al-}} = [(W_{\mathrm{Au}} - \chi)/(W_{\mathrm{Al}} - \chi)]^{3/2}$$

and calculating with the values of W and χ just quoted:

$$G_{\mathrm{f,Au-}}/G_{\mathrm{f,Al-}} = 1{\cdot}75.$$

This value is in good agreement with the ratio of the observed slopes, 1·76 according to table 3, supporting predictions of the theory on the influence of electrode material. These calculations yielded also the value of the product $(m^*/m)^{1/2} L_{\mathrm{f}} = 0{\cdot}94$, indicating that $m^* m \leqslant 1$. Unfortunately data found on m^*/m in the literature vary to such an extent that L_{f} could not be determined (Hurych 1970).

The influence of electrode material can be examined also by comparing τ_{s} values for opposite polarities on the Al_2O_3 specimen. For an applied field of $F = 6{\cdot}8$ Mv/cm the measured ratio $\tau_{\mathrm{s,Au-}}/\tau_{\mathrm{s,Al-}} = 4 \times 10^4$ (fig. 11). Calculating, this ratio is found to be $1{\cdot}7 \times 10^{10}$, if it is assumed that the L_{f} factors are independent of polarity and

$$\tau_{\mathrm{s,Au-}}/\tau_{\mathrm{s,Al-}} = \exp\,[(G_{\mathrm{f,Au-}} - G_{\mathrm{f,Al-}})/F]$$

according to table 2. This is a very large discrepancy, but may be caused by space charges in the insulator.

5.4. *Space Charges*

Until now we assumed an ideal situation where the only space charges affecting the event are those created by avalanching. In reality, breakdown observations are influenced considerably by space charges forming in the insulator in parallel with the breakdown clusters. These charges may be either ions or trapped electrons. The importance of the former has been pointed out already by von Hippel and Alger (1949), who assumed that the decrease in breakdown fields with time in alkali halides can be explained by positive ion drift towards the cathode.

The charges cause field distortions which may either enhance or decrease the breakdown rate. Such effects may explain why the measured ratio

$\tau_{s,Au^-}/\tau_{s,Al^-}$ was 4×10^5 times smaller in aluminium oxide than the calculated one. This decrease in the ratio may arise in the given case when the applied field of 6·8 Mv/cm at the gold cathode increases by a positive space charge in the insulator to 8·9 Mv/cm, or when the field at the aluminium cathode decreases to 4·9 Mv/cm due to a negative space charge. The occurrence of space charges in aluminium oxide has been observed on various occasions (Antula 1968, 1969, Davies and Collins 1969, Miles and Rawlings 1970).

The additional space charges further complicate the problem of how to obtain significant results on testing for beakdown. Let us consider, for instance, the test for breakdown rate R_b at constant voltage as illustrated for hafnium oxide in fig. 3. In this case R_b remains constant for more than 1000 breakdowns. However, in other insulators, for instance in aluminium oxide, changes occur in R_b on such a test (Albert 1972). The changes may be due to space charge variations. They may be caused also by completely different effects connected with the basic mechanism of breakdown, such as time dependent changes in material properties due to the high field. The constant voltage test can thus offer significant evidence on the breakdown mechanism, and it is important to verify whether changes in R_b are caused by space charges. Information on changes in space charges may be obtained by methods such as measurements of C–V characteristics, optical and thermal excitations and also by measurements of R_b with polarity changes.

Another experimental situation, where space charges are important is when intervals occur between voltage applications for a group of breakdowns. As the result of an earlier voltage pulse, positive charge clusters remain in the insulator. These are slow to drift to the electrodes when the external field has been removed. When a new pulse is applied, the positive charge clusters from the earlier pulse act as preferential sites for breakdowns as if new ' weak spots ' were created (Solomon 1972). The effect of these clusters is even more interesting when pulses of opposite polarity are applied. In such experiments not only temporary weakening, but also strengthening of the film may be observed (Albert 1972). Such experiments can shed additional light on the mechanism of breakdown.

5.5. *Temperature*

The temperature affects nearly every process in the development of the breakdown event. Some temperature effects promote and others oppose the development of breakdown.

Considering the relations in table 2 for the effect of the injection current on the time to breakdown instability, τ_s is found to decrease exponentially with the inverse temperature when injection at the cathode is by thermionic emission. The decrease in τ_s is less rapid with increasing temperature when injection is by thermionic-field emission and relatively small when injection is by pure field emission.

The influence of the increase in the mobility μ_p with increasing temperature, as given by eqns. (26) and (27), is to decrease the transit time and hence to increase τ_s. In analogy to the activation energies for hopping currents, the activation energy E_μ for μ_p may be assumed to vary from very small values to roughly 1 ev (Jonscher 1967, 1971). This magnitude of E_μ is smaller than usual barrier heights at the electrode-insulator interface, hence the temperature influence of the injection current predominates, except in the case of field emission.

The involved nature of the temperature dependence is revealed by table 2. It is found for the case of thermionic emission for instance, that when instability requires two successive avalanches

$$\tau_s \infty \exp\{(2E - E_\mu - bF^{1/2}[1 + \beta_1^{1/2}(1 - S)])/kT\} \equiv \exp(E_a/kT).$$

Experimental results giving the slope of $\log \tau_s = f(1/T)$ curve cannot be interpreted with this relation because most of the parameters in the brackets are not known. It may be of interest, however, to quote values found for E_a : A τ_s versus $1/T$ plot for the HfO_2 specimen of fig. 2 shows a small E_a below 100°K. The activation energy increases with temperature and becomes 1 ev around 350°K at low fields for breakdown. E_a for the Ta_2O_5 specimen of fig. 2 is 0·45 ev in the temperature range of 200–350°K at fields 5 to 6 Mv/cm.

The temperature dependence of τ_s is weaker in Al_2O_3 (Klein 1969 a), and τ_s and the breakdown field seem to be insensitive to temperature changes in both vapour grown silicon oxide (Klein and Burstein 1969) and thermally grown silicon dioxide (Chou and Eldridge 1970). An increase in breakdown field with temperature was observed in thick alkali halides (Franz 1956, Stratton 1961, O'Dwyer 1964) and in silicon p–n junctions (Lee *et al.* 1964, Crowell and Sze 1966). In both cases the increase was ascribed to the decrease of the ionization coefficient with increasing temperature. Thus the temperature behaviour of the electronic breakdown events is quite variable and is determined by how such parameters as the injection current, the hopping mobility and the ionization coefficient change with temperature.

5.6. *The Influence of Thickness*

The thickness dependence of breakdown does not appear explicitly in the relations of table 2. To investigate this dependence, the range of thickness to be considered first is $w > a_v$, where a_v is the length of a full size avalanche given by eqn. (5).

According to eqns. (5)–(8) and (25), the avalanche data a_v, A_v and N are independent of thickness. The average cathode field during the transit of the charge cluster, however, increases with w according to eqns. (10) and (11), increasing the average value of the factor β_1. The relations of table 2 show that, for a given value of τ_s, a decrease of the breakdown field corresponding to the increase in β_1 is expected with increasing w.

For thickness larger than roughly twice the avalanche length a_v a further effect may influence the breakdown field. When the electrons leave a fully grown charge cluster of length a_v in the insulator, the retarding effect of the opposing field on the electrons is large only close to the positive charges (Davies *et al.* 1964, Davies and Evans 1967). The opposing effect decreases with distance from the positive charges, leading to renewal of impact ionization and formation further away of a second avalanche and charge cluster. The second avalanche increases the transit time of the positive charges and the probability of breakdown and the breakdown field for a given value of τ_s is expected to decrease also for this reason with increasing insulator thickness. In ' thick ' insulators a string of avalanches originating from the first electron is expected to arise in series along a channel. Considering the magnitude of a_v in table 2, series of avalanches may occur roughly in insulators thicker than 10^{-4} cm.

As pointed out earlier, when $w < a_v$, so thin that full size avalanches cannot develop, the probability of a breakdown becomes very small. To restore the possibility of incidence of breakdown even at relatively low rates, the field has to be increased to make $a_v \leqslant w$, or with eqn. (5), $F \geqslant \kappa i V_i / w$. This implies that in very thin insulators the field at which breakdowns begin to occur is inversely proportional to w. For thicknesses w larger than a_v the rate of decrease in the breakdown field with increasing thickness for all values of τ_s becomes smaller. These predictions are supported by the earlier mentioned observations of Agarwal and Srivastava (1971) in barium stearate films. Investigating 23 thicknesses between 25 to 2000 Å, the breakdown field was found to be proportional to $w^{-1.01}$ for $w < 250$ Å and proportional to $w^{-0.59}$ for $w > 250$ Å.

5.7. *Concluding Remarks*

Predictions of the localized electronic breakdown model were compared with the experimental observations on oxides, halides and mica. The comparisons produced the following results:

(1) According to the theory the incidence of breakdown is a localized random event both in space and in time. This is confirmed by detailed observations on hafnium dioxide.

(2) Breakdown above a low field limit often occurs over a wide range of fields. The mean time to instability for a breakdown event decreases as some exponential function of increasing field. The breakdown field may be electrode dependent and can either decrease or increase with increasing temperature. These predictions of the breakdown model agree qualitatively with observations. In addition the breakdown field at large τ_s decreases first linearly and then more slowly with increasing film thickness in agreement with the observations on barium stearate films.

(3) The breakdown relations of table 2 show that the slope of the breakdown characteristics $|G|$ is larger than the absolute value of the slope of the injection current characteristic. This was confirmed by the

breakdown characteristics of the oxides of hafnium, tantalum and silicon. The ratio of the two slopes, observed to vary from 2·1 to 3·2, indicates plausible processes for the initiation of breakdown by a succession of a few avalanches in the high field range of the observations. The rate of electron injection is in silicon dioxide many orders of magnitude larger than the breakdown rate, indicating that breakdown is a chance event.

(4) The predicted electrode dependence of breakdown is supported by the good agreement of calculated and observed ratios of slopes G of breakdown characteristics in aluminium oxide with differing electrodes. The measured ratio of the breakdown rates for different electrodes, however, does not agree with calculation. It is possible that the discrepancy is caused by space charges in the insulator.

(5) The activation energies of the temperature dependence of the breakdown characteristics are expected to be smaller than the energy barrier at the electrode insulator interface. This is indicated by the observations on the oxides of hafnium and tantalum.

Qualitative and some quantitative agreements were found between the predictions of theory and experimental observations supporting the validity of the probabilistic model of localized electronic breakdown in insulating films.

ACKNOWLEDGMENT

Thanks are due to Professor J. K. Feldman for critical reading of the manuscript and for many very helpful suggestions.

APPENDIX

The Cathode Field Enhancement on Instability According to Continuum Theory

To determine the conditions for current instability, O'Dwyer (1969) solved the set of equations

$$j = en\mu_n F + ep\mu_p F, \quad . \quad . \quad . \quad . \quad . \quad . \quad (35)$$

$$\frac{\partial F}{\partial x} = e(p-n)/\epsilon_0\epsilon, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (36)$$

$$\frac{\partial j_n}{\partial x} = -\frac{\partial j_p}{\partial x} = -e\alpha n, \quad . \quad . \quad . \quad . \quad . \quad . \quad (37)$$

and

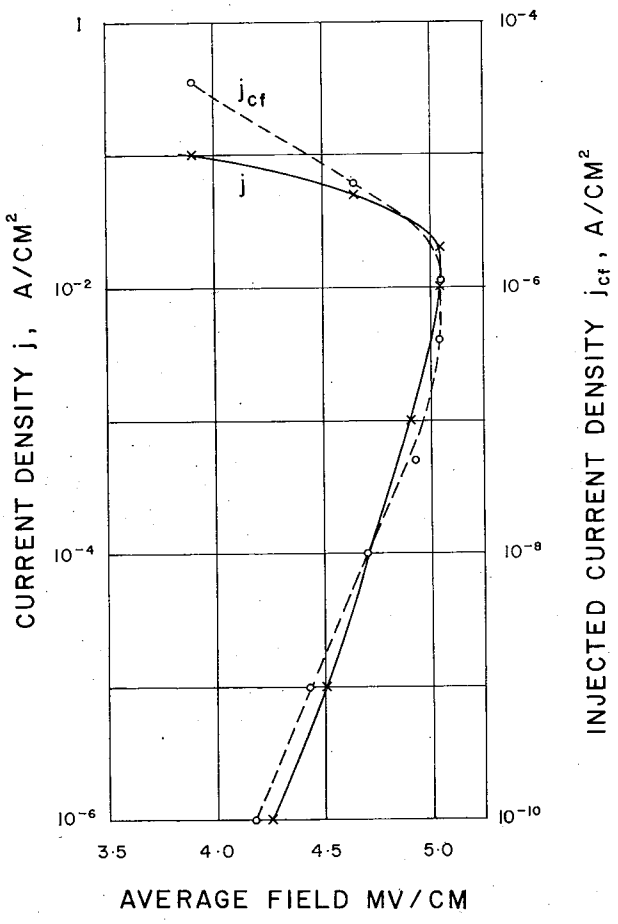
$$\alpha = \frac{B}{F} \exp(-H/F). \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Here n and p are the carrier densities, μ_n and μ_p the mobilities and j_n and j_p the current densities of the negative and positive charge carriers, respectively. Solutions were obtained assuming that the anode is blocking to positive charges, but the cathode is emitting electrons; only

electrons were assumed to cause impact ionization and it is assumed that $\mu_p \ll \mu_n$. As at the boundary the current density was given by eqn. (1), or (2), generalized solutions were not available. Solutions were evaluated for special cases, as illustrated by the F–j characteristics in fig. 5.

The characteristics of fig. 5 were calculated for the cases of field emission at the cathode and avalanching in the insulator, with the data $B = 10^{15}$ v/m^2, $H = 10^9$ v/m, $E = 1{\cdot}25$ ev in eqns. (3) and (2 a), respectively, $\epsilon = 2$, $\mu_n = 10^7$ m^2/v sec and $\mu_n/\mu_p = 10^6$. In the part of the characteristics where the differential resistance is positive, space charges and field distortions in the insulator are negligibly small. Field distortion starts only in the vicinity of instability leading to breakdown. In the special cases of fig. 5, the field distortion at instability was found to be small, the field enhancement at the cathode being only a few per cent of the average

Fig. 13



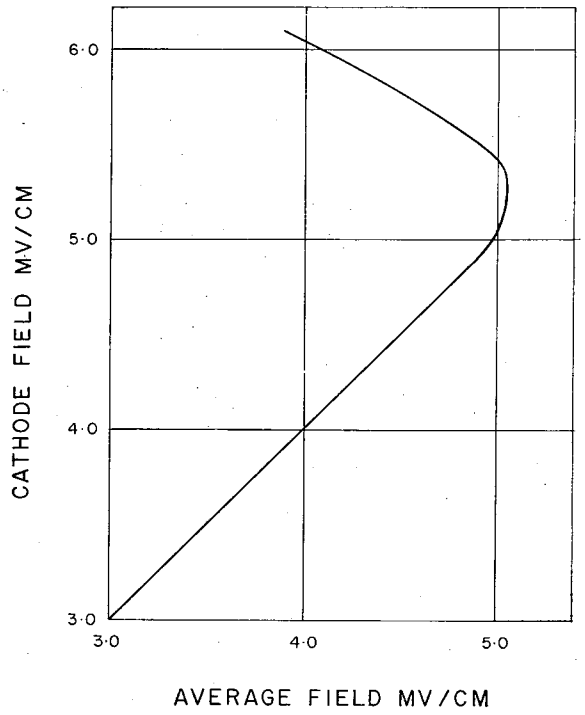
Calculated F–j and F–j_{cf} characteristics in the insulator film with field emission at the cathode and avalanching in the insulator, causing electronic instability. α is given by eqn. (4).

field. This small enhancement was obtained also for a specimen with thermionic emission at the cathode.

O'Dwyer assumed that the ionization coefficient α obeys the exponential relationship of eqn. (3). The present localized theory is based, however, on eqn. (4), where $\alpha = F/V_i$. To verify whether the difference in the relations for α is significant for the results of the continuum theories, the set of equations was again solved by the method shown by O'Dwyer (1969), with α given by eqn. (4). The resulting relations were applied to the specific case, when the ionization voltage $V_i = 10$ v, $\epsilon = 10$, $\mu_p = 10^{-11}$ m^2/v sec, the insulator thickness $w = 2 \cdot 10^{-5}$ cm, injection is by field emission at the cathode with the barrier $E = 2{\cdot}3$ ev and $m^*/m = 0{\cdot}47$. Calculations resulted in the F–j and the F–j_{cf} characteristics of fig. 13 and the F_c versus F curve of fig. 14, j_{cf} being the injected current density. In this case at instability $F_c = 1{\cdot}04F$ and the difference between F_c and F is again a few per cent only.

It is conjectured that the small differences between F_c and F calculated for several cases at instabilities are the general rule. Although this cannot be shown directly, as an explicit relation connecting F_c with F is not available the assumption of small differences is made plausible by the sharp increase of injection current for small increases in F_c.

Fig. 14



Cathode field F_c versus applied field F in the specimen of fig. 13.

References

Agarwal, V. K., and Srivastava, V. K., 1971, *Thin Solid Films*, **8,** 377.
Albert, M., 1972, Private communication.
Antula, J., 1968, *Phys. Stat. Sol.*, **28,** 395; 1969, *Thin Solid Films*, **4,** 281.
Austen, E. W., and Whitehead, S., 1940, *Proc. R. Soc.* A, **176,** 33.
Baessler, H., Riehl, N., and Spannring, W., 1969, *Z. angew. Phys.*, **27,** 321.
Berglund, C. N., and Klein, N., 1971, *Proc. I.E.E.E.*, **59,** 1099.
Budenstein, P. P., Hayes, P. J., Smith, J. L., and Smith, W. B., 1969, *J. vacuum Sci. Technol.*, **6,** 289.
Carslaw, H. S., and Jäger, J. C., 1959, *Conduction of Heat in Solids*, 2nd ed. (London: Oxford University Press), p. 55.
Chou, N. J., and Eldridge, J. N., 1970, *J. electrochem. Soc.*, **117,** 1287.
Cooper, R., and Elliott, C. T., 1966, *Br. J. appl. Phys.*, **17,** 481.
Crowell, C. R., and Sze, S. M., 1966, *Appl. Phys. Lett.*, **9,** 242.
Davidson, A. T., and Yoffe, A. D., 1968, *Phys. Stat. Sol.*, **30,** 741.
Davies, A. J., and Evans, C. J., 1967, *Proc. I.E.E.*, **114,** 1547.
Davies, A. J., Evans, C. J., and Llewellyn Jones, F., 1964, *Proc. R. Soc.* A, **281,** 164.
Davies, W. L., and Collins, R. E., 1969, *Electron. Lett.*, **5,** 462.
Davydov, B., and Shmushkewitch, J., 1940, *J. Phys.*, **3,** 359.
Egawa, H., 1966, *I.E.E.E. Trans. electron. Devices*, **13,** 754.
Forlani, F., and Minnaja, N., 1964, *Phys. Stat. Sol.*, **4,** 311.
Franz, W., 1952, *Z. Phys.*, **132,** 285; 1956, *Handbuch der Physik*, edited by S. Flügge, Vol. XVII (Berlin: Springer), p. 155.
Fritzsche, C., 1967, *Z. angew. Phys.*, **24,** 48.
Fröhlich, H., 1940, *Rep. Br. elect. all. Ind. Res. Ass.* L/T, 113.
Goodman, A. M., 1967, *Phys. Rev.*, **164,** 1145.
Heller, W. R., 1951, *Phys. Rev.*, **84,** 1130.
Hill, R. M., 1971, *Phil. Mag.*, **20,** 59.
Hurych, Z., 1970, *Solid-St. Electron*, **13,** 683.
Inuishi, Y., and Powers, D. A., 1957, *J. appl. Phys.*, **28,** 1017.
Jonscher, A. K., 1967, *Thin Solid Films*, **1,** 213; 1971, *J. Vacuum Sci. Technol.*, **8,** 135.
Kawamura, H., Ohkura, H., and Kikuchi, T., 1954, *J. Phys. Soc. Japan*, **9,** 541.
Klein, N., 1966, *I.E.E.E. Trans. electron. Devices*, **13,** 788; 1969 a, *J. electrochem. Soc.*, **116,** 963; 1969 b, *Adv. Electronics Electron Phys.*, **26,** 309; 1971, *Thin Solid Films*, **7,** 149.
Klein, N., and Burstein, E., 1969, *J. appl. Phys.*, **40,** 2728.
Klein, N., Gafni, H., and David, H. J., 1965, *Proc. Symp. Phys. Failure Electron.*, **3,** 315.
Klein, N., Schuller, J M., and Basseches, H., 1970, *I.E.E. Conf. Publ.*, No. 67, 81.
Lee, C. A., Logan, R. A., Batdorf, R. L., Kleimack, J. J., and Wiegmann, W., 1964, *Phys. Rev.*, **134,** A761.
Lenzlinger, M., and Snow, E. H., 1969, *J. appl. Phys.*, **40,** 278.
Lomer, P. D., 1950, *Proc. Phys. Soc.* B, **63,** 818.
Miles, H. T., and Rawlings, I. R., 1970, Fourth Thin Film Conf., Inst. Phys. and Phys. Soc., Reading.
Misawa, T., 1966, *I.E.E.E. Trans. electron. Devices*, **13,** 137, 143.
Mott, N. F., 1969, *Contemp. Phys.*, **10,** 125.
O'Dwyer, J. J., 1964, The Theory of Dielectric Breakdown in Solids (London: Oxford University Press); 1967, *J. Phys. Chem. Solids*, **28,** 1137: 1969, *J. appl. Phys.*, **40,** 3887.

PLESSNER, K. W., 1948, *Proc. Phys. Soc.*, **60,** 243.
POLLACK, S. R., and MORRIS, C. E., 1964, *J. appl. Phys.*, **35,** 1503.
RAETHER, H., 1964, *Electron Avalanches and Breakdown in Gases* (London: Butterworth).
RIEHL, N., BAESSLER, H., HUNKLINGER, S., SPANNRING, W., and VAUBEL, G., **1969,** *Z. angew Phys.*, **27,** 261.
SCHMIDLIN, F. W., 1970, *Phys. Rev.*, B, **1,** 1583.
SEITZ, F., 1949, *Phys. Rev.*, **76,** 1375.
SHOCKLEY, W., 1961, *Solid-St. Electron.*, **2,** 35.
SIDDALL, G., 1959, *Vacuum*, **9,** 274.
SMITH, J. L., and BUDENSTEIN, P. O., 1969, *J. appl. Phys.*, **40,** 3491.
SOLOMON, P., 1972, Private communication.
STRATTON, R., 1961, *Prog. in Dielect.*, **3,** 235.
SVIRIDOVA, A. A., and SVIKOVSKAYA, N. V., 1967, *Optics Spectrosc.*, **22,** 509.
VAN GEEL, W. Ch., PISTORIUS, A. C., and BOUMA, B. C., 1957, *Philips Res. Rep.*, **12,** 6.
VON HIPPRL, A., and ALGER, R. S., 1949, *Phys. Rev.*, **76,** 127.
WATSON, D. B., HEYES, W., KAO, K. C., and CALDERWOOD, J. H., 1970, *I.E.E.E. Trans. electl. Insulation*, **5,** 58.
WILLIAMS, R., 1962, *Phys. Rev.*, **125,** 850; 1965, *Ibid.*, **140,** A569; 1967, *Phys. Lett.* A, **25,** 445.
WIJSMAN, R. A., 1949, *Phys. Rev.*, **75,** 833.
WOLSCHENDORF, K., 1970, *J. appl. Phys.*, **41,** 5350.
YAMASHITA, J., and WATANABE, M., 1952, *J. Phys. Soc. Japan*, **7,** 334; 1954, *Prog. theor. Phys., Osaka*, **12,** 443.
YOUNG, L., 1961, *Anodic Oxide Films* (London; Academic Press).