Electrical phenomena in amorphous oxide films

G. DEARNALEY,† A. M. STONEHAM,† AND
D. V. MORGAN‡

† AERE, Harwell, Didcot, Berks.
‡ Department of Electrical and Electronic Engineering, University of Leeds

Contents

	Page
1. Introduction	1130
1.1. Charge transport in insulators	1131
1.2. Differential negative resistance	1134
1.3. Dielectric breakdown	1137
2. The forming process	1138
2.1. Observations of forming	1138
2.2. Models of the forming process	1142
3. Conduction, switching and memory phenomena	1146
3.1. Observations of conduction, switching and memory phenomena	1146
3.2. Theories of switching and memory in oxide films	1151
3.3. Observations on triodes	1162
4. Electron emission, luminescence and other phenomena	1163
4.1. Electron emission	1164
4.2. Electroluminescence in M–I–M (metal–insulator–metal) structures	1165
4.3. Theories of electron emission and electroluminescence	1168
4.4. Injection electroluminescence in MOS (metal–oxide–semiconductor) structures	1169
4.5. Noise	1171
5. Comparison with other systems	1172
5.1. Current-controlled negative resistance and bistable switching in oxides	1172
5.2. Comparison with the Ovshinsky devices	1176
5.3. A comparison of switching in crystalline, microcrystalline and amorphous films	1178
6. Practical implications	1179
6.1. Switches	1179
6.2. Memory arrays	1180
6.3. Hot electron, cold cathode emitter	1181
6.4. Alpha-numeric display panel	1181
6.5. Electroluminescent MOS devices	1182
6.6. Oxide-coated cathodes	1182
6.7. MOS capacitors and transistors	1183
6.8. Surface barrier diodes	1184
6.9. Anodic oxides and corrosion	1186
6.10. Anodic electret microphones	1187
7. Conclusions	1187
References	1188

Abstract. Thin layers of insulators, from 100 Å to 20 000 Å thick, have a number of interesting electrical properties. These properties include forming, which is a profound and essentially permanent change produced by an electrical field, and, after forming, differential negative resistance together with switching and memory phenomena. Electroluminescence and electron emission also occur.

These features are shown by most amorphous or microcrystalline insulators of the appropriate thickness, and are particularly clearly shown by the oxides, such as

Al_2O_3, Ta_2O_5 and SiO. In our review we survey the observed properties of metal–insulator–metal devices involving such insulating layers, and discuss the mechanisms which have been proposed for their operation. Further, since the unusual properties may be technologically important, we outline some of the possible uses of these devices and the relevance of the phenomena to the mechanism of corrosion of certain metals.

1. Introduction

Thin layers of insulators, between about 100 Å and 2 μm in thickness, frequently possess rather unusual electrical properties. Some of these were at first dismissed as due to 'breakdown' of an unspecified nature, 'pinholes' through the film, and so on. It is now recognized that the effects fall into a more definable category and that they are not electrical breakdown in the strict sense. These properties include the phenomena of 'forming', which is a profound and essentially irreversible change in conductivity induced by a high electric field. After forming, differential negative resistance together with switching and memory phenomena are observed. Furthermore, in suitable structures, electroluminescence and electron emission may occur.

These properties are displayed by a rather wide range of amorphous or microcrystalline binary inorganic insulators prepared under suitable conditions in films of the appropriate thickness. They are not confined to the oxides but are particularly well demonstrated by oxides such as SiO_x, Ta_2O_5, Nb_2O_5, Al_2O_3—hence our choice of title. Oxide films are common on metals and semiconductors and for this reason the effects have been studied more frequently in oxides than in the halides and sulphides which also sometimes show a similar behaviour. Apparently similar phenomena have been reported in films of a few more complex substances such as cryolite, polymethyl methacrylate (Lewowski *et al.* 1965) and polydivinylbenzene (Gregor 1966).

In this review, we shall survey the observed properties of metal–insulator–metal and metal–insulator–semiconductor structures, and discuss the varied mechanisms which have been proposed to explain their operation. We shall dwell only briefly on the properties of very thin films, below about 50 Å in thickness, where it would appear that electron tunnelling processes can account satisfactorily, and in a relatively conventional manner, for the observed conductivity. Such phenomena have been discussed at length by Duke (1969). Nor shall we be much concerned with the low-field conduction, prior to forming. These low-current conduction mechanisms, while still not fully understood, appear to make only a minor contribution to the electrical behaviour once the field-induced behaviour has been established. Hill (1967, 1971) has given a good review of the conduction mechanisms of unformed insulators. We shall concentrate most attention on the class of structures which exhibits voltage-controlled behaviour in negative resistance and switching. There is another class of devices, exemplified by the chalcogenide glass films and the so-called 'Ovshinsky devices', which shows current-controlled behaviour. This class has received a great deal of attention of late, and we feel that a reasonable explanation of the switching action has now been put forward (Warren 1970). We shall try to put into perspective these two classes of behaviour, both of which are observed in oxide films, and to derive a criterion for deciding which will occur. It may well be that both classes of device share common precursor processes, before conduction fully commences, and that these are easier to observe in the oxide structures.

Although far less publicized attention has been given to amorphous oxide devices, there has been a longer history of observation and interest in them dates back over more than ten years (Kreynina *et al.* 1960); Hickmott in the USA has been particularly active in the field. Moreover, the effects we shall describe are probably significant in several devices of current technological importance, as well as in novel structures which have yet to find a place alongside conventional semiconductor device technology. They may be involved, too, in corrosion and anodic oxidation to an extent not hitherto realized. A recent report by the Science Research Council (SRC 1970) has summarized work in the UK on amorphous materials, yet it reveals exceedingly sparse interest in oxide films. It would be a pity if the glare of publicity afforded to a particular development in the field rendered us blind to other approaches, and determined the shaping of research effort for several years.

For this reason, it seems timely to adopt a broad viewpoint and to summarize for a wider readership the entangled literature on thin oxide films and the problems that remain. There is still a considerable controversy about the mechanisms involved and the degree to which the medium can be regarded as homogeneous and amenable to band theory.

Before considering these matters in detail we shall, by way of introduction, consider rather generally the current-carrying mechanisms in insulators and the nature of dielectric breakdown. After this, the forming process will be considered, and two further sections discuss the associated phenomena. A discussion of the status of the models and a comparison with Ovshinsky devices follows in § 5, and the device implications of the phenomena are considered in some detail in § 6.

1.1. *Charge transport in insulators*

The insulators we shall be considering will largely be noncrystalline and may contain a significant proportion of impurities. The current may be carried by electrons or holes, or by ionic motion. We begin by considering electronic motion in such systems.

Several of these processes show a pronounced temperature dependence, since they involve thermal activation in some form. One common process is Schottky emission, or the field-assisted emission of electrons over the barrier at a metal–insulator interface. The effect is thermally activated and gives a current,

$$J_s = AT^2 \exp - \left(\frac{\Delta_B}{kT} - \alpha \frac{E^{1/2}}{kT}\right) \tag{1.1}$$

where A is the effective Richardson constant, Δ_B the barrier height and E the applied field. The process is said to be 'electrode-controlled'. Generalizations of this result are discussed by Simmons (1965) for cases where the electron mean free path is small. A related phenomenon is associated with emission over barriers which may exist within the bulk of the insulator, typically at localized trapping sites. The field-enhanced thermal excitation into the conduction band of electrons trapped by immobile positive charges is called Poole–Frenkel emission, and gives rise to a current

$$J_{PF} = BE \exp - \left(\frac{\Delta_B}{kT} - \beta \frac{E^{1/2}}{kT}\right) \tag{1.2}$$

where $\beta = 2\alpha$. A more exact expression is given by Jonscher (1967).

Thermally activated hopping between localized electronic states may occur in impurity conduction. For hopping conduction

$$J_{\mathrm{H}} = J_0 E \exp\left(-\frac{\Delta_{\mathrm{H}}}{kT}\right). \tag{1.3}$$

Here Δ_{H}, the effective activation energy, is less well defined, and it is inappropriate to think of electrons propagating with some typical mean free path. Mott (1967, 1968 a, b, c) has observed that the activation energy Δ_{H} tends to zero as the temperature tends to zero, since the only hopping transitions which will be energetically allowed will be those which involve traps equivalent in energy. The activation energy also tends to zero at a critical concentration of traps, when the insulator becomes a conductor. Hopping conduction always requires some degree of compensation (Mott 1956, Miller and Abrahams 1960); further, the ac conductivity always exceeds the dc conductivity and increases with frequency.

If the rate of carrier injection at a contact exceeds the rate of transport away from the contact, we have the bulk-limited situation described as space-charge-limited current flow (Mott and Gurney 1948). A general review is given by Lampert and Mark (1970). For very low injected carrier densities much less then the thermally generated carrier density, Ohm's law is obeyed. The space-charge-limited current has nonlinear dependence on the applied field

$$J_{\mathrm{SCL}} = C\mu \frac{E^2}{d} \tag{1.4}$$

where μ is the electron mobility and d the insulator thickness. The constant of proportionality depends, among other things, on the presence of traps within the insulator. At higher fields $E > (8\pi/\epsilon_\infty) N_0 ed$, where N_0 is the carrier density at the metal–insulator interface, J ceases to be space-charge limited and varies linearly with the field at the cathode interface. The critical voltage for the change to ohmic behaviour is given by

$$V = 3{\cdot}5\left(\frac{N_0}{10^{18}}\right) \mathrm{cm}^{-3} . \left(\frac{d}{100\ \text{Å}}\right)^2 \mathrm{V} \tag{1.5}$$

so that the change may occur within the range of parameters we are considering.

Besides the space-charge-limited current we have already considered there are other electronic processes which do not exhibit a pronounced temperature dependence. In very thin films (typically less than 100 Å) tunnelling can be important (Mead 1961), but it need not concern us for the range of film thickness under discussion in this review.

There are some circumstances in which an insulator may become conducting, for example, by virtue of a phase transition, either as a whole or in localized regions. It is clear that in these cases, the current–voltage relation can have the usual Ohm's law form once the conducting phase has become established.

Ionic conduction shows the thermally activated form:

$$J_{\mathrm{ion}} = \left(\frac{GE}{T}\right) \exp - \left(\frac{\Delta_{\mathrm{I}}}{kT}\right). \tag{1.6}$$

Usually the ionic conductivity is small, since the ionic mobilities are small. If the ionic mobility (drift velocity per unit applied field) is typically $10^{-3}\ \mathrm{cm^2\,s^{-1}\,V^{-1}}$ at

1200 K and Δ_I is 1 eV, then at 300 °C the mobility is only 10^{-16} cm^2 s^{-1} V^{-1}. In this case the mean time to cross a film of thickness 100 Å under an applied field of 5 V is about half an hour, considerably longer than any of the characteristic times with which we shall be concerned. However, we cannot rule out the possibility of rapid diffusion which results from the amorphous or microcrystalline nature of the insulator. In such cases the mobility should be sensitive to the detailed mode of preparation of the specimen.

In deciding the mechanism of charge transport, difficulties arise similar to those which occur in discussions of tunnelling (Duke 1969). Theoretically, it is customary to use a phenomenological approach. This uses one-electron theory, often in the form of the effective-mass approximation and a simple model of the band structure. The one-electron potential contains image-force terms in the insulator and a

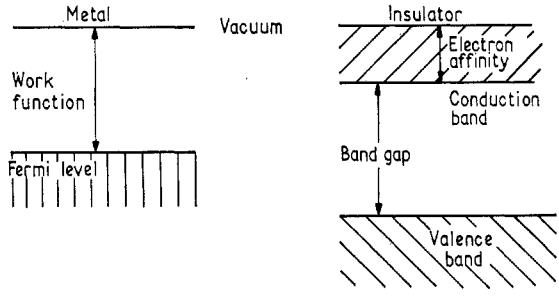

Figure 1. A schematic representation of the band structure of a metal and an insulator. The vacuum levels of the bulk components are equated at the interface so that the results are expressed in terms of the metal work functions and the insulator electron affinities.

contribution, calculated from Poisson's equation, to allow for local deviations from charge neutrality. It is difficult in practice to distinguish between surface charges and bulk space-charge terms, so further assumptions are necessary. Further, fluctuations in the position, potential energy discontinuity and electric field at the interfaces are ignored. The neglect of fluctuations may lead to qualitative errors, and possibly to a misunderstanding of effects which arise, in part, from such fluctuations. Dielectric breakdown, for instance, is affected by the details of the interface (Alpert *et al.* 1964). The relative positions of band structures in adjacent components are treated as adjustable parameters. At best, these are fixed by independent experiment. However, it is more common to equate vacuum levels of the bulk components, so that the results are expressed in terms of metal work functions and insulator electron affinities (figure 1). In general, this equation is not accurate, since the discontinuity at the interface can disturb the local electronic structure of the components considerably. Moreover, even nominally 'symmetric' devices may not have symmetric properties if the method of preparation differed for the two metal–insulator faces. This asymmetry is usually associated with the kinetics of oxide growth (Smith and Miles 1963, Fromhold and Cook 1967 a, b). It is only possible to assume general trends of behaviour with work function and electron affinity. Experimentally, one must ask how a particular mechanism may be verified, whether the dependence on fabrication (if any) is consistent and, in particular, whether the interpretation is unique. The mechanism of carrier transport in the thin films under discussion has not been conclusively identified, and the several interpretations are very different.

1.2. *Differential negative resistance*

Systems which show differential negative resistance are divided into two classes: the current-controlled (S-type) and the voltage-controlled (N-type) types. The current–voltage characteristics are shown in figure 2. The thin-film devices we shall primarily be concerned with have an N-type characteristic.

The mechanisms which give rise to differential negative resistance fall into three broad categories: (i) processes in which the Joule heating of the conduction electrons causes a change in their number or in their mobility, (ii) processes in which special semi-permanent space-charge distributions are set up, and (iii) processes which involve a phase change or atomic rearrangement of the host insulator. A number of these processes are reviewed by Volkov and Kogan (1969).

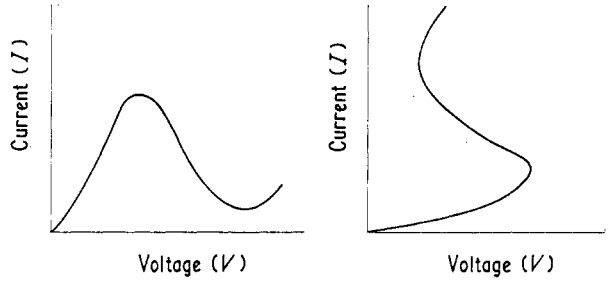

Figure 2. Two types of differential negative resistance: (*a*) voltage controlled (*b*) current controlled

The most commonly considered mechanisms are those in the first category. One example is the *superheating* mechanism. Certain scattering processes, such as electron collisions with charged ions or electron interactions with piezoelectric phonons, become less efficient at high electron energies. The electron mobility can then rise suddenly, giving an S-type characteristic, once enough of the electrons have sufficient energy. A second example is the *Gunn effect*. In certain semiconductors, such as GaAs and InP, the conduction band has a minimum at $k = 0$ together with higher-energy minima with larger effective masses. As the electrons are heated they populate the higher minima, lowering their mobility ($\mu \propto m^{*-1}$) and giving an N-type characteristic. In a third example, also giving an N-type characteristic, higher-energy electrons are more easily captured by trapping centre, so that their mobility is reduced. Thus Au^{2-} in Ge can trap an electron, but because of the long-range repulsive field, such trapping is only possible for electrons of higher energy. A fourth process can occur in compensated semiconductors. Impact ionization can increase the number of carriers; this, combined with the decrease of scattering at higher electron energies, gives rise to an S-type characteristic.

The second category, in which a special space-charge distribution is set up and which gives rise to the differential negative resistance, includes three models of particular relevance. These are the models of Hickmott (1962, 1964 a, b, 1970 private communication), of Simmons and Verderber (1967 a) for oxide films and of Mott (1968 a, b, 1969) for the chalcogenide 'threshold' switch. We shall also include the model of Barriac *et al.* (1966, 1968, 1969) in this category, although it does involve other features, such as ionic conduction. All these models will be discussed in more detail later.

Hickmott proposed that the initial part of the current–voltage characteristic was space-charge limited conduction in an impurity band roughly in the middle of the insulator bandgap. The electron density of states is shown in figure 3. At a sufficient voltage the number of impurity centres which contribute to the conduction is then reduced by some process, so that the conductivity is reduced. An N-type characteristic results. One possible process would be the neutralization of the impurity centres by field-induced inter-impurity tunnelling from a lower lying band of defect levels.

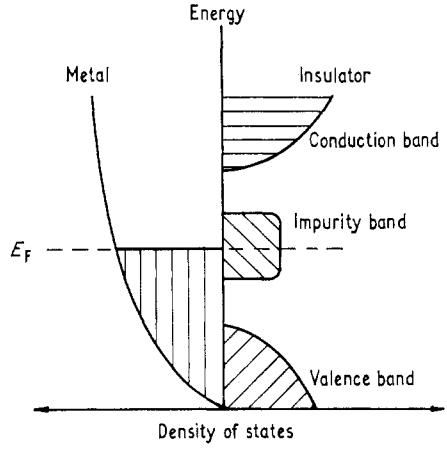

Figure 3. The density of states at a metal–insulator interface, with an impurity band roughly in the middle of the insulator band gap.

Simmons and Verderber suggested that a defect band would be formed with a density of states shown in figure 4. The electron transport is assumed to be hopping between traps of very similar energies. At voltages less than ϕ_I electrons can tunnel from one electrode to the other via impurity states. When the applied voltage exceeds ϕ_I, electrons entering the insulator with energies close to the electrode Fermi energy can only cross the insulator if they make an (improbable) transition to traps of lower energy. Consequently the current is dominated by the electrons with energies more than $(\phi_I - V)$ below the electrode Fermi level. The density of states of electrons within the electrode decreases with energy below the Fermi level, and so the number of electrons contributing to conduction falls rapidly, giving the negative resistance.

Barriac and co-workers introduce the extra feature of ionic motion. As with the last two models, they predict an N-type characteristic. At low voltages the current is suggested to be a space-charge limited *ionic* current. At higher voltages electron tunnelling occurs, which neutralizes the positive ionic space charge by being trapped so that the two current-carrying components are largely cancelled.

Mott's model for the threshold switch differs from these in that it predicts an S-type characteristic. A Schottky barrier is set up at both interfaces, with space charge, giving a strong field at the boundaries. Electrons and holes tunnel through the two interfaces, through a barrier which increases in height but decreases in thickness as the voltage increases. The holes form a space charge near the electrode at which the electrons are injected, and vice versa. This space-charge, when formed, enhances the current, giving the high-current state of the S-type

characteristic. The space-charge is sufficiently stable against electron–hole recombination because the holes are trapped in regions where the electron velocities are high. When the current drops below a certain minimum, however, it is no longer possible to sustain the space-charge, and the low-current state results.

Finally in this section, we turn to theories which involve an atomic rearrangement. This may be a bulk change, or it may involve only small regions of the insulator. Examples in which there are bulk transitions include the switching of vanadium oxides, where Joule heating is sufficient to raise the oxide to a temperature

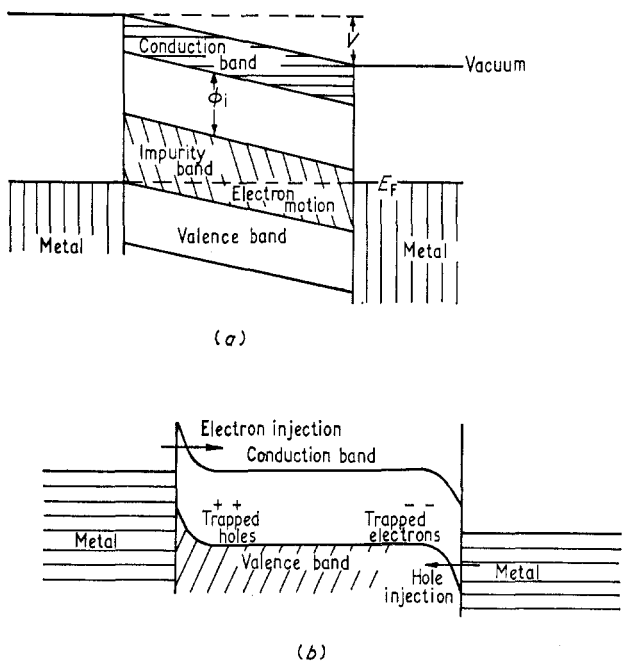

Figure 4. (*a*) A suggested defect band formed by the diffusion of metal ions from the anode into the insulator (after Simmons and Verderber). (*b*) The suggested form of conduction and valence bands in the low-impedance state of a threshold switch. The curved lines show how the potential energy of an electron, at the top of the valence band or bottom of the conduction band, varies with position across the glass film, under the influence of the applied field produced by the trapped electrons (after Mott 1968 a).

above that for a phase change (Cope and Penn 1968). The high-temperature state is more highly conducting, and an S-type characteristic results. Other models involve the changes in behaviour of conducting filaments. In the 'bistable' switches filaments have been observed directly, and their behaviour appears to be associated with a glass transition, and possibly with recrystallization. Mott has suggested that the increase in conductivity, giving an S-type characteristic, results from an atomic rearrangement which leaves some atoms with incompletely saturated bonds. The low-current state of these devices corresponds to an amorphous atomic configuration in which essentially all atoms have saturated bonds. Two approaches which have features in common are those of Gibbons and Beadle (1964) and of Dearnaley (1967). In these models the conduction is ohmic (that is, with zero activation energy) along filaments which are physically different from the host matrix. In Dearnaley's model these filaments can fracture, becoming nonconducting, and the differential negative

resistance results from this fall in conductivity on fracture. Gibbons and Beadle postulate one filament; switching occurs but there is no negative resistance. One plausible cause of fracture would be the Joule heating of the filament to beyond its melting point at some point along its length.

Despite the diversity of mechanisms for differential negative resistance there are certain general results. Ridley (1963) has shown that a system showing S-type differential negative resistance is unstable against the formation of current filaments, whereas an N-type system is unstable against the formation of high-field domains. These general results do not apply directly to systems in which there can be atomic rearrangement, and so do not rule out models which invoke filaments and predict N-type behaviour from structural changes.

1.3. *Dielectric breakdown*

The forming process, which we discuss in the next section, bears some resemblance to dielectric breakdown. Forming is *not* breakdown, although there are similarities, described here, but is rather a distinct and novel high-field phenomenon which occurs near breakdown in certain amorphous and microcrystalline dielectrics.

The classical theory of breakdown, reviewed by Whitehead (1951) and O'Dwyer (1964), is based on the concept of an electron avalanche. The simplest forms (for example Fröhlich 1947, Mott and Gurney 1948) indicate that breakdown occurs when electrons in the insulator gain energy from the applied field at a rate which exceeds their energy loss. Developments of this theory have been made by O'Dwyer and by Forlani and Minnaja (1964). These avalanche theories predict a breakdown field as a function of film thickness and temperature. Although the field is not independent of thickness except in certain limits (for instance, the thickness is much greater than the electron–hole recombination distance in the Forlani–Minnaja theory) the predictions always indicate a thickness-dependent breakdown *voltage*. This feature is also true of the different approach of Klein and Gafri (1966). Detailed experimental studies of thin films (see, for example, the work on SiO, MgF_2, CaF_2 and some other films by Budenstein *et al.* 1969) confirm that the breakdown voltage does indeed depend on film thickness.

In a recent review of dielectric breakdown Klein (1969) has commented on the anomalous behaviour of oxide films. Unlike the well-behaved dielectrics, these exhibit a wide range of breakdown fields, each breakdown event destroying a very small part of the insulator. Klein concludes that "as the nature of the avalanche-producing breakdown is not yet known, the development of a satisfactory theory requires further experimental work, mainly aimed at the identification of the processes forming a conducting channel".

Some features appear to be electrode-dependent. The breakdown is not a sudden process at a critical field, but may occur with varying speed over a range of fields. These features are important in some of the models of forming. Worthing (1968) has studied the Si–SiO_2–Au system and finds that, for Au negative, the breakdown is indeed the usual intrinsic or disruptive breakdown. When the gold was positive, however, a time-dependence of the breakdown was observed, with (time for breakdown)$^{1/4}$ $\propto$ (applied voltage + constant). The same effect was found for Al or Pt instead of Au. This time-dependence vanished at nitrogen temperature. During the incipient breakdown, the small current passed shows considerable noise, and light emission is seen from local spots. These local spots do not seem to be associated

with pinholes existing before, and appear to be related to the formation of conducting paths. Similar results have been reported by other workers, including Laverty and Ryan (1969).

2. The forming process

Once the metal–insulator–metal device has been constructed, a radical and essential permanent change can be caused in its electrical properties by application of a voltage greater than a minimum, V_F. This process is called *forming*, and V_F is known as the *forming voltage*. We now survey the observed features of the forming process before outlining the various models proposed.

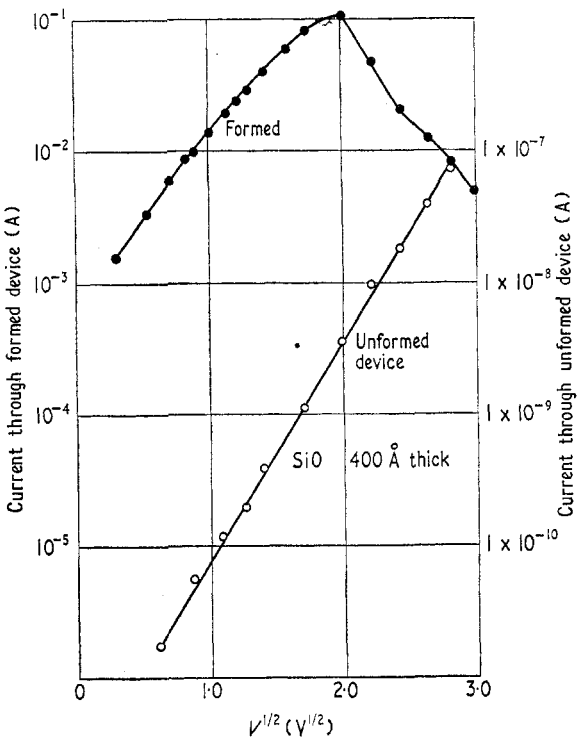

Figure 5. Current–voltage characteristics of unformed and formed device (after Simmons and Verderber 1967).

2.1. *Observations of forming*

Prior to forming the conduction in the insulator is by one of the thermally activated mechanisms, giving a current I where

$$\ln(I) \propto (V)^{1/2} \tag{2.1}$$

though in some impure materials the behaviour may tend towards a linear dependence upon V. The thermal activation energy is typically a few tenths of an electron-volt. Application of a voltage pulse $V \geqslant V_F$ of a few seconds' duration at room temperature and in vacuum produces a profound change in the current voltage characteristic. The current at a given voltage is increased by a factor up to 10^8, and the new characteristic shows differential negative resistance at low frequencies. The two characteristics—formed and unformed—are shown in figure 5. We shall later

discuss this current–voltage behaviour in detail. Instead of by a voltage pulse, the forming may be achieved by applying a sinusoidal voltage of amplitude greater than V_F. The development of the forming process can then be followed as the amplitude is increased through V_F. This behaviour is shown in figure 6, after Barriac *et al.*

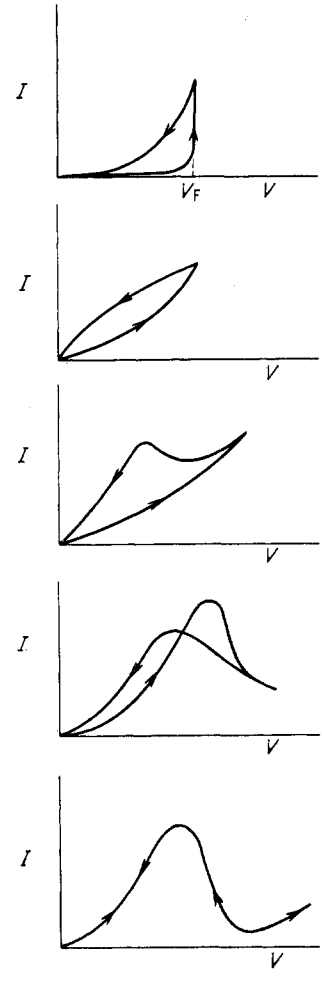

Figure 6. The development of the forming process as the amplitude of the applied sinusoidal voltage V is increased through the forming voltage V_F (after Barriac *et al.* 1969).

(1966). Similar behaviour has been described by Hickmott (1964 a, b). For clarity, the very considerable noise has been omitted in the diagram. In certain cases forming has been achieved by x irradiation instead of applying a field.

The occurrence of forming and its extent (measured by the increase in conductivity) depends on the insulator, on the electrodes and on the temperature and atmosphere during forming. It is of especial importance that *the forming voltage does not depend on the thickness of the insulator* (figure 7). The forming process is voltage-controlled. We shall return to this point later. The degree of forming does depend on the thickness, however, and the peak current obeys the approximate relation:

$$I_{max} \propto d^{-3} \tag{2.2}$$

due to Verderber *et al.* (1967) and shown in figure 8. Forming tends to take longer for thicker devices, and it is not usually achieved if $d > 1$ μm. At the other extreme, with $d < 100$ Å, forming is also not possible. This is partly because a voltage V_F

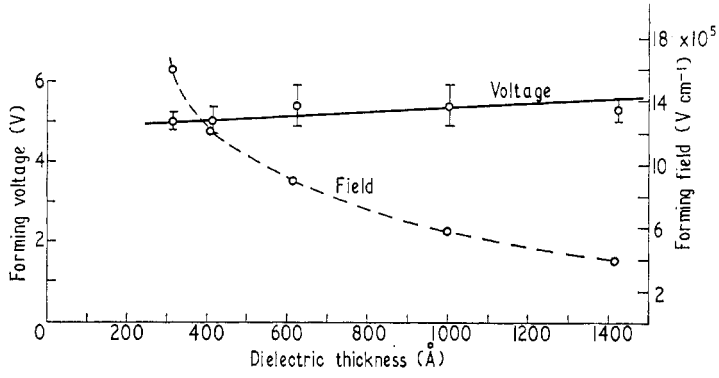

Figure 7. The dependence of the initial forming voltage and forming field (V_F divided by thickness) upon insulator thickness (after Verderber *et al.* 1967).

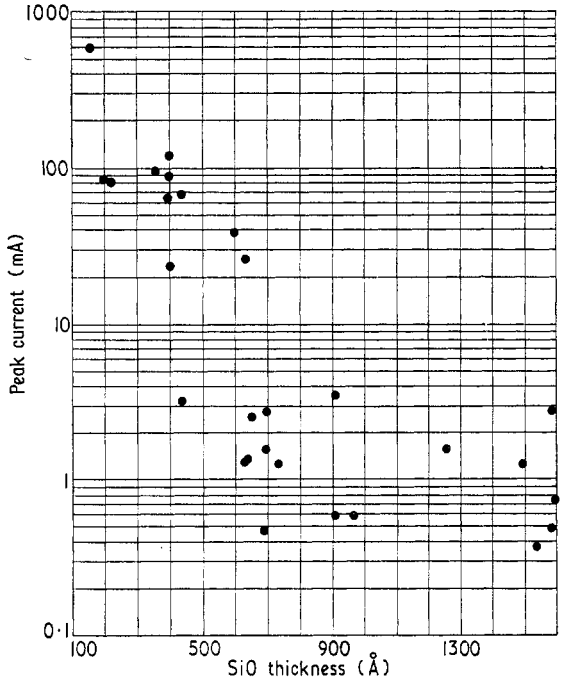

Figure 8. Plot of peak current through a formed sample for different insulator thicknesses (after Verderber *et al.* 1967).

causes a field in the dielectric in excess of that for dielectric breakdown, E_{BD}

$$V_F \geqslant E_{BD}\, d \tag{2.3}$$

and partly because electrode–electrode tunnelling becomes significant.

Forming occurs most readily in insulators with reactive anions, such as oxides (SiO_x, Ta_2O_5, Al_2O_3) and fluorides (CaF_2, MgF_2, MnF_2), and appears to be easier when the insulator is nonstoichiometric. Some oxides, for example Cu_2O, show

S-like character, particularly when the thickness is several microns (see, for instance, Cook 1970). During the forming process there is emission of gaseous oxygen from the oxides and fluorine from fluorides—which produces bubbles between the insulator and an electrode. This localized formation of bubbles may be related to other localized phenomena—luminescence and electron emission—discussed later. The bubbles may subsequently coalesce and spoil the electrical contact between insulator and metal. The forming voltage V_F shows a number of correlations with bulk insulator properties. Greene *et al.* (1968) observed that V_F was related to the Gibbs free energy for formation of the insulator. Thus, for a crystal of monovalent ions MX

$$V_F \simeq (\text{Gibbs free energy of formation of MX}) \times (\text{Faraday constant}).$$

Other correlations can be found; in particular we note that

$$V_F \simeq \tfrac{1}{2}(\text{bandgap}) \tag{2.4}$$

$$V_F \simeq \tfrac{1}{2}(\text{Madelung energy}) \tag{2.5}$$

also appear to give good results (table 1). However, these formulae involve perfect crystal properties, whereas V_F does depend on the impurity content.

Table 1. Forming voltages compared with other crystal parameters

Crystal	Observed V_F	From Gibbs free energy	$\frac{1}{2}$ bandgap	$\frac{1}{2}$ Madelung energy
LiF	6·1	6·0	6·8	6·2
KCl	4·5	4·2	4·45	4·0
KBr	4·1	3·9	4·0	3·8
CsI	3·8	3·4	3·2	3·2
CaF_2	6·2	6·0		
MgF_2	5·4	5·4		
MnF_2	3·8	3·8		
SiO	5		2	
Al_2O_3	4·1†		3·6 to 4·1	

† Hickmott (1970 private communication) gives 3·2 to 4·1 V, dependent on thickness (usually 3·7–4·1 V when d = 200 Å). He notes also that V_F tends to be higher in very pure materials.

The dependence on electrode material is different for the anode, to which forming is sensitive, and the cathode, which does not appear to affect forming. The forming of a device M_A–I–M_B may fall into one of three categories. In the first forming occurs readily and at the same voltage whether M_A or M_B is the anode. In the second forming occurs only at higher voltages when one electrode (say M_B) is the anode than at the usual forming voltage. The third class will not form at all when one electrode (say M_B) is the anode. In table 2 we list the various anodes together with the class of junction they give. Results are given for Al_2O_3 (Hickmott 1965 b) and SiO (Simmons and Verderber 1967 a). The SiO results are listed in order of ease of forming, measured by the time needed to form and the degree of forming. The Al_2O_3 list is chosen to parallel the SiO one. It is clear that the general division is much the same in the two lists. Elements with high work functions that do not readily form stable oxides or fluorides generally give easy forming. In other instances the oxide is gaseous, or will itself undergo forming. By any other standards the appearance of carbon and silicon in the first category is surprising (see table 3).

The occurrence of forming is also sensitive to the atmosphere. Oxygen inhibits forming, even in fluorides, whereas inert gases have no effect. The maximum tolerable gas pressure is not given with any consistency in the literature. Hickmott found 1 torr or better adequate; Simmons and Verderber usually used 10^{-6} torr, but found 10^{-2} torr adequate, and Barriac *et al.* find 10^{-5}–10^{-8} torr necessary, with complete inhibition at 10^{-3} torr or above. The formed device is still sensitive to the atmosphere, at least when $V > V_T$. Even dry inert gases at atmospheric pressure can have some effect although these changes are not permanent.

Table 2. Ease of forming: Dependence on anode

	SiO (Simmons and Verderber)	Al_2O_3 (Hickmott)†
Easy forming at voltage V_F fixed largely by the insulator	Pd, Ir, Pt, Au, Ag, Cu, C, Si, Ni, Co, Fe, Sn	Ag, Au, Cu, Sn
Forming only at higher voltages		Co, Pb, In, Bi, Al
No forming	Zn, Pb,‡ Cr, In, Mn, Be, Al, Mg	Mg

† Hickmott (1970 private communication) data for *impure* oxides agree; for pure oxide he usually gets breakdown first except for Au, Ag.

‡ Hickmott (1970 private communication) reports forming of SiO under a Pb anode.

Table 3. Work functions

Metal	Capacitance–voltage	Photo-response	Vacuum work function
Ag	5·1	5·05	4·3
Au	5·0	5·0	4·8
Cu	4·7	4·7	4·25
Ni	4·55	4·6	4·5
Al	4·1	4·1	4·25
Mg	3·35	3·15	3·7

Differences appear in part because one cannot simply equate vacuum levels at the interface; there will be a modification of local charge density, and there may be surface states—see Duke 1969, p. 21. Metals are listed in order of effectiveness as electrodes in the forming process, and the data are taken from Goetzberger and Sze (1969).

Oxygen, however, can have permanent effects, and essentially reverses the forming process. We note at this stage the similarity to the activation and poisoning of oxide-coated cathodes (see § 4.3).

Changing the temperature has only small effects on V_F (figure 9). However, the rate of forming falls off rapidly and there is no forming in SiO at liquid nitrogen temperatures. The noise in the forming process increases as the temperature rises.

2.2. *Models of the forming process*

We now survey five different mechanisms which have been proposed to explain forming. For convenience a concise summary of them is given in table 4. The mechanims are generally designed to lead to one of the explanations of differential negative resistance, outlined earlier, and which we discuss in more detail in § 3.

Table 4.

Author	Model proposed	Local or general forming	Dependence on electrode	Dependence on insulator	Dependence on atmosphere
Hickmott	Schottky ionization of impurities near middle of bandgap; formation of localized states for impurity band and space charge to aid injection of electrons	Uncertain	No detailed model; probably depends on work function	$V_F \simeq \frac{1}{2}$(energy gap); voltage-dependent since ionization of impurities is so	No detailed model; probably impurity–oxygen reaction
Simmons–Verderber	Injection of ions from anode to form broad impurity band with a sharp top. Band bending from space charge	Not local	Impurity levels depend on species; band bending depends also on work function	V_F usually less than $\frac{1}{2}$(energy gap); depends on electrode, through impurity levels. Voltage-dependent because impurity spectrum and bandgap fix V_F	Apparently none since the injected ions do not react strongly with oxygen
Dearnaley	Propagation of conducting filament through film	Local	Forming initiated at anode and can be inhibited if electrolytically released gas reacts readily with anode	V_F probably related to crystal formation energy or to breakdown strength. Voltage-dependent because forming initiated in small region	Oxygen causes competing reactions which prevent filament propagation or initiation
Barriac *et al.*	Local fusion of insulator and electrode, with subsequent ion injection. Also traps formed	Probably not local	Not discussed	Not discussed	Oxygen neutralizes the ions which constitute the space charge
Greene *et al.*	High-field electrolysis with injection of vacancies from cathode	Probably local	If the anode reacts readily with gas released growth of insulator at anode may inhibit forming	$V_F \simeq$ Gibbs free energy of formation of insulator	Oxygen causes other competing reactions at cathode

Hickmott (1964 a, b) suggested that there were immobile neutral impurities in the insulator, with ionization energies of about half the bandgap. In the forming process the impurities are ionized by the Poole–Frenkel process, giving an impurity band roughly in the middle of the bandgap, together with surface charges which allow electrons to enter the crystal without thermal activation. The voltage, rather than field, dependence of forming appears because the ionization of the impurities becomes significant at

$$V > V_F \simeq \frac{E_g}{2}. \tag{2.6}$$

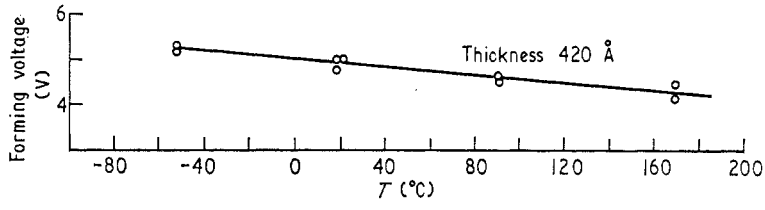

Figure 9. Forming voltage at various temperatures for a Au–SiO–Al device (after Verderber *et al.* 1967).

Like the other authors, Hickmott confirms that forming is not the same as electrical breakdown, since breakdown occurs at higher voltages. The difficulties with this model lie in part in the features which are not explained explicitly, such as the dependence on atmosphere, the release of gas and the dependence on electrode, and in part that rather severe demands are placed on the impurity levels. The impurities must be sufficiently dense that a satisfactory impurity band be formed, and in the correct energy range for injection to occur without thermal activation. It seems surprising that these features could occur in all the very wide range of insulators studied (if indeed the mechanism is the same in all cases). This point is emphasized further because his model of differential negative conductivity also involves an additional impurity band.

Simmons and Verderber (1967 a) proposed that forming consisted of the injection of metal ions from the anode into the insulator. This gave a broad band of impurity states which allowed conduction. Their model explains the dependence on electrodes, since the energy spectrum of the impurity levels depends on which species of ion is injected. The temperature dependence of forming follows since the ionic mobility is temperature dependent. Further, by making assumptions about the distribution of ionic space-charge and the resulting fields it is possible to argue that forming should be voltage dependent; this suggests that $V_F < E_G/2$. In more detail, Verderber *et al.* (1967) argue that V_F is that potential which (allowing for space charge) gives a non-zero injecting field at the insulator–metal interface. They show this gives

$$V_F = 2\,(\text{metal work function}) - (\text{insulator electron affinity}) - (\text{activation energy for unformed device}). \tag{2.7}$$

The objections to this theory are twofold. On the one hand, the forming process seems to show none of the dependence on preparation or on insulator thickness and on electrode ionic mobility which would be expected for the wide range of cases studied. Nor is the dependence on atmosphere satisfactorily treated.

More telling, however, is the fact that the concentration of injected ions can be measured by Rutherford back-scattering. This is particularly straightforward for Au–SiO–Si devices, and shows that the Au concentration is at least a factor of 30 below that necessary for this theory (Dearnaley *et al.* 1970).

Dearnaley developed a filamentary model of forming. He argued that since the anode–insulator interface is not completely smooth there will be places where the electric field is particularly high. Near these places a structural change in the insulator occurs, giving a conducting filament. The electric field in the insulator near to the tip of the filament will be especially high, so that the conducting filament can propagate through the insulator until it forms a conducting link between the electrodes. This theory predicts a voltage-dependent forming if the initiation of the filament is a localized phenomenon and requires a higher electric field than the propagation of the filament, once initiated. Suppose the region in which initiation occurs at the electrode–insulator interface has a radius of order R. If the insulator thickness is d, then the electric field at the boundary of the region is

$$E(R, d) = E(R, d = \infty)\left(1 + \frac{\eta R}{d} + \ldots\right) \tag{2.8}$$

where η is of order unity and depends on the detailed structure of the interface. The observed dependence of V_F on d then allows an estimate of R, and the data in figure 7 suggest $R \simeq 20$ Å. The filaments would be hard to detect microscopically. The temperature dependence of the forming rate follows naturally, since the structural changes involve thermally activated ionic motion. The gas emission also occurs if the structural changes are electrolytic, and the dependence on ambient atmosphere is presumably a consequence of the equilibrium of the electrolytic reaction being affected. The dependence on the electrode results from the reaction of the gas released (usually fluorine or oxygen) with the anode to form a layer which prevents further filaments being started. Thus, as observed, the less reactive elements are the best anodes.

However, it is hard to see immediately what microscopic picture of the filament should be adopted so as to give the observed conductance. Conduction along ...S–O–Si... chains via π-electron overlap was originally suggested for SiO (Dearnaley 1967, Revesz 1967), but clearly this cannot be extended to all the systems which exhibit forming. A defect configuration, ...M–□–M–□... consisting of alternate anions and cation vacancies was also proposed (Dearnaley 1969, Greene *et al.* 1969). It is feasible that conduction then takes place by hopping between adjacent localized anion sites—this phenomenon would be the (more-or-less) one-dimensional analogue of the Mott insulator–metal transition (§ 1.1), and might satisfactorily account for the low activation energy. There are some similarities between the model of Dearnaley *et al.* and that of Gibbons and Beadle (1964), proposed to explain switching in thin NiO films. In the latter case, however, a single filament consisting entirely of metallic nickel was considered to form in the oxide—switching occurs but negative resistance does not result in this case. In the model of Dearnaley *et al.*, very many filaments are postulated to develop at nucleation sites at the anode interface: this is feasible as long as (i) each filament introduces only a small change in the circuit resistance, and (ii) the applied voltage exceeds the forming voltage, V_F.

Greene *et al.* (1969) argued that the forming mechanism is a high-field electrolytic process in which anion vacancies are injected into the insulator. These vacancies give rise to localized conduction paths through the layer. At low fields the

reactions at the (assumed inert) electrodes for an insulator MX are

$$X^- \longrightarrow X + e^- \quad \text{(anode)}$$

$$M^+ + e^- \longrightarrow M \quad \text{(cathode)}$$

with the anions given off as gas. At high fields a further cathode process occurs:

$$(\text{X vacancy}) + e^- \longrightarrow (\text{X vacancy with trapped electron})$$

and vacancy injection results. Since atomic transport is involved the forming rate should decrease with temperature, as observed. Moreover, the forming voltage, V_F, should be related to a Gibbs free energy of formation, as is shown in table 1. The dependence on electrodes can be explained by noting that if the anode is reactive, the insulator can grow at the anode through combination of insulator with anions with the electrode, and this can interfere with subsequent forming. The interference of an oxygen atmosphere follows from the existence of other competing cathode reactions, such as

$$\tfrac{1}{2}O_2 + 2e^- \longrightarrow O^{2-}.$$

Greene's model accounts rather well for the forming in ionic crystals. It is not clear if the mechanism can be adapted to other insulators, such as SiO_2, which show forming.

The model of Barriac *et al.* (1968) involves the injection of ions from the anode, giving the ionic conduction and space charge needed for their mechanism of differential negative resistance. The injection is initiated by a local fusion of the insulator and electrode by a spark with large heat release. Electron traps are also created during the forming process that can only trap electrons when there is a sufficient electric field. The influence of the atmosphere arises because the oxygen can permanently neutralize the ions which constitute the space charge. This model suffers from the objections to ion injection which were mentioned in connection with the Simmons–Verderber model. It also leaves questions unanswered, such as the nature of the electron traps, and the dependence of V_F on the electrodes and on the insulator.

3. Conduction, switching and memory phenomena

The formed metal–insulator–metal device shows N-type differential negative resistance and, associated with this, may be switched between states of different conductivity. We first survey the experimental data, and then review the various models which attempt to explain these conduction phenomena.

3.1. *Observations of conduction, switching and memory phenomena*

We begin by considering the low-frequency current–voltage characteristics (frequency $\ll 100$ Hz for SiO). As the voltage increases from zero the characteristic passes through several stages

(i) The current increases with voltage rather faster than expected from Ohm's law.

$$I \propto I_0 \sinh\left(\frac{V}{V_0}\right) \quad \text{(Simmons and Verderber)}$$

$$I \propto I_0\left(\frac{V}{V_0}\right)^2 \quad \text{(Hickmott)}$$

where I_0 and V_0 are constants.

(ii) The current goes through a maximum at a voltage V_{max} and then shows differential negative resistance. The decrease in current is roughly exponential in voltage. The current is then very noisy, the current fluctuations being as much as 10%, as opposed to 1% or less in other voltage regions.

(iii) The current passes through a minimum at V_{min}, and rises slowly as the voltage increases. Ultimately dielectric breakdown occurs over the whole device. As the voltage is reduced the current follows essentially the same characteristic as before, sometimes with slight hysteresis and sometimes with a slight displacement of V_{max}. The current characteristic is more or less symmetric in a formed device, and does not depend on which electrode is biased positive. This is also true of the switching and memory phenomena treated later. Figure 10 shows typical characteristics for SiO.

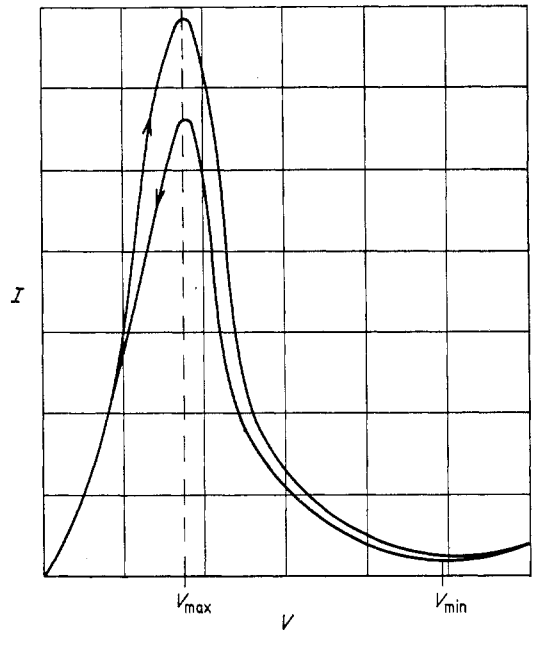

Figure 10. Current–voltage (I–V) characteristics of a formed SiO_x device for increasing and decreasing voltage sweeps (after Simmons and Verderber 1967 a).

The various memory states of the device may be achieved by applying a voltage $V > V_{max}$, and then reducing the voltage rapidly. For SiO times of less than 10^{-4} s are necessary. The conductivity is then fixed by the voltage prior to switch-off. This is shown in figure 11. The new current–voltage characteristic is essentially permanent provided the voltage does not exceed V_T, which is rather less than V_{max} and characteristic of the insulator. The memory state is long-lasting (it lasts for at least two years) and is stable against mild heating and optical illumination. It can be erased by applying a voltage in excess of V_T, when the current rises in steps to the original characteristic (figure 12). The current in the memory states can be varied continuously by a factor up to 1000.

The switching to and from the memory state can be rapid. A pulse longer than 10^{-7} s is needed to switch from the high impedance (low current) to the low impedance state in SiO. The reverse process, from low impedance to high, can be

achieved with a pulse of only 2×10^{-9} s. However, the device cannot be cycled at a frequency corresponding to the longer time. There is a 'dead time' which must elapse before the device can be returned to the low impedance, high current state. The dead time, τ_D varies with temperature and details of manufacture; it is usually in the range 10^{-1}–10^{-6} s. The recovery process associated with it is thermally activated, the activation energy being in the range 0·5–1·25 eV. Consequently at

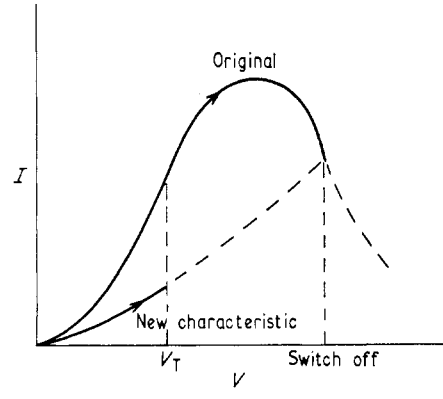

Figure 11. The high- and low-impedance memory states which can exist, up to a threshold voltage V_T, for a formed metal–oxide–metal structure.

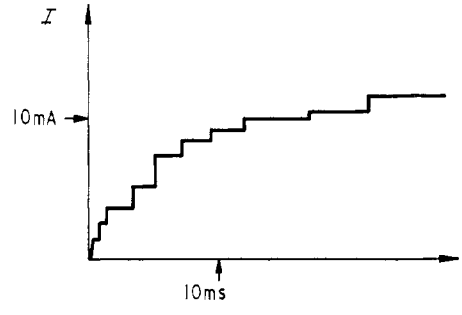

Figure 12. Step-like increase of current (I) as a function of time (t) in the switching of a formed SiO_x film (after Greene *et al.* 1969).

very low temperatures, it is very difficult to switch the device from its high impedance, low current state. The characteristic for $V < V_T$ is very little affected, even at temperatures down to a few degrees K (figure 13) and the changes seem to be due to changes in the populations of the electron states near the Fermi level of the electrodes.

As the frequency of the applied voltage is raised the hysteresis of the current–voltage characteristic increases. The peak voltage and current also change (figure 14). The differential negative resistance disappears, usually between $\nu = 100$ Hz and $\nu = 1000$ Hz, giving a steady-state conductance which depends on the voltage amplitude. This is shown in figure 15.

In table 5, we collect together observed values of V_{max}, V_{min} and V_T for a number of systems. Hickmott has observed two correlations:

$$V_{max} \propto (\text{energy gap in eV})^{1/2}\ \text{V}$$

$$V_{max} \simeq (10{\cdot}3 - 0{\cdot}18\ [\text{dielectric constant}])^{1/2}\ \text{V}.$$

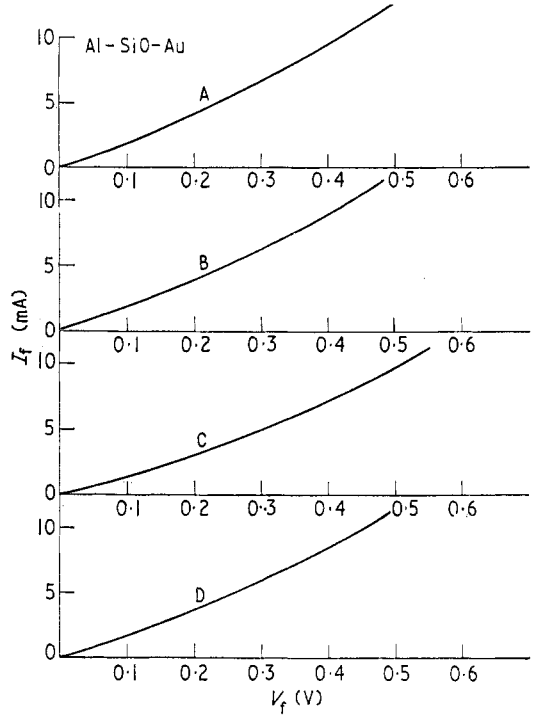

Figure 13. The low-voltage conductivity of an Al–SiO_x–Au device at temperatures of (A) pumped helium, (B) liquid helium, (C) liquid nitrogen and (D) room temperature. The oxide thickness is 400 Å (from Hickmott 1971).

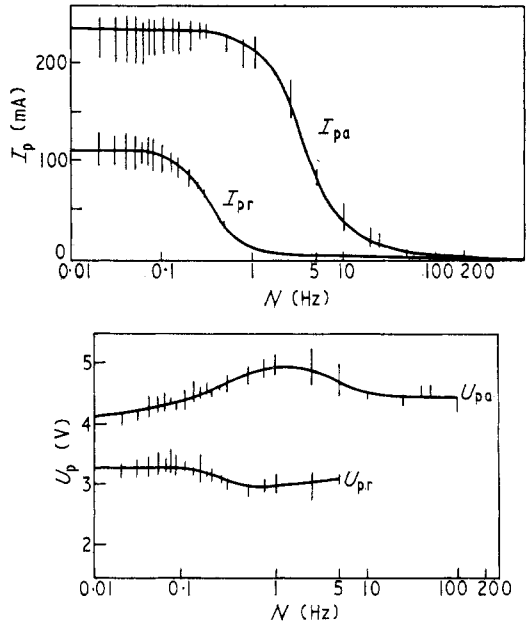

Figure 14. The variation in the maximum current, I_p, and the corresponding voltage, U_p, as a function of frequency, N, in a formed Ag–Al_2O_2–Al structure. The subscripts a and r denote increasing and decreasing current, respectively (from Barriac *et al.* 1969).

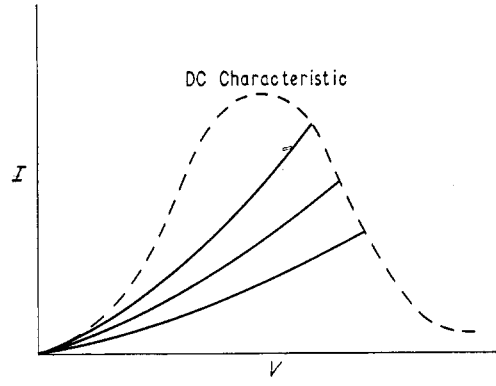

Figure 15. Broken line: the dc or low-frequency current–voltage (I–V) characteristics of a formed Au–SiO_x–Al device. Solid lines: the ac (1 kHz) characteristics of the same device for several voltage amplitudes (after Simmons and Verderber 1967 a).

Table 5.

Device	Source	V_T	V_{max}	V_{min}
Al–SiO–Au	H		3·1	8–9
	SV	2·9	4–5	8
Pb–SiO–Pb	H		3·0	7·2
Al–Al_2O_3–Au	H	3	2·9	7·5–8·5
Al–Al_2O_3–Ag	B		4	8
Ta–Ta_2O_5–Au	H		2·2	5–6
Zr–ZrO_2–Au	H		2·1	4–5
Ti–TiO_2–Au	H		1·7	3·5–4

H —Hickmott.
SV—Simmons and Verderber.
B —Barriac *et al.*

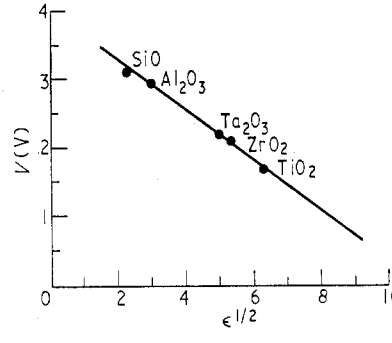

Figure 16. The dependence of the voltage (V) for maximum current flow through formed metal–oxide–metal structures on the dielectric constant (ϵ) of the oxide (after Hickmott 1971).

These correlations are not entirely independent, since the dielectric constant and bandgap are related by

$$\epsilon_\infty = 1 + \left(\frac{\text{constant}}{E_g}\right)^2.$$

The correlations are shown in figure 16 and table 6. The various voltages are

independent of insulator thickness. Hickmott has verified this, for example, for Al_2O_3 films in the range $d = 150$ to 1000 Å.

Table 6. Dependence of V_m on insulator band gap

Insulator	V_m(volts)	V_m^2(volts^2)	E_g(eV)
Al–SiO–Au	2·9–3·1	8·4–9·5	—
Al–Al_2O_3–Au	2·8–2·9	7·9–8·4	(8·4)
Ta–Ta_2O_3–Au	2·2	4·8	4·6
Zr–ZrO_2–Au	2·1	4·4	(4·3)
Ti–TiO_2–Au	1·7	2·9	3·0
Nb–Nb_2O_5–Au	1·7	2·9	3·4

3.2. *Theories of switching and memory in oxide films*

Our discussions of forming and of differential negative resistance showed that there were as many theories as there were groups working in the field. The same is true of the conduction phenomena. We review the models in turn. The results are summarized in table 7.

Hickmott's model of conduction invokes space-charge limited conduction in an impurity band near the middle of the bandgap. The differential negative resistance is produced by some mechanism (ionization of another set of impurity levels has been suggested) which depends on a high field and which reduces the number of impurity states which contribute to the conduction. The mechanism must be temperature independent and with a short time constant. The re-formation of the impurity centres as the voltage is decreased has a longer time constant, and it is this longer time constant which determines the range of frequencies over which negative resistance may be observed. The processes he discusses are shown in figure 17. Switching and memory effects are not discussed by Hickmott, but the natural extension of his ideas would suggest that these phenomena arise from redistributions of electrons over the various impurity and band states. However, such a model would be open to objections which apply to all models which do not involve atomic rearrangement—the memory state lasts for several years without significant decay, and it is hard to see that a nonequilibrium charge distribution could exist for so long in a thin-film device.

Simmons and Verderber also proposed a mechanism without atomic rearrangement. The differential negative resistance resulted from the top of the impurity band falling below the Fermi level of the injecting electrode, thus inhibiting hopping conduction at constant electron energy across the device. In our earlier description of this process we ignored band-bending from the presence of space charge for simplicity of exposition. It is essential to include these effects in their model of switching. Their description of band-bending assumes a broad distribution of electron traps with density $n(x, E)$ per unit energy range and unit volume, where x measures distance through the insulator. These states are not populated uniformly in space, and the local deviations from charge neutrality give the space charge of interest. We shall discuss the resulting potential when there is no bias voltage and when the Fermi level E_F lies within the impurity band. The local deviations from charge neutrality are

$$e\rho(x) = \int_{-\infty}^{E_F} \mathrm{d}E\, n(x, E) - (\text{space-averaged charge density}). \tag{3.1}$$

Table 7.

Author	Mechanism of differential negative resistance	Mechanism of switching	Predicted dependence of V_{max}	Non-ohmic behaviour at low voltages	Temperature dependence
Hickmott	Some process dependent on a high field which reduces the number of impurity centres which contribute to conduction, e.g. neutralization from hole centres	Not discussed	Fixed by relative positions of impurity levels. Observes relation to bulk properties (band-gap, dielectric constant)	Space-charge limited electron or hole current	Trapping is temperature dependent
Simmons–Verderber	Top of impurity band in the centre of the insulator falls below the anode Fermi level. Hence charge trapped in depletion layer near anode, and reduction in current since only electrons below the top of the impurity band contribute	When $V > V_{max}$ charge is injected into the depletion layer. Rapid reduction of V, in conduction with band bending, sends charge to the centre of the insulator, giving low conductivity. Slow reduction of V allows charge to trickle to electrode	eV_{max} is the energy difference between the top of the impurity band and electrode Fermi level. Suggests (i) V_{max} depends on electrode and (ii) impurity states should have sharp maximum energy	Tunnelling from electrode or within depletion layer.	Trapping of space charge is temperature-dependent
Barriac *et al.*	Injection of electrons immobilizes ions by trapping	Not discussed	Rapid dependence of V_{max} on temperature	Space-charge limited ionic current	Trapping of electrons by ions is temperature-dependent
Dearnaley	Rupture of conducting filaments by Joule heating	Rupture and regeneration of filaments	Depends on distribution of filament resistances; hence on host, but not strongly on electrodes	Not discussed; either tunnelling or space-charge limiting behaviour	Regeneration of filaments involves dielectric relaxation and ionic motion

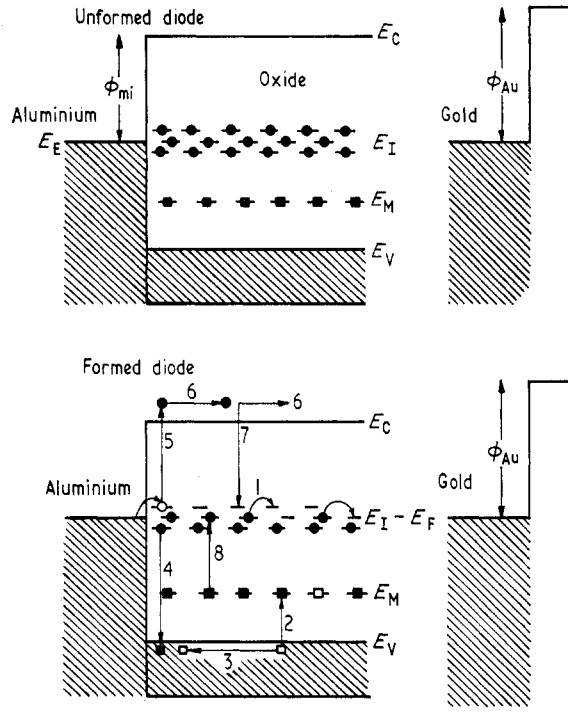

Figure 17. Schematic illustration of conduction processes in the high-field region of a metal–oxide–metal structure (after Hickmott 1964 b).

But the effect of the space charge is to produce a potential $V(x)$ which simply shifts the energy levels, and so displaces $n(x, E)$ in energy:

$$n(x, E) \rightarrow n(x, E + V(x)). \tag{3.2}$$

If we now assume $n(E)$ has a square distribution, that is $n(E) = N$ within the impurity band and is zero outside, then

$$e\rho(x) = eNV(x) + \text{constant terms}. \tag{3.3}$$

Substituting into Poisson's equation,

$$\frac{d^2 V}{dx^2} = \frac{eN}{\epsilon} V + \text{constant terms} \tag{3.4}$$

so that

$$V(x) = A \exp\left(\frac{x}{x_1}\right) + B \exp\left(-\frac{x}{x_1}\right) + \text{constant} \tag{3.5}$$

in which

$$x_1 = \left(\frac{\epsilon}{eN}\right)^{1/2}$$

and A and B are fixed by the boundary conditions. If the device is symmetrical and the barrier thickness $d \gg x$, then

$$V(x) \simeq A\left\{\exp\left(-\frac{x}{x_1}\right) + \exp\left(-\frac{d-x}{x_1}\right)\right\} + \text{constant terms}. \tag{3.6}$$

Thus the band structure is as shown in figures 18 and 19. The bands bend downwards because of the (attractive) defects in the insulator. Simmons and Verderber quote $N = 10^{19}\ \mathrm{cm}^{-3}$ (with an impurity band width of 3 V) which, with $\epsilon = 4$, gives $x_1 \simeq 50$ Å; this may be compared with the insulator thickness of several hundred Ångstroms.

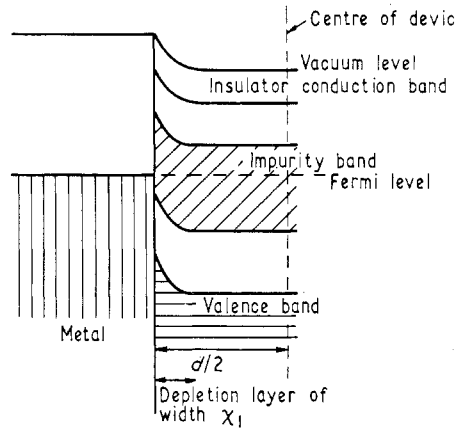

Figure 18. Band bending of the energy levels at a metal–insulator boundary when the applied voltage is zero and the Fermi level lies within the impurity band. Local deviations from charge neutrality give rise to a space charge. The bands bend downwards due to attractive defects in the insulator.

When the applied voltage V exceeds $V_{\max}$, electrons injected at energies near E_F hop between impurity states close in energy and are trapped in the depletion layer, since transitions to traps of very different energy are considered improbable. If the applied voltage is rapidly reduced, this charge does not return to the electrode from which it came, but moves by a thermally activated process to the centre of the insulator. This electronic space-charge reduces the conductivity of the device, as in the memory state. The memory state can, of course, be removed by removing the stored charge. The processes are shown in figures 20 and 21. The rate of switching is limited by the mean hopping time, τ_H from trap to trap. If τ_H is short compared to the time in which the applied voltage V is reduced from above $V_{\max}$, the space charge can trickle back to the nearest electrode; the space charge follows the field roughly adiabatically. However, if τ_H is long compared with the voltage reduction time the charge does not move significantly until after the reduction is complete, when the fields now favour a movement of charge to the centre of the insulator. The dead time is also τ_H. Since τ_H involves thermal activation, we can also understand the temperature dependence of these processes. Simmons and Verderber postulate that the normal (high current) behaviour, which also involves hopping, does not show thermal activation because transitions between sites with the same energy (that is, with zero activation energy) are dominant. This hypothesis does not seem to be consistent with the general arguments of Mott (Mott 1967, especially p. 71) who argues that

$$\tau_H^{-1} \simeq \omega_H(\boldsymbol{r}) \exp\left(-\frac{\Delta W + \frac{1}{2}W_p}{kT}\right) \tag{3.7}$$

in which $\omega_H(\boldsymbol{r})$ is a 'tunnel factor' depending on the separation of the two sites

involved. Even when $\Delta W = 0$, corresponding to two equivalent sites, thermal activation is needed; W_p appears because the polarization field moves with the electron. This contribution only vanishes above a critical concentration, when metallic behaviour appears. Further, Mott (1968 a, b) observes τ_H^{-1} should be frequency dependent ($\tau_H^{-1} \propto \nu^2$) and the thin film devices show no evidence of this.

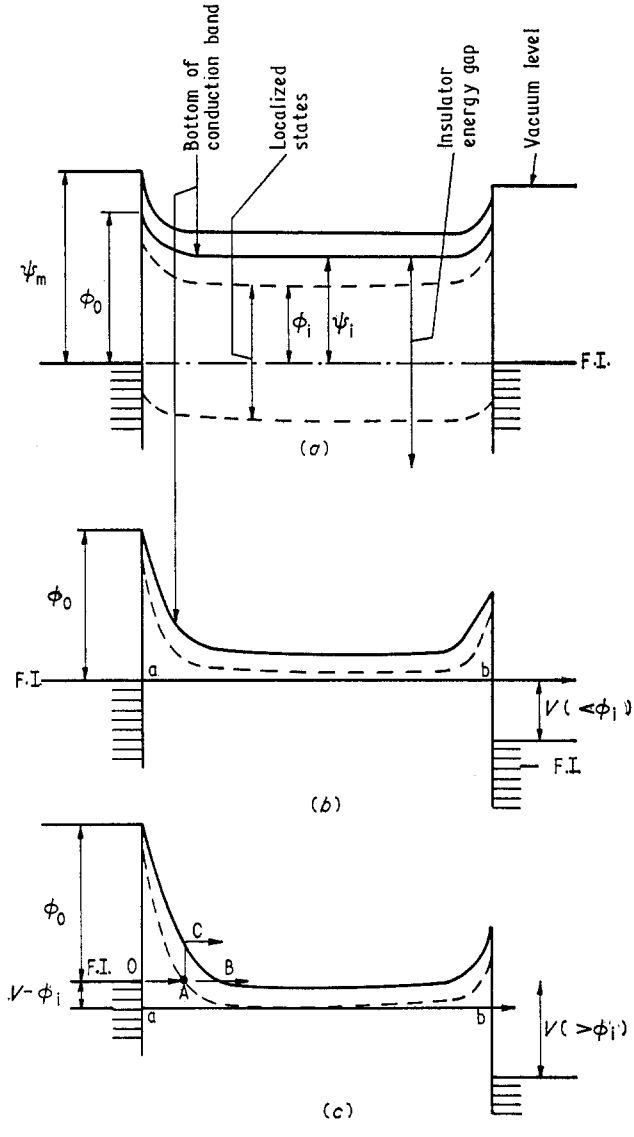

Figure 19. Energy diagram of formed metal–insulator–metal sample for various voltages: (*a*) $V = 0$, (*b*) $V < \phi_i$, (*c*) $V > \phi_i$ (after Simmons and Verderber 1967 a).

Dearnaley's model of conduction invokes the formation of filaments in the insulator. These filaments exhibit ohmic conduction, without thermal activation, perhaps because the sites involved are considered to be very close together so that hopping without activation energy may occur (Mott 1967). The filaments are further characterized by two important features. The first is that the filament may

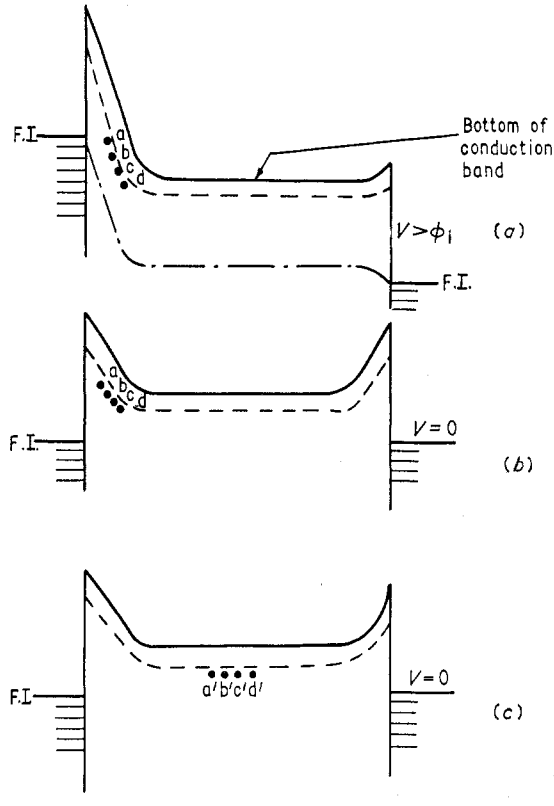

Figure 20. Energy diagram showing the position of the stored charge relative to the Fermi level: (*a*) at a voltage bias of V volts, (*b*) immediately after voltage has been rapidly reduced to zero, (*c*) after a longer time interval (after Simmons and Verderber 1967 a).

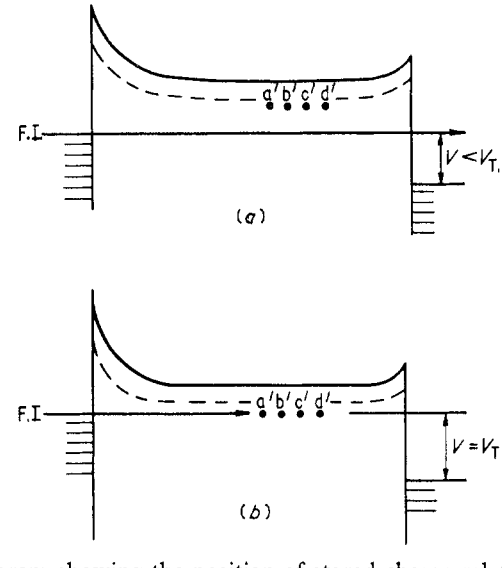

Figure 21. Energy diagram showing the position of stored charge relative to the Fermi level: (*a*) before switching occurs, (*b*) at onset of switching (after Simmons and Verderber 1967 a).

fracture and cease to contribute to the conduction; the second is that the filament can re-form under suitable conditions. Detailed atomic models of these processes are still speculative, but it is clear that the re-forming will be related to the basic forming process of propagation of the filament through the insulator. The fracture of the filament is probably a consequence of Joule heating, raising the temperature of a part (at least) of the filament so that there is local atomic reorganization, as in melting or other phase changes. In addition, the phonon scattering of electrons also increases as the filament temperature rises. We shall see that the fracture is probably localized in some small region of the filament, since then the switching is voltage, rather than field-dependent. The localization of the fracture results from the nonuniformity of the filaments, which are postulated to have weak spots, a consequence of their growth and the amorphous nature of the insulator.

We now consider the current–voltage characteristics of an array of conducting filaments embedded in an insulating matrix. Each filament will be assumed to exhibit ohmic conduction and we will ignore hopping conduction, tunnelling and space-charge effects which probably play a role; by ignoring these effects we shall be unable to predict the nonlinear dependence of current upon voltage at very small voltages. The novel current–voltage characteristics occur because filaments can rupture.

Once each filament bridges the gap between the electrodes, electronic conduction begins and Joule heating will occur. We then associate a local lattice temperature T with each filament and assume that rupture will occur if T exceeds $T_{\max}$. The filaments can re-form subsequently, but this will depend on their temperature, the ambient temperature, the electric field and possible space-charge effects; the precise conditions will not be specified in detail. The filament temperature, T is not a thermodynamically well-defined quantity. It will only be used as a measure of the local heating in the filament. Physically both T and the resistance ρ of each filament can best be imagined by associating them with local regions of high resistance on a highly conducting filament or chain of defects. Two terms contribute to changes in T, one due to the usual Joule heating and the other to heat loss to the insulator, which we assume simply to be proportional to the temperature difference between the filament and its environment. Thus the local temperature is given by

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \alpha\frac{V^2(t)}{\rho} - \frac{1}{\tau_c}(T - T_0) \tag{3.8}$$

where α measures the heating and τ_c characterizes the cooling.

We first consider the static characteristic. From the steady-state solutions of (3.8) we see that, at voltage V, only filaments with resistance greater than

$$\rho_{\min} = \frac{\tau_c \alpha V^2}{T_{\max} - T_0} \tag{3.9}$$

remain unfractured. The total current is thus

$$I = \int_{\rho_{\min}}^{\infty} \mathrm{d}\rho \frac{V}{\rho} P(\rho) \tag{3.10}$$

where $P(\rho)$ gives the probability distribution of the various resistances, and

$$\int_0^{\infty} \mathrm{d}\rho P(\rho) = 1. \tag{3.11}$$

We have assumed implicitly that the filaments are either permanently broken or permanently unbroken at the given voltage, and that an equilibrium has been achieved so that the previous history of the specimen can be ignored. Clearly the detailed shape of the characteristic is determined by the distribution $P(\rho)$. Two features of $P(\rho)$ are immediately apparent. First, there are very few filaments with resistance below

$$\rho_{\mathrm{T}} = \frac{\tau_{\mathrm{c}} \alpha V_{\mathrm{T}}^2}{T_{\max} - T_0} \tag{3.12}$$

since the devices can be cycled below V_{T} without showing signs of filaments breaking.

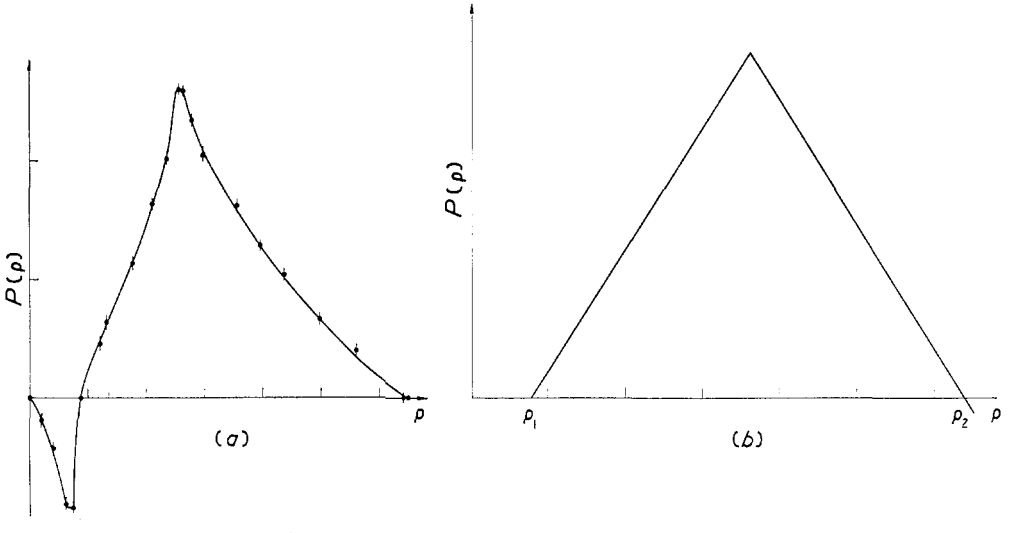

Figure 22. (*a*) The calculated distribution of filament resistances, ρ, deduced from experimental current–voltage behaviour. Negative values near the origin arise because tunnelling at the electrodes has been neglected. (*b*) The simple triangular distribution of filament assumed as an approximation in fitting the current–voltage characteristic (after Dearnaley *et al.* 1970).

Secondly, there are very few filaments with resistances above

$$\rho_{\mathrm{u}} = \frac{\tau_{\mathrm{c}} \alpha V_{\mathrm{u}}^2}{T_{\max} - T_0} \tag{3.13}$$

where V_{u} is the voltage at which $I(V)$ is a minimum. Above V_{u} the conductivity is small, and the filamentary contribution cannot be appreciable. We can derive $P(\rho)$ from $I(V)$ by noting

$$\frac{P(\rho)}{\rho} \propto \frac{\mathrm{d}(I/V)}{\mathrm{d}(V^2)} \tag{3.14}$$

this distribution is shown in figure 22. The apparently negative value of $P(\rho)$ for small ρ simply reflects the contribution of tunnelling or space-charge effects at small voltages. The absence of very low-resistance filaments is just a consequence of finite specific resistance; the absence of very high-resistance filaments presumably reflects a natural selection mechanism in the forming process; that is to say that a filament will cease to propagate rather than continue as high-resistance path, if conditions (such as the local structure of the insulator) are unfavourable.

The distribution of resistances shown in figure 22 is merely qualitative unless we can estimate N, the number of filaments per unit area. We now make a rough estimate of N assuming that all filaments have the same resistance, $\bar{\rho}$. The initial resistance of the device per unit area is $\bar{\rho}/N$, apart from small tunnelling effects, and the filaments all burn out when $V > V_{\max}$. Thus, from equation (3.8),

$$\bar{\rho} \simeq \frac{\tau_c \alpha V_{\max}^2}{T_{\max} - T_0}. \tag{3.15}$$

The cooling time, τ_c, is estimated by regarding each filament as a cylinder of radius R and temperature T cooling to a cylindrical sink radius D and temperature T_0; D is taken to be the mean filament separation defined by $ND^2 = 1$. It is then easy to show that

$$\tau_c \alpha = \frac{\ln(D/R)}{4{\cdot}18(2\pi KL)} \tag{3.16}$$

where K is the thermal conductivity of the insulator and L is the length of the filaments. Note that the choice of D and R is not very critical in this equation. Combining (3.15) and (3.16) we may find $\bar{\rho}$ and N. Reasonable values of most of the parameters are easily introduced, e.g. an initial resistance $\bar{\rho}/N$ of 5 Ω for a filament length of 400 Å, a thermal conductivity of $0{\cdot}002\ \mathrm{cal\,cm^{-2}\,s^{-1}\,K^{-1}}$ and $(T_{\max} - T_0)$ of 1000 K. It is harder to estimate R, but our earlier discussion of the dependence of forming voltage upon film thickness suggests $R \simeq 10$ Å. Using these values we find $\bar{\rho}$, the mean resistance per filament, to be about $2{\cdot}5 \times 10^7$ Ω and N to be 5×10^6 filaments per cm². These lead to $D \simeq 2{\cdot}5$ μm and a specific resistance of about 0·2 ohm cm for the material of the filament. We emphasize that these calculations give only order of magnitude estimates since, for example, end effects have been neglected in deriving τ_c. However, the results appear to be both plausible and consistent. We have since learned that microphotographs of formed silicon monoxide films (R. Hall, private communication) do in fact reveal about 10^6 minute holes per square cm in close agreement with our estimate.

We now turn to consider the dynamic characteristics of formed devices, in much the same manner as the dc behaviour was treated. If the applied voltage has amplitude V_0

$$V = V_0 \sin \omega t \tag{3.17}$$

then after a time of order τ_c when the transients have disappeared, the filaments with resistance greater than

$$\rho_{\min}(\omega, V_0) = \rho_{\min}(0, V_0)\, \phi(\omega, \tau_c) \tag{3.18}$$

remain unfractured. Here $\rho_{\min}(0, V_0)$ is equivalent to (3.9), and

$$\phi(\omega, \tau_c) = \tfrac{1}{2}[1 + \{1 + (2\tau_c)^2\}^{-1/2}]. \tag{3.19}$$

At higher frequencies more filaments survive. The current is given by

$$I = \int_0^\infty \mathrm{d}\rho \frac{V}{\rho} P(\rho)\, f(\rho, \omega) \tag{3.20}$$

where $f(\rho, \omega)$ is the fraction of the time for which a filament of resistance ρ operates at frequency ω. Clearly $f(\rho, \omega)$ is unity if ρ is greater than $\rho_{\min}(\omega)$. The dynamic characteristic differs from the static characteristic in that the filaments which are broken re-form only after a number of cycles, and the conduction is dominated by

those which remain unbroken throughout. There is no negative resistance régime in the dynamic case. As long as ω is sufficiently large for the re-forming of a ruptured filament to be unlikely during the cycle, we may treat $f(\rho, \omega)$ as zero for ρ less than $\rho_{\min}(\omega)$. Other cases can, and do, arise at lower frequencies. We may relate the dynamic and static characteristics and, using a simplified triangular distribution of resistances (figure 22(*b*)) with $\phi(\omega, \tau_c) = 1$, we obtain

$$I = \frac{2V}{\rho_T + \rho_u}\left(\frac{1-(V_0/V_u)^2}{1-(V_T/V_u)^2}\right). \tag{3.21}$$

These ac characteristics are shown in figure 23 (these are straight lines since we

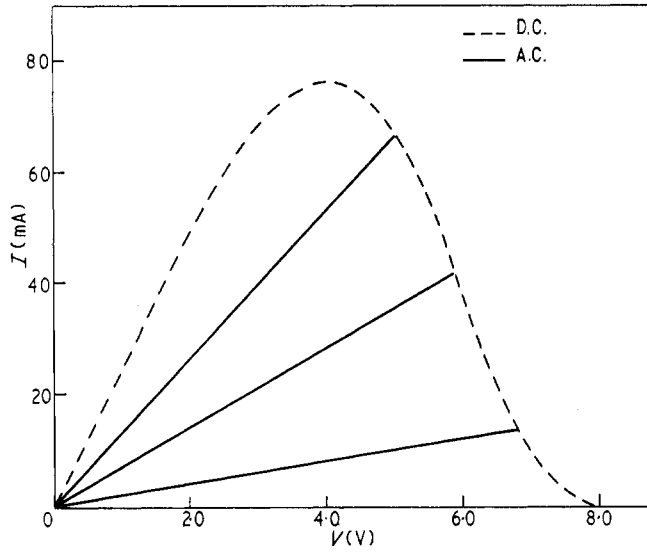

Figure 23. The ac (1 kHz) current–voltage (I–V) characteristic calculated using the filamentary model, assuming a simple triangular distribution of resistances (figure 22(*b*)), and after the transient period has elapsed.

have again ignored the effects of space-charge and tunnelling). Further, we find that for V_0 less than $V_T/\phi^{1/2}(\omega, \tau_c)$ then no filaments should break, and the initial slopes of the static and dynamic characteristics should be the same. At 1000 Hz it seems that $\phi(\omega, \tau_c)$ is close to unity, since the fracture of filaments occurs at very similar voltages in the two cases. Using $\phi(\omega, \tau) \simeq 1$ we find that, for V_0 greater than V_T, the dynamic and static characteristics should meet at $I(V_0)$, that is

$$I(V)_{\text{dynamic}} = \frac{V}{V_0} I(V)_{\text{static}}. \tag{3.22}$$

This has been observed in SiO by Simmons and Verderber (1967 a, b) and it is illustrated in figure 23. Their static characteristics were measured at 50 Hz. From the observations at two frequencies we can deduce that τ_c is shorter than 1 ms and the mean re-forming times of filaments are typically longer than 1 ms but shorter than 20 ms at room temperature. Re-forming appears to be thermally activated, as it occurs slowly, if at all, at low temperatures. As re-forming does not occur for an applied voltage (rather than an applied field) below V_T it is likely that both the

fracture and the re-forming only involve a well-localized region of space. After fracture the filament must be in one of several distinct states. The dead time, discussed earlier, is a measure of the time between the filament breaking and reaching a state from which it may re-form. This transition appears to be thermally activated.

One possible microscopic model which predicts a dead time is the assumption that, during the rupture of the filament at a weak spot, electrons are scattered into the surrounding matrix. The insulator will be highly polarized locally and will return to equilibrium by thermally-activated processes, such as Poole–Frenkel emission from traps. At low temperatures this 'polarized' state will exist for a long time, and re-forming will not occur. At higher temperatures the system may relax rapidly to a state in which the filament is broken but the polarization is absent. The filament may then regrow by ionic migration in the electric field. The relaxation will take a finite time and hence we can explain the transient phenomena which appear at frequencies between 50 and 1000 Hz. Furthermore, filaments which have been broken remain nonconducting for almost the whole of the time since the time constant for the heating process is short compared with the time to re-establish conduction. In this way we can understand the relatively long switching time from the high to the low impedance state (of the order of 10^{-7} s) while switching back to the high impedance state is rapid (of the order of 10^{-9} s). It is likely that the difference occurs because in the former case ionic migration must take place, while in the latter the controlling mechanism is the scattering of 'hot' electrons leading to polarization of the dielectric. Likewise, we can understand why the 'dead time' of many milliseconds must elapse before the cycle can be repeated—this corresponds to the relaxation of space-charge polarization.

Barriac *et al.* (1968, 1969) argue that the differential negative resistance in dc characteristics results from a combination of ionic and electronic conduction. They do not discuss switching.

Their analysis of the dc characteristic divides the voltage into four regions—regions A and B below the maximum current, and regions C and D when there is differential negative resistance. Ionic conduction alone is important below the voltage for maximum current, $V_{\max}$. In this regime

$$I = A \exp(\alpha V) \tag{3.23}$$

where $A \simeq \lambda_s \exp(-U_s/kT)$, U_s being an activation energy and λ_s the diffusion length (the diffusion coefficient is $D = \lambda_s^2 \nu \exp(-U_s/kT)$, λ_s being the site separation multiplied by an entropy factor). The parameter α depends on temperature and the thickness of the depletion layer

$$\alpha = \left(\frac{\lambda_s}{\delta}\right)\frac{q}{kT} \tag{3.24}$$

where q is the ionic charge. The dependence on the potential V is complicated because δ is a function of V. At the lowest voltages (region A), δ changes rapidly with V, saturating to a value Δ independent of V to give region B. This is illustrated in figure 24. Thus below $V_{\max}$ the current is a space-charge limited ionic current. Once V reaches $V_{\max}$ electrons begin to tunnel into the insulator through the electrode barriers. The electrons in turn are trapped within the insulator to form immobile defects, thus reducing the electronic and ionic currents. The reduction of the current (basically the loss of ionic current, assuming all electrons are trapped)

is then

$$I_{\text{imp}} = -BV^2 \exp\left(\frac{-\beta}{V}\right) \tag{3.25}$$

which is the usual Fowler–Nordheim expression. B and β both depend on $\delta(V)$ and also on the height of the metal–insulator barrier. As the ionic space-charge becomes increasingly neutralized the depletion layer starts to increase in thickness from Δ once more, giving the transition to region D from region C.

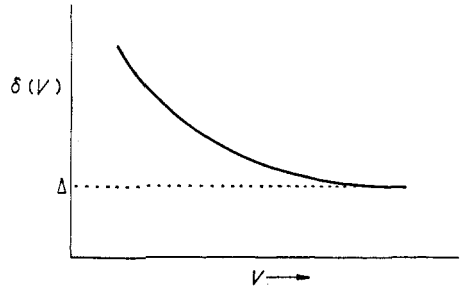

Figure 24. The dependence of the parameter δ on the potential V, in the theory of Barriac *et al.* (1969). At low voltages δ changes rapidly with V, but saturates to a value Δ.

The traps are assumed to have a mean trapping time τ which is temperature-dependent, and the neutral units of an electron and ion are assumed unstable in the absence of a field. The hysteresis seen in the characteristics at low frequencies is a direct result of trapping. Finally, the effect of oxygen is to give a permanent neutralization of the ions which give rise to the space charge, so explaining the sensitivity to the atmosphere.

One prediction of this model is the value of the voltage of maximum current. The current is the sum of two terms:

$$I = A \exp(\alpha V) - BV^2 \exp\left(\frac{-\beta}{V}\right) \tag{3.26}$$

and $V_{\max}$ is readily found from the zero of $(\partial I/\partial V)$. The result leads to a major disagreement with experiment. Both A and α depend on temperature, so that $V_{\max}$ is sensitive to the temperature. But $V_{\max}$ is observed to be essentially temperature-independent from helium temperatures to above room temperature. Thus both on this point and on the earlier objections to models invoking ion injection it seems unlikely that the Barriac–Pinard–Davoine mechanism is correct.

Greene *et al.* do not discuss the switching or conduction mechanisms in detail, but suggest they involve trapped space charge produced when the applied voltage is sufficient to give some electrons enough energy to leave the conducting path and enter the surrounding unformed dielectric.

3.3. *Observations on triodes*

The experimental data discussed so far all refer to diodes, that is to two-electrode metal–insulator–metal devices. Hickmott (1964 a, b) has made an important set of experiments on triodes, or three-electrode devices. The structure he used is shown in figure 25.

Before forming, the electrical behaviour of the triodes is orthodox, and the potentials applied between the plate and cathode divide as expected between the plate-grid and grid-cathode regions. However, after forming the potential distributions are highly nonlinear (figure 26). Nearly all the potential drop occurs at the negative electrode, between grid and cathode. When the voltage is reversed the potential drop remains in the same region, although the high conductivity and differential negative resistance are still observed. The high-field region can be shifted to between the plate and the grid; this does not affect the plate-cathode characteristics greatly provided that a high-field region remains somewhere in the device.

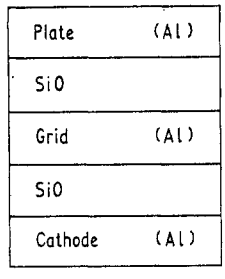

Figure 25. The metal–oxide–metal–oxide–metal triode structure.

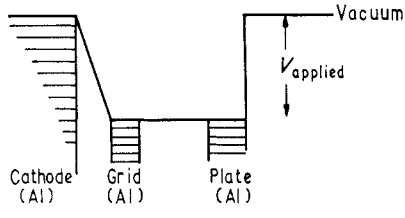

Figure 26. Schematic representation of the potential distribution in oxide triodes after development of plate-cathode conductivity (after Hickmott 1964 b).

Thus the high conductivity is *not* a consequence of field emission at the cathode from this high-field region. The potential in the low-field region is linear in the current in that region; that is, conduction is ohmic.

One remarkable feature is that the development of conductivity between plate and cathode (the forming) does not necessarily result in conductivity to the grid. Grid conductivity may be developed separately or it may develop during the establishment of plate conductivity. In one device, for example

$$R_{pg} = V_{pg}/I_{pg} > 2.10^5 \text{ ohms} \quad \text{and} \quad R_{gc} = V_{gc}/I_{gc} > 2.10^5 \text{ ohms}$$

whereas when the voltage is applied from cathode to plate $R'_{pg} = V_{pg}/I_{pc} \simeq 2$ ohms. We know of no satisfactory explanation of these phenomena.

4. Electron emission, luminescence and other phenomena

We now consider some of the effects of forming other than those on the current–voltage characteristic. In particular, formed devices are rather efficient electron emitters, the emitted current being up to 1% of the current in the device in some cases. Other related phenomena are electroluminescence and the noise in the device.

4.1. *Electron emission*

The electron emission from a formed device depends strongly on the voltage. At voltages below V_{max} there is very little emission, but a sharp rise occurs at $V \simeq V_{max}$. As the voltage is further increased the emitted current either rises slowly, as with the Al–SiO–Au device (see figure 27) or decreases, as with the Al–Al_2O_3–(Al

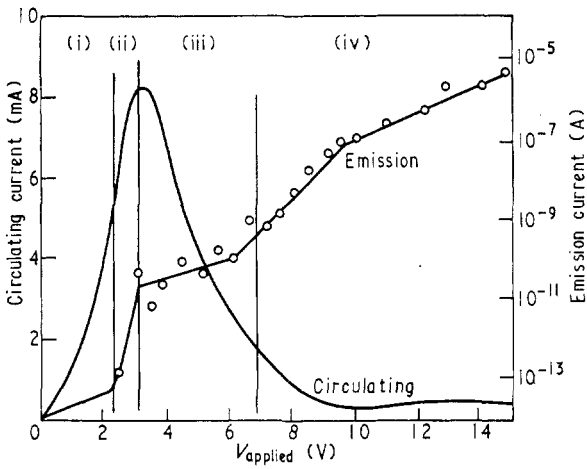

Figure 27. Curves showing the circulating current and emission current as a function of applied voltage in a formed Au–SiO_x–Al device (after Simmons and Verderber 1967 b).

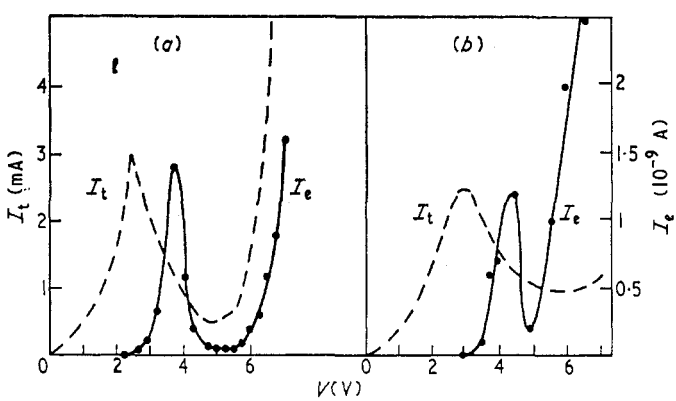

Figure 28. The circulating current (I_t) and emission current (I_e) as a function of applied voltage for (*a*) a Au–Al_2O_3–Al device and (*b*) an Al–Al_2O_3–Al device (after Barriac *et al.* 1969).

or Au) device (see figure 28). The emitted current then starts to increase again at a voltage near to the work function of the electrode through which emission occurs. The detailed behaviour depends both on the nature of the insulator (Hickmott 1963) and, to some extent, on its thickness (Hickmott 1965 a). The emitted current may also show slight structure when eV is equal to the insulator bandgap; this appears in Al_2O_3 and Ta_2O_5.

The emission is not uniform over the device and appears to come from small localized regions. There is some correspondence between the luminescent regions we discuss later, but there is not a clear one-to-one correspondence between the two.

The emission is also temperature-dependent; at lower temperatures the fraction of current through the device which is emitted decreases, the decrease being greater at lower voltages.

The electrode through which emission occurs also affects the observed external current. In the simplest cases the current is simply attenuated; the attenuation length in gold being typically between 200 and 300 Å, without strong dependence on the applied voltage. However, more subtle effects may be seen, particularly at lower voltages. These include the diffraction of electrons by pinholes in the electrode, which affects the spatial pattern of the emission (Verderber and Simmons 1966, Simmons and Verderber 1967 a, b). One of the reasons for the emission efficiency of these oxide devices as opposed to thinner films (10–50 Å) which has been suggested is that higher voltages can be applied to thicker films without breakdown, and these higher voltages give emission with less attenuation by diffraction.

The important points in interpreting the emission data are these. First, a significant fraction of the emitted electrons have energies close to eV, where V is the applied voltage. Second, much of the emission is localized at small regions in the electrode. Third, the emission is not simply a result of the insulator becoming hot; Verderber and Simmons eliminate this possibility by observing that the electron emission does not follow the power dissipation in the device.

A number of phenomena which occur in conventional oxide-coated cathodes may be related to the effects we have been discussing. Thus the electron emission of such a cathode is low prior to 'activation', which consists usually of the application of an electric field to the cathode under conditions in which some of the oxide may be reduced. Oxygen is released during this process. The electron emission is seen (Hanney *et al.* 1949) to follow the growth of electrical conductivity through the oxide. Heinze and Wagener (1939) studied the distribution of the emission from an oxide cathode at successive stages of activation. They formed electron images of the cathode and thereby observed that emission takes place from small bright centres. Initially only a few centres were present but as activation proceeded the number of sites increased very greatly but their size varied little. Later Mecklenburg (1943) used projection-tube microphotography to demonstrate that these emission centres are less than 400 Å in diameter and are distributed in a pattern related to the grain size in the oxide. It appeared that the emission centres were grouped along grain boundaries in the microcrystalline oxide of the cathode. A characteristic of the oxide-coated cathode is the phenomenon of 'flicker' noise. This is a low-frequency electrical noise with a power spectrum showing a $1/f^{\alpha}$ frequency dependence, where α is about 1·0–1·5. This noise is a minimum if the cathode is well activated.

4.2. *Electroluminescence in* M–I–M (*metal–insulator–metal*) *structures*

In this section, we consider electroluminescence which occurs inside the insulating oxide as opposed to the electroluminescence in a semiconducting substrate (the M–I–S system of § 4.4). Oxide films that develop voltage-controlled negative resistance (Al_2O_3, SiO, Ta_2O_5 and ZrO_2) have many features in common. Prior to forming, low-voltage electroluminescence in the near infrared has been reported in Ta–Ta_2O_5–Au (Hickmott 1966 a, b). However, visible electroluminescence develops simultaneously with the development of voltage-controlled negative resistance.

In the formed oxide, the electroluminescence parallels electron emission to some extent. The luminescence first appears at voltages just below V_{max}, and falls again in the differential negative resistance régime, before starting a slow rise in intensity with applied voltage. The light, which covers a wide spectral range, is emitted from a few bright spots; these spots are usually stable under repeated tracings of the current–voltage characteristic, although a few spots appear or disappear. The spots are typically a few microns across and there is no conclusive evidence that they are the spots from which electron emission occurs. The electroluminescence and its spectrum are shown in figure 29. Hickmott has analysed the spectra from Al–Al_2O_3–Au and Ta–Ta_2O_5–Au structures; one of the conspicuous peaks at about 4 eV for Al_2O_3 could be related to the strong peak, extending from 3·5 to 6·0 eV, observed in x ray induced luminescence (Runciman 1968) and attributed to iron impurity. However, atomic aluminium in its various charge states also has a number of bands in this range, and these have probably been detected in the breakdown of Al–SiO–Al (Budenstein *et al.* 1969).

Electroluminescence is not confined to oxide films which develop voltage-controlled negative resistance. In anodically grown Zr–ZrO_2–Al diodes, for example, Mochniak (1966) has observed electroluminescence where the integral electroluminescence intensity increases with the current density flowing through the oxide. The light appears for both the forward and reverse bias. Hickmott (1966 a, b, 1969 a, b) has studied in detail the electroluminescence in Nb_2O_5 diodes which exhibit current-controlled negative resistance. The diodes are electroluminescent before forming; after the forming process, increased electroluminescence accompanies switching and emitted light is both polarity- and voltage-dependent. Microscopic observations indicate that the forming and switching occur at small conducting regions.

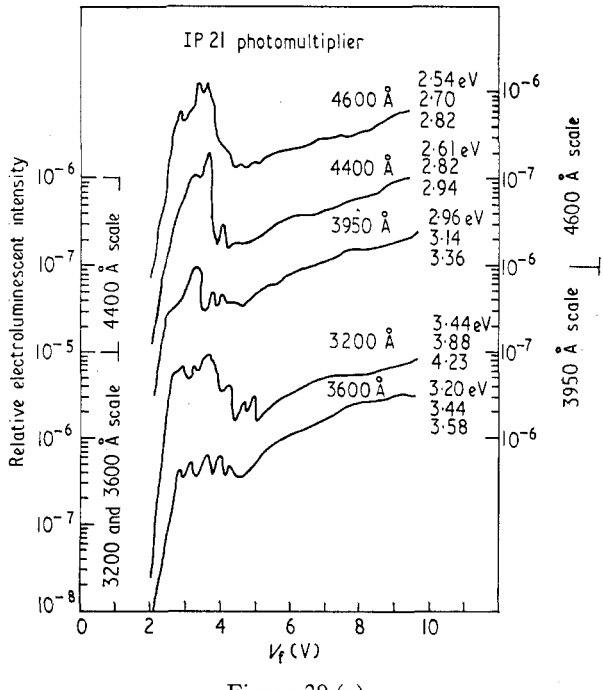

Figure 29 (*a*)

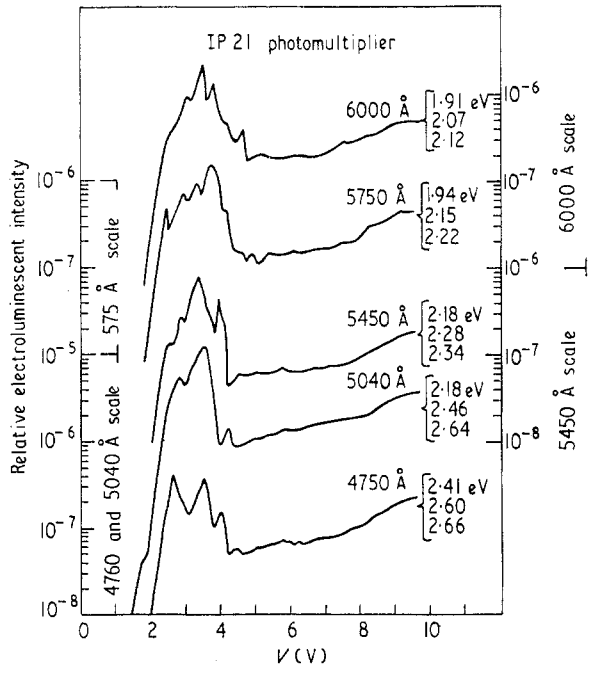

Figure 29 (*b*)

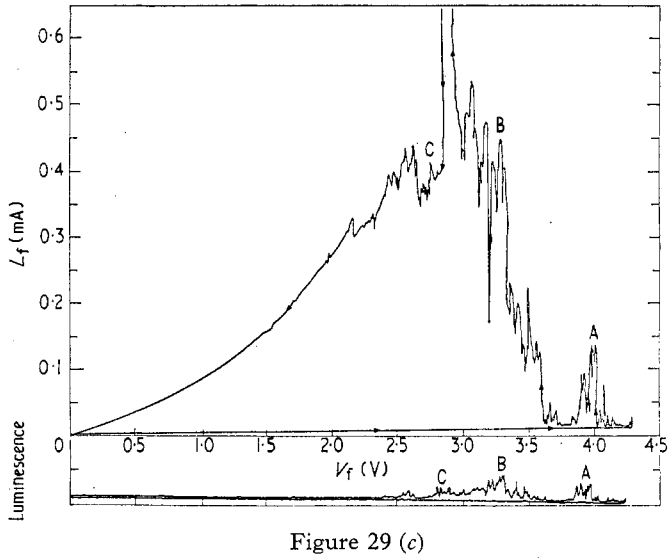

Figure 29 (*c*)

Figure 29. (*a*) and (*b*): the electroluminescent intensity at different wavelengths as a function of the voltage applied to a Au–Al_2O_3–Al structure. The oxide thickness was 550 Å, the gold thickness 200 Å (after Hickmott). (*c*) Pen-recorder plot of the development of conductivity and of electroluminescence in an Au–Al_2O_3–Al stucture. Current bursts at A, B, C coincide with flashes of light (after Hickmott 1971).

4.3. *Theories of electron emission and electroluminescence*

In both the Hickmott and the Simmons–Verderber theories electrons are excited by some means into the conduction band, where they move through the insulator without attenuation. Emission can then occur at energies up to the sum of the energy received from the applied field and the energy received, E_x from the excitation mechanism. The mechanisms suggested by Hickmott, although tentative, include the impact ionization of levels in the insulator, and some kind of recombination process. The recombination process may be the same as that giving luminescence, but this is by no means certain. Simmons and Verderber, on the other hand, suggest that the excitation mechanism is thermal.

The voltage at which emission should start is given by

$$eV + E_x > \chi = \text{electrode work function.}$$

If E_x does not depend on V (the Simmons and Verderber case) then emission can occur for a voltage bias *less* than the metal work function:

$$eV > (\chi - E_x) < \chi.$$

If the excitation does depend on V, as in recombination processes (for example, one model discussed by Hickmott involves the field ionization of impurity levels) then the minimum voltage is fixed by the efficiency of the mechanism as a function of voltage.

Dearnaley has proposed a different model. Electron emission is due to the generation of hot carriers by the high fields which exist at high-resistance spots in the filaments. In some of these spots, the voltage drop may be very nearly the whole bias voltage. If the 'weak' spot is close to the anode, electrons may be accelerated to nearly the full applied potential difference. This approach suggests emission should become particularly strong near $V = V_T$, when the filaments begin to rupture, or at any rate at some voltage between V_T and V_{max}, the voltage corresponding to maximum current. Hickmott (1965 a, b, 1970 private communication) has observed in Al_2O_3 films that electron emission commences at about 2·3 V while $V_T = 1{\cdot}8$ V and $V_{max} = 2{\cdot}8$ V in this material. Further, since high-energy electrons (with energies of the order of $100kT$) are being scattered into the insulator, it is possible to understand luminescence at spots over the device.

Hickmott has also discussed the luminescence, and has proposed models based on recombination involving the impurity levels in the insulator. The origin of the decrease in the luminescence in the differential negative resistance régime is by no means certain, and may be a population effect (a decrease in the number of electrons capable of radiative transitions or a decrease in the number of sites which can receive the electron), or an effect on the transition probability itself.

Table 8.

Author	Electron emission	Electroluminescence
Hickmott	Excitation of electrons to conduction band; energy supply by recombination or impact ionization	Recombination of electrons with impurity levels.
Simmons–Verderber	Thermal excitation of electrons from top of impurity band to conduction band	
Dearnaley	Acceleration of electrons in high field at weak spots of filaments	Scatter of fast electrons into insulator from filaments

A model for the oxide-coated cathode analogues to the one developed to account for the behaviour of amorphous oxide films has been proposed by Dearnaley (1969). He suggests that the electronic conduction is developed during activation by filamentary propagation through the oxide, assisted by reducing conditions and temperature. The filaments are believed in this case to be a defect state, possibly a chain of alternate cations and anion vacancies, or surface F-centres, clustered along grain boundaries of the material (see also § 5.3). In this way the small emission centres are satisfactorily explained, together with the mechanism of 'poisoning' of the cathode by oxygen. The high thermionic emission is explained by the field emission from such highly localized 'patches' coupled with the local Joule heating which is a consequence of conduction through filaments.

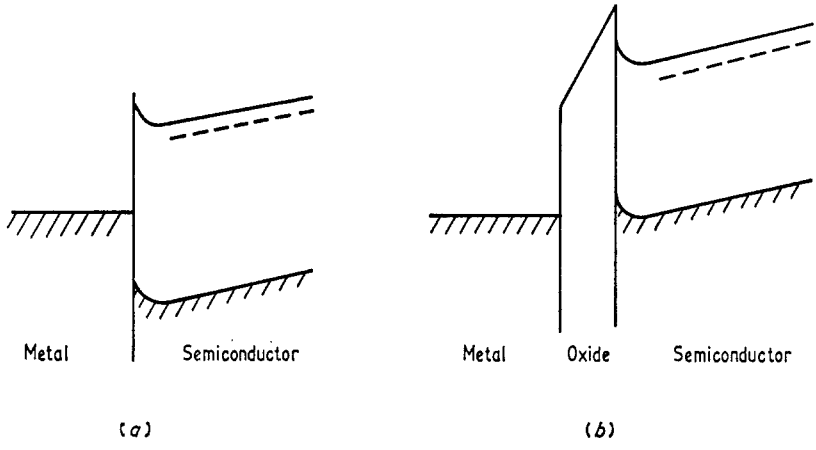

Figure 30. The electronic energy levels in (*a*) a metal–semiconductor and (*b*) a metal–oxide–semiconductor contact (after Jaklevic *et al.* 1963), illustrating the greater ease of tunnelling of holes brought about by the electric field in the oxide.

4.4. *Injection electroluminescence in* MOS (*metal–oxide–semiconductors*) *structures*

The injection of carriers into suitable wide bandgap semiconductors has often been studied as a potentially efficient means of generating electroluminescence. One method which has been used in studying this has been to deposit a thin film of SiO or CaF_2 on to a single crystal of semiconductor. Thus Jaklevic *et al.* (1963) deposited 50–150 Å layers of SiO_x on CdS and after forming of the oxide film, observed green electroluminescence at 77 K. If the oxide layer was omitted, no luminescence was seen. It was noticed that the light originated at numerous tiny spots, distributed over the surface in a random manner, the smallest being less than 10 μm in diameter. These authors attributed the behaviour to hole injection by tunnelling through certain localized regions, the holes entering the valence band as indicated in figure 30. Some observations by other authors cannot, however, easily be attributed to tunnelling, even assuming the existence of 'pinholes'.

Fischer and Moss (1963) made a series of similar experiments with a range of materials, including CaF_2, MgO, ZnF_2, CdF_2, on semiconductors. The layers on this case were as much as 500 Å in thickness, yet tunnelling of carriers through the film was still proposed as the injection mechanism.

Yee and Condas (1968) observed electroluminescence in MOS structures on CdS. An evaporated layer of SiO_x several hundred ångströms thick was employed, with

a gold layer deposited on top. High current densities (of the order of 30 A cm^{-2}) were achieved in the dc mode, and the emission spectrum, while varying with injection current, showed several features characteristic of CdS emission. The carrier injection mechanism was not identified.

Ralph and Tansley (1970) have examined the behaviour of Au–CaF_2–CdS and Tansley and Ralph (1970) Au–SiO_x–GaAs diodes. Here, oxide and fluoride layers up to 5000 Å in thickness were used without drastic alteration in device behaviour. While tunnelling seemed an unlikely mechanism in some previous work, this finding rules it out as an explanation of the injection process. The internal quantum efficiency of some of the structures used as photodiodes was very high (as much as 28% at 4800 Å). The Au–SiO_x–GaAs diodes exhibited forward-bias electroluminescence with a spectrum characteristic of radiative recombination on GaAs. Such emission demands the injection of holes into the n-type semiconductor. In addition a variety of switching and negative resistance phenomena was observed in the devices, dependent upon the nature of the anode metal.

It seems clear from the nature of the insulating films used in these experiments, their thickness and mode of preparation and the type of metal electrodes deposited on them, that phenomena similar to those we have discussed above will be occurring in the oxide (or halide) layer. The current densities are such as can only be explained if the layer is 'formed'. The localized electroluminescence observed by Jaklevic *et al.* (1963) is very reminiscent of the localized conduction and emission phenomena we have already discussed. We shall now consider how the various models of conduction and forming in oxides can account for these effects.

Hickmott (1964 a, b) has proposed that conduction and negative resistance in a thin oxide film are high-field phenomena, with a nonuniform potential distribution within the material. Near the negative electrode, a high-field region between 100 and 150 Å thick is believed to develop, and accounts for the electron emission and electroluminescence of metal–oxide–metal diodes or triodes. This same model could allow the injection of electrons or holes into a semiconductor if this were substituted for the cathode material. It is less easy to understand the process in terms of Simmons and Verderber (1967) although they present a clearer picture of the physical mechanisms taking part. The difficulty lies in the fact that energy loss processes in the insulator are likely to be dominant, and no mechanism is proposed whereby the carriers can gain energy efficiently enough.

Dearnaley *et al.* (1969) have proposed a filamentary conduction model, which immediately explains the localized electroluminescence observed by Jaklevic *et al.* The size of the emitting spots will naturally be larger than the filament diameter, since minority carriers injected into semiconductor will have a certain mean free path before recombination. Since, in their model, the filaments are considered to possess weak spots at which the field is high, and undergo rupture at these points, conditions for efficient heating of carriers in localized high-field regions are met. Ralph (1970) has recently pointed out that, in the propagation of a filament from the anode towards the cathode, termination of the atomic rearrangements will naturally occur when the distance between the filament and the cathode is short enough (say, about 30 Å) to allow tunnelling. Therefore, it is natural that a high-field region should exist very close to the cathode, and this may in turn explain some of Hickmott's (1964) observations on the field distribution in oxide layers. The presence of a fine filament terminating in a high-field zone (figure 31) gives ideal conditions for the injection of hot holes into the semiconductor, and therefore the

possibility of efficient electroluminescence. If this is correct, the electroluminescence efficiency will tend to increase with the degree of forming of the oxide layer. This mechanism of injection resembles that proposed by Fischer (1963) as a means of creating hot carriers on ZnS crystals, by field emission from the tips of fine crystallites embedded in a structurally different matrix (figure 31(*b*)).

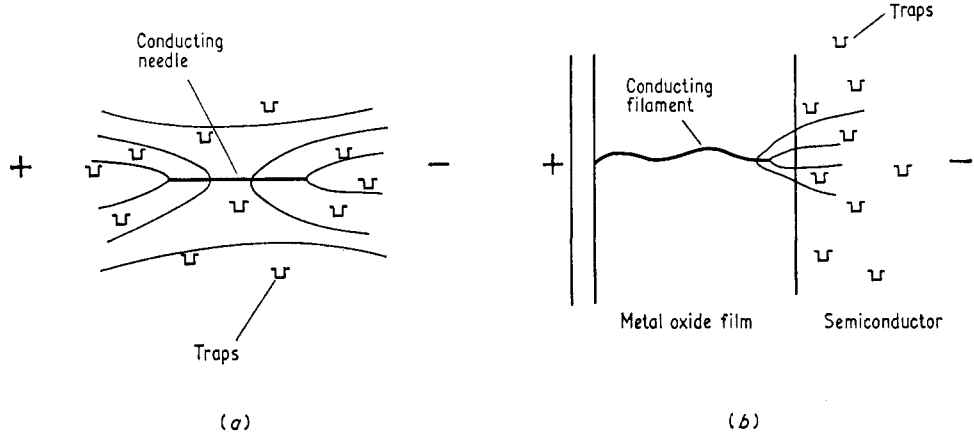

Figure 31. The analogy between (*a*) the alternate field emission of electrons and holes from a conducting needle embedded in a luminescent medium subjected to an alternating electric field (after Fischer 1963) and (*b*) field emission of holes from the end of a conducting filament embedded in an insulating matrix, near to a semiconductor interface. In both cases, radiative recombination may take place at traps.

4.5. *Noise*

The current–voltage characteristic of a formed metal–oxide–metal structure shows very considerable noise fluctuations (of the order of 10%) in the region of differential negative resistance, and a related phenomenon may be of the step-like recovery from the high-resistance to the low-resistance state during switching.

These features were discussed by Dearnaley *et al.* (1970), who observed that the noise will be large, particularly between V_T and V_u on the filamentary model since in this region filaments are rupturing and regrowing. Fluctuations will occur both in the number of mobile carriers and in their mobility. On the other hand, as is observed, the noise is expected to be much less at voltages below V_T (when essentially no filaments are breaking) and above V_u (when most of the filaments are broken and bulk conductivity dominates).

Below V_T one has an interesting situation which may perhaps be regarded as a one-dimensional analogue of the model put forward by McWhorter (1956) and by Flinn *et al.* (1967) to explain the noise behaviour in the conducting channel of MOS transistors. These authors consider that fluctuations in the number of charge carriers occur due to tunnelling out of the semiconductor into traps situated in the gate oxide layer, from which they are subsequently released. A similar mechanism may be expected to occur owing to tunnelling (or phonon scattering) of carriers out of a filament into traps in the surrounding amorphous matrix. Assuming a uniform distribution of traps, a similar $1/f$ noise should result, with a mean square current fluctuation proportional to V^2. It is interesting that the noise formed in metal–oxide–metal structures, MOS transistors, metal-oxidized silicon surface-barrier

diodes and in oxide-coated cathodes exhibits a $1/f$ dependence on frequency, often referred to as 'flicker noise', the origin of which has never been well understood. On the basis of such a filamentary model it is easy to see why the $1/f$ noise should be minimized by activation of an oxide cathode: if there are more filaments present the statistical fluctuation in current will be reduced. Measurements of the noise in two oxide cathodes pressed together (Hannam 1956) showed a variation in mean square noise proportional to I^2, as would be expected by analogy with the model of Flinn *et al.* (1967).

It does not seem possible at the present time to offer a simple interpretation of the size of the steps seen (Greene *et al.* 1969) in the recovery of conductivity in a metal–oxide–metal device. The difficulty, in terms of a filamentary model, is that the steps are much too large (figure 12) to correspond to individual filaments regrowing, since the steps amount to resistance changes of the order of 10^8 ohms rather than the 10^7 ohms predicted earlier (§ 3.2). Of course, some of this discrepancy may merely reflect the difficulty of observing small resistance changes, and the dominating effect of low-resistance filaments. Without some further development, the alternative models of switching cannot yet account for this phenomenon.

5. Comparison with other systems

5.1. *Current-controlled negative resistance and bistable switching in oxides*

In this review, we have been primarily concerned with oxide films which exhibit voltage-controlled negative resistance (VCNR). However, a number of oxides develop current-controlled negative resistance (CCNR) and bistable switching after a forming process. A mere division of oxides into two classes, exhibiting the two types of negative resistance (figure 2) is unsatisfactory since some oxides can show both (table 9). Thus titanium oxide and tantalum oxide have been found to show VCNR (Hickmott 1962, 1970 private communication) or CCNR (Chopra 1963, 1965, Beam and Armstrong 1964, Argall 1968). These apparently conflicting results probably arise owing to different methods of preparation and therefore different structure and stoichiometry of the oxide film. Table 9 summarizes the findings of different workers.

Prior to forming, diodes of Nb, Ta and Ti oxides have characteristics as shown in figure 32(*a*). Chopra (1965) reports that by increasing the reverse current to above 100 mA cm^{-2} the characteristic is transformed irreversibly from a rectifying characteristic to a symmetrical one, exhibiting current-controlled negative resistance (figure 32(*b*)). In the region below peak current the characteristic has the form $I \propto V/d^3$, which is similar to the behaviour in SiO reported by Simmons and Verderber (1967) (figure 8). Chopra suggests that this indicates a space-charge limited current. Beyond the negative resistance region the dynamic impedance of the device is effectively zero. A holding voltage V_h is required to keep the device in this state, where V_h is independent of electrode material and of the oxide thickness (as in Ovshinsky devices). Furthermore, this holding voltage is only slightly dependent upon temperature, being about 10% greater at 77 K than at 300 K. V_h varies with the relative permittivity ξ_r as ξ_r^{-2}.

Hickmott (1966), working with Nb–Nb_2O_5–metal diodes, reports a second stable high conductivity state which was overlooked in previous work. Bistable switching can be achieved between the two states but, unlike the current-controlled negative resistance reported by Chopra, the switching voltage and details of the

Table 9.

Oxide	Voltage–current characteristics	Method of preparation	Thickness (Å)	Author
SiO	VCNR bistable + memory	Vacuum evaporation	200–3000	Simmons *et al.* (1967)
SiO	VCNR bistable + memory	Vacuum evaporation	200–3000	Hickmott (1962)
Al_2O_3	VCNR bistable + memory	Evaporate Al film followed by anodization in aqueous solution	300	Hickmott (1962)
ZrO_2	VCNR bistable + memory	Evaporation of Zr on to glass followed by anodization in electrolyte, or evaporation of oxide	1000–2000	Park and Basavaiah (1970)
ZrO_2	VCNR bistable + memory	Anodization of metal in electrolyte	500	Hickmott (1962)
Ta_2O_5	VCNR bistable + memory	Anodization of sputtered film	500	Hickmott (1962)
Ta (oxide)	CCNR	Thermal anodization of metal in oxygen	50–500	Chopra (1965)
Ta (oxide)	VCNR bistable	Electrolytically oxidized films	0·1–1 μm	Kreynina (1962)
TiO_2	VCNR bistable + memory	Electrolytically oxidized	500	Hickmott (1962)
Ti (oxide)	CCNR	Thermal oxidation	50–500	Chopra (1965)
Ti(oxide)	CCNR (during forming unpredictable) bistable memory	Anodized in electrolyte	1000	Argall (1968)
Nb (oxide)	CCNR	Thermal oxidation	50–500	Chopra (1965)
Nb_2O_5	CCNR bistable switching	Anodization in electrolyte	90–3000	Hickmott (1965 a, b, c)
Nb (oxide)	CCNR	Anodization in electrolyte	45–2200	Beam and Armstrong (1964)
Nb_2O_5	VCNR	Relatively sputtered films		Pinto and Shahan (1968)
Cu (oxide)	VCNR	Anodization in electrolyte	0·1–1 μm	Kreynina (1962)
Mg (oxide)	VCNR	Anodization in electrolyte	0·1–1 μm	Kreynina (1962)
CeO	VCNR	—	—	Simmons (1970)
Cu_2O (single crystal)	CCNR bistable		—	Cook (1970)
NiO	No CCNR bistable switch	Thermal oxidation	—	Gibbons and Beadle (1964)
VO, V_2O_3, VO_2	CCNR	Fusion of oxide beads in reducing atmosphere	—	Cope and Penn (1968)

new low impedance state depend on the nature of the anode metal. Hickmott's forming procedure is somewhat different to that of Chopra. The voltage across the device is raised by using a constant-current supply in the direction of easy current flow; at the forming voltage sudden breakdown occurs to the high-conductivity state. The forming voltage and current are not well defined; they are both sensitive functions of the anodizing electrolyte, the metal electrode and the detailed procedure used for forming. The procedures for switching Nb–Nb_2O_5–Bi and

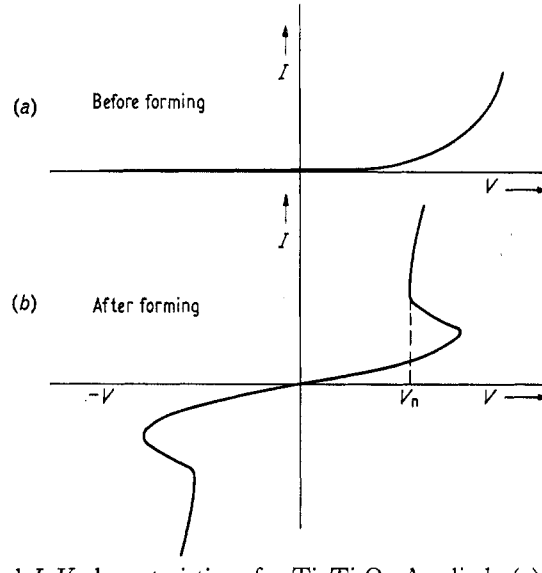

Figure 32. Typical I–V characteristics of a Ti–Ti_2O_3–Au diode (*a*) before forming, (*b*) after forming (from Chopra 1964).

Nb–Nb_2O_5–In from the high- to the low-impedance state and back is illustrated in figure 33. When switching from the high- to low-impedance state, a low-impedance, constant-current supply is used; conversely, switching in the reverse direction is achieved by reversing the polarity and using a constant-voltage supply. Typically, the impedance ratio is 20 : 1. The switching voltages are independent of oxide thickness. In contrast to the high-impedance state, in which the I–V characteristic is independent of the electrode metal, the characteristics of the low-impedance state are electrode-dependent. Thus when metals of low-work function are used (for example, In, Zn, Cd) the current-controlled negative resistance observed with Bi or Au electrodes is replaced by a small voltage-controlled inflection at 0·1 to 0·2 V (Hickmott and Hiatt 1970). The short relaxation time (of the order of 10^{-6} s) determined by self-sustained oscillations suggests (Chopra 1963) that the negative resistance is electronic in origin. Hickmott (1970 private communication) reports that under favourable conditions the bistable device can be switched from low to high impedance in about 2×10^{-6} s and in the reverse direction in 10^{-6} s, though in some devices this may take as long as 10^{-3} s. The switching speed appears to depend on both the preparation and forming procedure of the insulator film. Beam and Armstrong (1964) report that conduction in titanium oxide occurs at filamentary singularities. For the same material, Argall (1968) observes that the current after forming is proportional to the diode area, suggesting either that the whole device is active or that there is a fairly high density of fine filaments.

As a general classification of negative resistance devices, it appears that current-controlled (S-shaped I–V) characteristics can result from filament formation which is regenerative; creation of a single conducting filament leads to conditions in the medium which bring about a rapid radial growth or conductivity rise of the filament (for example, by a strong thermal dependence of the conductivity or by a phase

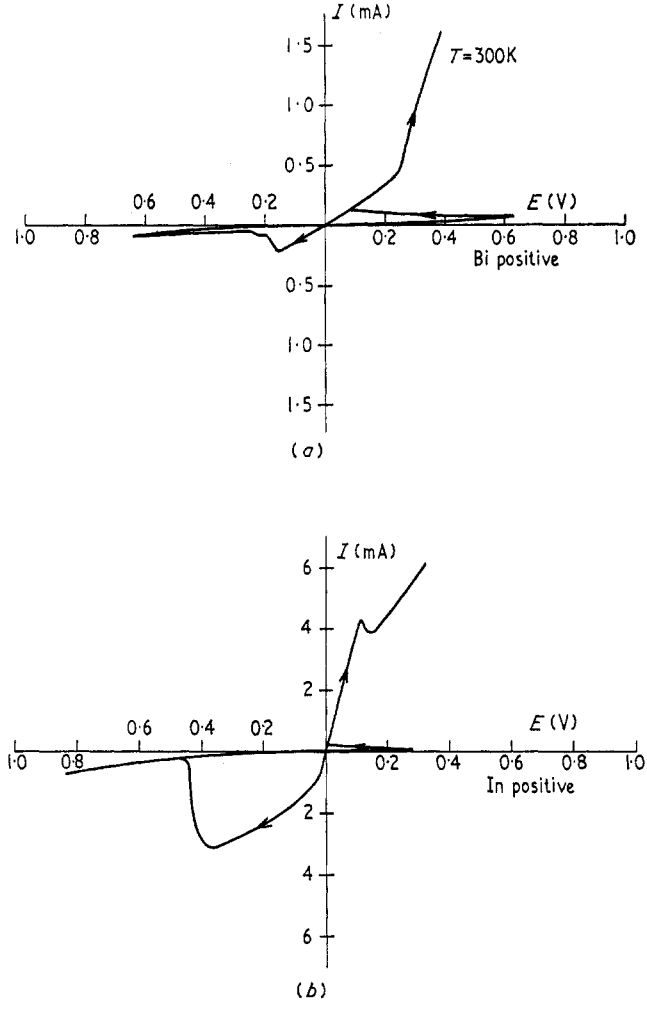

Figure 33. Bistable switching of (*a*) Nb–Nb_2O_5–Bi diode with oxide thickness 1250 Å, (*b*) an Nb–Nb_2O_5–In diode (after Hickmott 1965 c).

change). On the other hand, voltage-controlled (N-shaped I–V) characteristics, we have seen, can result from nonregenerative filament formation. In this case the creation of a filament is not followed by a sudden rise in its conductivity and, if it remains of small dimensions, the resistance of the device is only slightly perturbed. Conditions are therefore retained for the further development of other filaments wherever appropriate nucleation and propagation can occur. These ideas offer some explanation for the inconsistencies observed by different workers in what is ostensibly the same material. The regenerative behaviour of conduction in an oxide film may

well be critically determined by stoichiometry, strain or impurity content. It is difficult to be more specific since most workers do not identify precisely the content of the oxide films used. In niobium alone can one be reasonably sure (Hickmott 1965 a, b) that the oxide is Nb_2O_5. Clearly further experiments are needed to relate the observed behaviour to the detailed nature of the oxide film.

5.2. *Comparison with the Ovshinsky devices*

The switching and memory properties of the chalcogenide glass films are at present the subject of intense research and controversy in laboratories throughout the world. The observed switching phenomena can be conveniently divided into two categories: (i) *The threshold switch.* Here the glass film can be made to switch from a high-impedance state (typically of the order of 300 megohms) to a low-impedance state (of the order of 30 ohms), by increasing voltage across the film above a critical value. The device will only remain in this low-impedance state provided the current is maintained above a critical value called the 'holding current'. The I–V characteristic is S-type; (ii) *The memory or bistable switch.* In this case the device, once switched into the low-impedance state (the 'on' state), remains there. It can be switched back to the high-impedance ('off') state by a high-energy pulse of short duration. A glass film may exhibit either threshold or memory switching, depending on its mode of fabrication; alternatively, a given device may show both types of behaviour when subjected to appropriate electrical pulsing.

The memory switch shows rather slow switching times, of the order 10^{-3} s, and it is believed that the memory state involves atomic rearrangement within a highly conducting filament bridging the electrodes. It may be that this is in a crystalline state. Such filaments can then be destroyed by a pulse of current—the rapid heating and quenching returning the material to its nonconducting state. Ovshinsky and his co-workers have observed directly the growth of such filaments (Sie 1970).

In the threshold switch, on the other hand, the duration of the switching process (or at least the final stage of it) is only about 10^{-10} s and for this reason it was considered to be electronic in nature. However, Warren (1969) has shown that a simple thermal runaway, taking proper account of the strong temperature dependence of the conductivity of chalcogenide glasses, can allow the rapid switching observed. Boer and Ovshinsky (1970) argue that this explanation is inadequate, since it should be feasible to stabilize the transition with a sufficiently large load resistor; instead, one observes self-sustained oscillations. Warren counters this by relying upon the capacitance of the switch to store enough energy to sustain the switching; if the external resistance is very high, the conducting channel will cool, and the inverse of runaway will occur; that is, the switch will go 'off'. A series of relaxation oscillations could then occur. The thermal runaway model may not provide the whole explanation, however, since there are indications of nonthermal field-induced effects in the thinner films (Warren and Male 1970).

It seems evident that the behaviour of oxide films exhibiting voltage-controlled negative resistance has little in common with that of the ovonic chalcogenide devices. However, there are certain similarities between the oxide structures described in § 5.1 and the ovonic switches (table 10). Oxide films, for example, of Nb_2O_5 show a threshold switching (Chopra 1965) and current-controlled negative resistance. The holding voltage is independent of film thickness, as in the case in ovonic threshold switches. The switching is rapid (of the order of 10^{-7} s) and with

Table 10.

	Oxide films			Chalcogenide glasses	
	Class I	Class II			
Current–voltage characteristics	VCNR and switching Analogue memory	Threshold switch (CCNR)	Bistable switch	Threshold switch	Bistable memory switch
				Any given device can show both switching and memory effects when subjected to the appropriate electrical pulsing	
Switching time	In SiO—high–low required 100 ns; low–high $\simeq$ 2 ns but has a dead time associated with it before it can be switched again, ie dead time 10^{-6}–10^{-1} s	Of the order of 1 μs relaxation time	1 μs to 1 ms, unpredictable	10^{-10} s	10^{-6} s
Device thickness	100–4000 Å; occasional report of VCNR	Up to 1 μm		5000 Å–1 cm	5000 Å–1 cm
Forming and voltage switching voltage	The switching voltage covers a very wide range from 1 V in Nb_2O_5 to several hundred in NiO. Switching voltage independent of oxide thickness			Switching tens of volts, threshold voltage increased linearly with thickness	
Other phenomena	Electron emission and electroluminescence in formed device	Electroluminescence in formed device			
Method of preparation	Evaporation, thermal oxidation, electrolytic oxidation	Thermal oxidation.	Electrolytic oxidation	Evaporation, either thermal or sputtered films	
Reliability	Switching eventually destroys device. Analogue memory unreliable due to noise	Switching characteristics very erratic Indefinite switching	Switching eventually destroys diode, no. of switches depends on oxide and method of preparation	Indefinite switching, switch for months without notable change in their characteristics	10^8 switches before device fails
Range of memory states	Analogue (varying between the two limits of the full on and off states) varies by a factor of 10^8 in SiO	Two, requires a holding voltage in the conductive state	Bistable	Two, requires a holding current in the on state. Resistivity change of the order of 10^7 ohms	Bistable

a high series resistor, self-sustained oscillations are observed. Several workers have reported localized conduction at filaments, but it is not clear in these cases whether one or more filaments are involved in the conduction process at any one time; Hickmott (1969 a, b) obtained evidence that switching in $Nb–Nb_2O_5–Au$ structures occurs at single points. Since heating can cause evaporation of electrode material and even some of the oxide, it is possible for a filament to be destroyed (as in 'self-heating' breakdown), whereupon another may develop at another point. Subsequent examination of the film can in this way reveal several pinholes (Park and Basavaiah 1970). Both oxide films and chalcogenide glass structures give rise to memory switches. This is a relatively slow switching process in each case, and in both instances is believed to be accompanied by ionic movement within a filamentary conduction path. The ovonic devices are generally much thicker than the oxide films, however. Other direct comparisons between the device performance in these two types of material would be misleading, however, since at the present time the ovonic devices have been subjected to a much more intensive study resulting in many technological improvements. It should be remembered, perhaps, that the oxides are simpler in composition and mode of preparation. Since they may be more homogeneous it could be that the processes which lead to the initiation of localized thermal runaway (for example, electrode topography) are more readily studied in oxide films than in thin chalcogenide structures.

5.3. *A comparison of switching in crystalline, microcrystalline and amorphous films*

Although the majority of results demonstrate that switching and memory phenomena occur most readily in amorphous or polycrystalline films, threshold and bistable switching does occur in some single crystals. Thus Cook (1970) observed switching in Cu_2O crystals, Anderson (1968) saw similar effects in thin ZnS crystals, and Lee *et al.* (1969) report many observations of current-controlled negative resistance and bistable switching in the layer compounds SnS_2, ZrS_2 and PbI_2. In Cu_2O the current in the high-conductivity state travels along a filamentary path produced as a result of a localized structural change in the single crystal. This is very similar to the behaviour of the Ovshinsky bistable switch, and indicates that this type of switching is not confined to amorphous, glassy films.

It does seem, however, that the phenomena occur more often in disordered structures. Some possible reasons for this are: (i) growth of conducting paths may be easier; ionic diffusion along dislocations and grain boundaries is probably easier, and surface states may play a role, (ii) physical changes in the structure may proceed more readily (cf. Ovshinsky's suggestions), and (iii) localized electron states can occur in disordered solids which do not occur in crystals. The distribution of impurity levels in energy may give rise to a wide impurity band in a dirty amorphous solid.

Some theories of the structure of highly disordered covalently bonded media (for example, Grigorovici 1967) hold that small (~ 20 Å) regions exist in which there is a high degree of short-range order, but that this situation breaks down at longer ranges. The relatively ordered rings and networks are bounded by a number of dangling bonds or surface states (Brodsky *et al.* 1969). In this way we can see that the switching behaviour in amorphous and microscrystalline media may not be so very different, if the phenomena responsible for conduction occur principally at the grain boundaries. Thus covalent amorphous silicon oxide and microcrystalline

magnesium fluoride (an ionically bonded material) films may owe their conduction and switching behaviour to basically similar phenomena.

One possibility, not discussed previously, is that surface states which have been observed in microcrystalline oxides may be responsible for their forming and negative resistance. Nelson *et al.* (1964, 1967) have observed, by electron spin resonance measurements, the surface analogue of the bulk F-centre. Such states might be created at grain boundaries by anion transport as in the model of Greene *et al.*, and could be responsible for the propagation of filamentary conducting paths (cf. Dearnaley). They are poisoned by oxygen and should have optical transitions in the same energy range as that reported for CaO cathodes (Jones and Mee 1964). Moreover, it is clear why forming occurs best in ionic oxides and halides that are microcrystalline: the ratio of surface sites to bulk sites increases as the crystallite size decreases. In amorphous (noncrystalline) films, such as SiO_x, increased conduction has been attributed to strain near contacts, which could result in crazing of the material (J. E. Ralph, private communication). Further experiments are in progress to test these ideas.

6. Practical implications

In this section we shall consider the practical applications of the phenomena discussed above and shall conclude that they are both significant and widespread. They are, moreover, not confined to wholly novel devices, for we believe that the effects play an important role in the proper functioning of several existing types of device. They have, however, hitherto been dismissed as 'breakdown' due to 'pinholes' or other ill-defined causes. We feel that by gaining a better understanding of the nature of the processes involved it may be possible to improve fabrication techniques.

6.1. *Switches*

We have seen that voltage-controlled switching from one impedance state to another can be accomplished in a metal–oxide–metal structure in times which are of practical interest. Such switches can be manufactured very simply, and may have small dimensions and consume little power. It would be a simple matter to prepare them on the surface of integrated circuits or on other electrical components. Often, the devices can display several stable states and in principle could be used in analogue as well as digital applications. The materials employed are cheap and, being noncrystalline, are easily handled. For military applications there is the potential asset that, like the ovonic switches, the devices may be highly radiation-resistant (see below).

What, then, are their disadvantages? The chief objection lies in the fact that the characteristics are not reproducible, from one operation to the next, or over a useful life. Hickmott (1970 private communication) remarks that oxide switches prepared from impure materials show better reproducibility, but even then it is difficult to achieve uniform values of the impedance. Most structures cannot be used in atmospheric conditions, and require encapsulation which adds substantially to the cost, but even when so protected the switching threshold is variable and there is a lack of stability below the threshold voltage. Over successive operations the devices frequently deteriorate and show evidence of local overheating; they may

eventually persist in either the high- or low-impedance state. The switches are also rather noisy.

These defects are inherent in the mechanism proposed in the filamentary model, since ionic motion and local heating are very likely to result in instabilities and short life. The role of chemical impurities has not yet been established.

There is no firm evidence as to whether oxide switches are radiation-resistant or not, though a study of this would throw more light on the mechanisms responsible for their operation. Grannemann *et al.* (1960) report no change in the electrical resistivity of reduced TiO_2 for a dose of 5×10^{15} fast neutrons per cm^2, and aluminium and tantalum electrolytic capacitors withstand some 10^{15} fast neutrons per cm^2 without deterioration (Kircher and Bowman 1964). The oxidation of zirconium is unaffected by fast neutron fluxes up to 10^{20} per cm^2 (Harrop and Wanklyn 1967). These circumstantial pieces of evidence suggest that electronic conduction processes in oxide films are not easily affected by fast neutron bombardment; gamma rays are even less effective at producing change (Harrop and Wanklyn 1965).

6.2. *Memory arrays*

One area of application of fast, compact switches lies in memory storage. The fact that the metal–oxide–metal switch is a two-terminal device with long-term memory makes it immediately attractive from this point of view.

A thin-film resistive memory array based upon voltage-controlled negative resistance in SiO_x was first proposed by Nielsen and Bashara (1964) and such a device has been described by Simmons and Verderber (1968).

The array was prepared by vacuum deposition of SiO on to aluminium previously deposited in stripes on a glass substrate. Gold upper electrodes were next evaporated as stripes running at right angles. Appropriate signals fed to a pair of anode and cathode strips would thus switch the corresponding unit of the memory. The success rate of fabrication was claimed to be high and a packing density of 10^4 per square inch was envisaged. The idea has not been exploited, however, largely owing to the drawbacks listed in § 6.1.

Varker and Juleff (1967) described a read-only memory written in SiO_2 films by means of a finely focused scanning electron beam. The target consisted of a 4000 Å layer of SiO_2 grown on n^+ silicon on a p-type diffused epitaxial substrate. A 15 keV electron beam focused to a diameter of 0·25 μm and with a beam current of 10^{-9} Å produced permanent changes in the state of the SiO_2. These scanned regions allowed an increased penetration of electrons into the p–n junction below, and also showed enhanced etching properties. Varker and Juleff suggest that the SiO_2 is 'densified' and thinned by the bombardment.

It seems possible, however, that what is occurring is an enhanced conductivity of the oxide produced as a result of the electric field and heating of the material. The instantaneous power dissipation exceeded 1 kW cm^{-2}. Increased electron conduction would allow the negative charge built up on the dielectric surface to escape more readily to the p–n junction below, and produce a signal there. Whatever the mechanism, the device possesses a very high packing density of information, exceeding 10^7 bits/cm^2, and great simplicity in writing and reading. The system suffers, however, from the fact that the memory cannot, in the form described, be erased.

It is interesting to note that in each of these memory arrays, information is stored in the form of permanent changes in the interior of a thin film. Probably these are localized atomic rearrangements which may, we have seen, consist of fine filamentary conduction paths. This type of memory storage is believed also to occur in biological systems. In the human brain, it is generally held that changes in the synaptic junctions, induced electrochemically, influence the electrical coupling of axons. Electron microscopy has revealed that across the narrow synaptic cleft (about 2000 Å wide) are many fine filaments, called trans-synaptic filaments, which may be less than 100 Å in diameter. Assuming that these play some part in memory, it is remarkable how analogous this system may be to the oxide-layer memory.

6.3. *Hot electron, cold cathode emitter*

This is basically a metal–oxide–metal structure in which the anode film is sufficiently thin that hot electrons traversing the oxide are able to escape into the vacuum above. Some formed devices have emission efficiencies (defined by the ratio of the emitted current to the circulating current) as high as 1%, about two orders of magnitude greater than that achieved by tunnelling through thin oxide films. Some loss of efficiency is caused by attenuation of the slow electrons in the anode film. However, Gould and Collins (1970) have measured attenuation lengths in gold of about 1000 Å at 3 V bias across a 1000 Å SiO film. The electron emission is, as we discussed in § 4.1, highly localized and this is attributed (Dearnaley *et al.* 1970, Gould and Collins 1970) to the existence of filamentary conduction paths through the oxide. In order to explain the effects which are seen, it is necessary to assume that a small but significant fraction of the electrons are emitted from the cathode with an energy close to the full potential difference across the oxide. Hickmott (1963) came to the same conclusion regarding electrons from $Al–Al_2O_3–Au$ structures. Hickmott observes that, at low potentials, electrons are emitted despite the potential difference being less than the work function of gold. Since, at low voltages, the emission tends to be rather uniform across the cathode and does not show the localized bright emitting centres, this feature of the cathodes remains unexplained.

6.4. *Alpha-numeric display panel*

A simple extension of the devices described in § 6.3 has been to produce a display panel consisting of an array of 25 cold cathode emitters (of $Au–SiO_x–Al$) in conjunction with a phosphor screen. A high potential accelerates electrons from the cathode to the screen to provide luminescence clearly visible under normal lighting conditions (Lomax and Simmons 1968). This panel was on show at the 1968 Physics Exhibition organized by The Institute of Physics and The Physical Society (figure 34). A problem, as with oxide switches, has been to achieve a reasonable working life—in early devices the lifetime was only a few minutes. It was found that the lifetime depends on the composition of the films, and that those near to SiO_2 in composition had longer lives. Various insulator materials were tried out, and a working life of around a thousand hours has since been achieved. Clearly, the devices are cheap to make and in principle could be quite large.

In the course of this work, it was observed that the electron images on the phosphor screen appear as rings or arcs with a narrow range of radii. These effects were explained by Simmons *et al.* (1966) as due to coherent scattering of electrons

in the gold anode layer, and this interpretation has been confirmed by Gould and Collins (1970). It is necessary in this explanation that the emission takes place from small spots on the cathode. The intensity of the individual arcs fluctuates with time, and new emission points constantly appear and disappear.

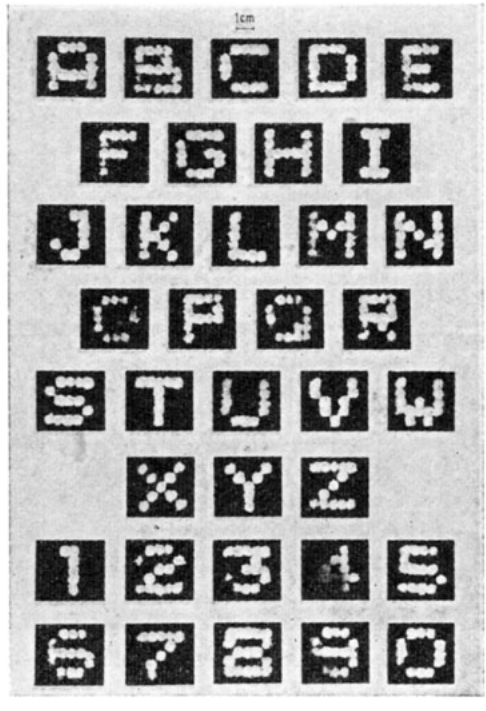

Figure 34. An alpha-numeric optical display panel using an array of 25 cold-cathode Au–SiO–Al units each of 10 mm² area (by courtesy of STL, Harlow).

6.5. *Electroluminescent* MOS *devices*

Structures consisting of a metal–oxide (or halide)–semiconductor sandwich have been discussed in § 4. Such an arrangement has yielded encouraging luminescent efficiency, and there is still industrial interest in these devices. While a number of II–VI semiconductors of great interest for electroluminescence exist which occur only in one conductivity type, so that junction injection cannot be used, other techniques of carrier injection are worth exploring. It is feasible that ion implantation may overcome the difficulties of achieving type conversion and junction formation in CdS, ZnO, etc., and indeed a recent paper (Hou and Marley 1970) claims that electroluminescent diodes have been produced in CdS in this manner. However, it is possible that the ion bombardment produced a high-resistance layer near the semiconductor surface, similar to that reported in the ion implantation of ZnTe (Donnelly *et al.* 1970). Under such circumstances an avalanche mechanism has been proposed to take place in what is essentially a p–i–p structure (Weiser and Woods, 1965).

6.6. *Oxide-coated cathodes*

We have earlier introduced the idea that much of the behaviour of the oxide-coated cathode may be attributed to phenomena analogous to those occurring in

metal–oxide–metal and MOS structures, and that 'activation' is related to 'forming' (Dearnaley 1969). It seems unlikely that this hypothesis, even if confirmed, could further the long-established empirical techniques of cathode fabrication developed over may years of trial and error. It is possible, however, that it might assist the general treatment of electronic noise in devices based on oxide layers (cathodes, MOS transistors, etc.) and provide a more physical explanation for its origin. It is possible that the correlation might suggest certain other materials, or combinations of them, which should be tested as oxide cathodes, perhaps with a better resistance to 'poisoning'.

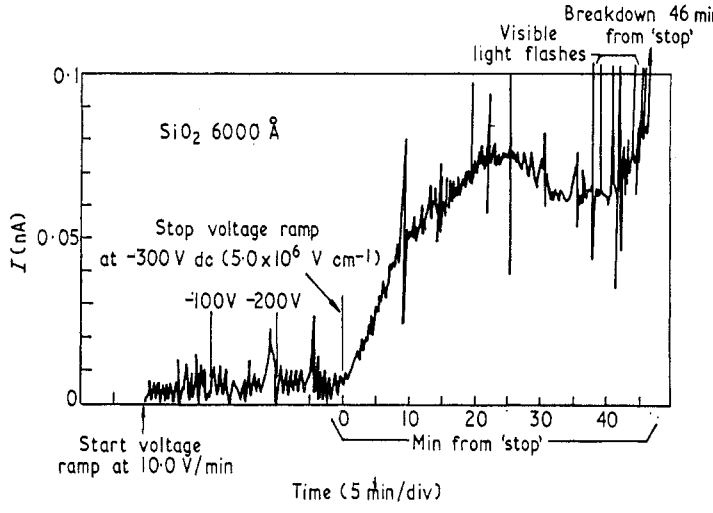

Figure 35. Variation of current (I) in a Au–SiO_2–Si MOS structure for which the gold potential was raised to 300 V. Conduction continued to increase until dielectric breakdown occurred after 46 min. Transient current pulses, accompanied by visible light flashes, occurred just prior to breakdown (from Worthing 1968).

6.7. MOS *capacitors and transistors*

So far we have discussed applications in which a high electrical conductivity in the oxide layer is desirable, at least at some stage of operation. In the case of metal–oxide–semiconductor capacitors and transistors this is not so, and every attempt is made to achieve a dielectric of high strength. Nevertheless, it appears that even carefully prepared thermal oxide on silicon exhibits some of the features we have associated with forming.

Worthing (1968) studied the behaviour of Au–SiO_2–Si capacitors in which the oxide was grown thermally in steam at atmospheric pressure, at 1050 °C for one hour. The oxide so produced had thicknesses averaging 6000 Å, and on to this gold dots 600 μm in diameter were deposited, after which the structures were baked in air at 400 °C for an hour. A different voltage breakdown was observed depending upon whether the gold electrode was positive or negative. With the gold positive, the current, near to breakdown voltage, was time-dependent (figure 35) and cumulative effects were seen. Thus, although the current was not allowed to rise above 10^{-10} A, subsequent tests revealed a progressive decrease in breakdown voltage. In other words, the film has a 'memory' of its previous history. When the current exceeded about 10^{-10} A, visible pinpoints of blue-white light were seen flashing at various points on the gold anode, and subsequent examination showed

tiny 'burned' spots. It was suspected that these might have originated at 'pinholes' or 'potential pinholes' (Lopez 1966) in the oxide, and a test was made for these by a chemical etching technique. From these test results, it was concluded that the oxide was relatively free from such defects and that the light flashes do not result from pinholes.

Breakdown of the SiO_2 with the gold negative, on the other hand, was not time-dependent and showed no unexpected behaviour. The observed dielectric strength (between 7×10^6 and 1×10^7 V cm^{-1}) agreed with that reported for quartz and vitreous silica. This was also the behaviour for structures tested (with either polarity) at 77 K.

It is an almost inescapable conclusion from this work that, under positive metal electrode conditions, permanent atomic rearrangement takes place at localized spots in the oxide and that these are accompanied by transient conduction and luminescence. Clearly, this allies the phenomenon to those we have discussed in oxides of lower stoichiometry but here it occurs to a very much lesser extent. We would attribute the reason for this to the reduction in the amount of 'forming' which takes place as the composition approaches SiO_2.

It is well known that many of the techniques of preparing oxide layers for MOS capacitors and transistors are unsuccessful and result in instabilities in the conduction and charge storage of the oxide. Thermally grown oxides have therefore been preferred for device fabrication. Anodically grown SiO_2 is probably the least satisfactory method of oxide growth, and such films have a high electronic conduction. Hitherto, these instabilities etc. have been attributed to chemical impurities, e.g. sodium, present in the oxide and so leading to ion drift. Admittedly, this is an important factor (Snow *et al.* 1965) and anodic oxides are particularly impure. Carefully grown thermal oxides, utilizing only pure ingredients and precautions to minimize sodium contamination, certainly have very few impurities. We would argue, however, that a secondary cause of instability lies in the 'forming' of small regions of oxide where, probably, the stoichiometry departs a little from that of SiO_2. A second reason why thermal oxides are superior is that the chemical reaction tends most nearly to produce SiO_2, under conditions of zero electric field. Anodic oxide, on the other hand, is grown under an extremely high electric field which is able to induce forming; we shall return to this point in § 6.9.

Another phenomenon has been observed in silicon planar n–p–n transistors in which a thermally grown 4000 Å SiO_2 film was prepared on top of the emitter–base junction (Verwey 1969). Application of a positive voltage to a metal gate electrode on the surface of the oxide produced incipient breakdown, which resulted in a negative charge becoming stored in the oxide and resulting in a shift in the transistor characteristics. It appears that conduction through the oxide has been accompanied by trapping of electrons in the amorphous oxide. Most of the conduction models we have discussed would allow such an effect to take place, but the example is of interest in demonstrating how oxide conduction phenomena may influence MOS and related transistors.

6.8. *Surface barrier diodes*

For many years, a simple technique of preparing high-quality diodes in n-type semiconductors, such as silicon, has been via the so-called surface barrier. A noble metal, such as gold, is evaporated on to the etched and cleaned semiconductor and, in most cases, a diode with excellent reverse bias characteristics results. It has been

observed, however (Siffert *et al.* 1964), that this is the case only if the structure is exposed to oxygen after the anode deposition. The non-injecting contact takes a measurable time to develop, and this time depends upon the nature of the metal layer. In the case of aluminium, the contact may remain injecting for a long time. Siffert *et al.* observed too that the rate of development of rectifying characteristics was influenced by the application of an electric field across the oxide layer.

We hold that these phenomena are related to the mechanisms of electronic conduction. The same conditions which are favourable to hole injection in a forward-biased MOS structure (§ 4.4) lead to a non-injecting contact in the reverse biased mode (figure 36). The electric field within the oxide layer raises the height

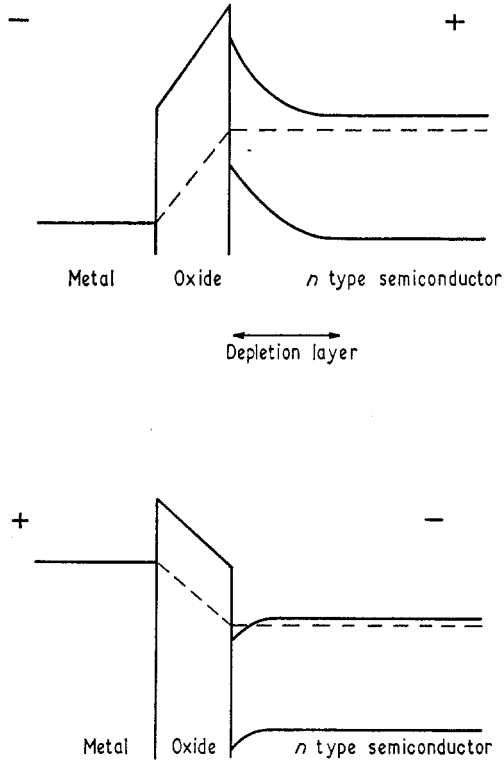

Figure 36. Electron energy levels in a reverse biased and (below) forward biased metal–oxide–n semiconductor structure. In the upper case a depletion layer is formed with an enhanced barrier height (as in surface-barrier diodes), while in the lower case, hole injection can take place.

of the barrier. The device is not, as it has sometimes naïvely been described, a Schottky barrier dependent for its operation upon the work function difference between metal and semiconductor. In the absence of an oxide, the barrier height is lower and the contact may be dominated by the metal work function. Thus evaporated aluminium contacts initially show a more or less ohmic behaviour (Andrews 1962) but over a period of time (which may run into years) the contact becomes non-injecting. In fact, given long enough, it seems that the characteristics of all metal–silicon diodes tend to become the same, and presumably at this stage the oxide film has dominated the behaviour of the device.

It is worth noting that this understanding of the behaviour of surface barrier diodes has already allowed much more reproducible results to be achieved in their fabrication. Rather than relying upon fortuitously attaining the correct condition of the oxide by spontaneous oxidation, one can now control it with a high degree of success.

6.9. *Anodic oxides and corrosion*

Anodic oxidation is an industrially important method of finishing metals (such as aluminium) and of preparing oxide films of controlled thickness. Some of these films have sufficiently good dielectric properties to allow capacitor fabrication, while others display a high degree of electronic conduction. Often this conduction has been dismissed as due to pinholes and unspecified defects in the film, unworthy of much consideration. It is true that the conduction is localized (Ramasubramanian 1968) and we have seen in § 4.2 that anodic oxide films frequently display electroluminescence during their growth; this has been observed in the oxidation of Al, Zr, Ta, Bi, W, Zr and Zn (Mochniak 1966). At low voltages this luminescence appears to be uniform, but at higher voltages a large number of bright pinpoints are seen. The ultimate breakdown of anodic oxide films under electric stress takes place at localized regions, and signs of intense heating appear at these spots.

Revesz and Zaininger (1968) consider that in certain materials the high electric field present during anodization creates conducting channels through the oxide, these channels having a greater degree of order than the surrounding amorphous matrix. Revesz (1967) has discussed a possible conduction mechanism in anodic SiO_2, based on π-electron overlap in chains of Si–O–Si–O atoms with an unusual bonding. This is precisely the mechanism put forward, among others, independently by Dearnaley (1967), to explain conduction in SiO_x and other oxides.

Corrosion is, of course, the undesired result of oxidation or other forms of chemical attack. The process involves the transport of atoms through the corroded layer to the point of interaction, and as a result electrochemical barriers are often established which inhibit further reaction. The electrical conductivity of the layer then plays a part in determining the rate of corrosion, and therefore the possibility of 'forming' of oxide and other films may influence corrosion and particularly the initiation and spreading of localized attack. Thus Shirvington and Cox (1970) have studied the oxidation of zirconium alloys. Oxide growth on Zr proceeds almost entirely by migration of oxygen ions through the oxide film to the metal surface, and this ionic current is balanced by a flow of electrons in the opposite direction. Most earlier investigators assumed that the rate of film growth was determined by the ion flow, but Shirvington and Cox showed that, at all film thicknesses, the rate-determining process is electronic conduction. Conduction through the oxide was found to be localized, and the inhomogeneities were attributed to impurities present in the oxide layer. The conclusion is drawn that the presence of localized routes for electron transport in oxide films where the bulk oxide is a good insulator can extend the range of thicknesses over which the oxidation rate is controlled by electron flow to much greater thicknesses than had been previously envisaged. Similar results were anticipated in the oxidation of niobium and tantalum. Another interesting observation has been that evaporated metal layers can increase the oxidation rate of zirconium specimens (Fiegna and Weisgerber 1968). Cox and Phillips (1970) found that gold layers increase oxidation by a factor of between 2·4 and 2·7 at 500 °C. The

effect may be due to increased surface conduction resulting from the metal layer, or possibly to an enhancement of conduction in the oxide via something akin to forming. Gold is one of the metals which best allows forming, and it would be interesting to substitute aluminium (which inhibits forming) in such experiments.

6.10. *Anodic electret microphones*

Davies and Collins (1969) have described a simple microphone based upon the behaviour of anodic aluminium oxide. Plane foils of pure aluminium 15 μm in thickness were anodized to an oxide thickness of about 1300 Å. The foils were then placed in contact with a roughened metal back plate, and in some devices the compliance to incident sound waves of the air layer between oxide and back plate was increased by connecting the air layer to a closed volume of air behind the back plate. Such microphones gave open-circuit voltages of a few mV at normal voice levels.

Davies and Collins pointed out that this property implies that the anodic oxide is an electret; that is, it possesses a permanent dipole moment as a permanently stored electric charge. This charge must be negative and its amount, of the order 10^{17} ions cm^{-3} in the layer, was found to vary with the anodizing voltage, the subsequent heat treatment and the time elapsed after anodization.

Dearnaley (1970) suggested that, if the ideas discussed in § 6.9 are correct and conducting filaments of some type exist in anodic oxide layers, it is reasonable to expect that, during anodization, some electrons will have been scattered or will tunnel out of the filaments to occupy traps in the surrounding amorphous matrix. Such trapped electrons may give rise to very long-lived negative stored charges which will gradually become neutralized via normal conduction processes, but which can be accelerated by heat treatment. Here the stored charge which had deleterious effects in transistors (§ 6.7) can prove useful in a device structure. Moreover, the behaviour of such microphones offers another way of studying the electrical properties of anodic oxide films.

7. Conclusions

We have seen that there is evidence for believing that electrical phenomena in amorphous oxide and other films are of wide significance in electronic devices, both new and old, and in the processes of corrosion and anodic oxidation. We have attempted in this review to draw together a range of diverse effects which appear to have aspects in common, so that observations made in one field can shed light on the behaviour in another. In doing so, we have been somewhat selective in focusing attention on those oxide systems which exhibit voltage-controlled negative resistance, while discussing more briefly the other classes of behaviour. In outlining the various models for conduction and other phenomena in such systems, we have seen that no single model can yet explain all the effects observed. It is not yet clear to what extent avalanche processes are involved, or to what extent ionic rearrangement is occurring within the oxide. The role of impurities in controlling the behaviour is not properly established, but it appears that similar negative resistance phenomena occur even in very pure oxides—this point needs to be verified.

Perhaps the most puzzling and inexplicable results obtained so far are those of Hickmott on thin-film triodes, and it would be interesting to extend this work. It is necessary, also, to investigate further the conditions which best favour the forming

of oxides, and to verify whether it is the state of the anode–oxide interface which is of prime importance. This would assist in determining the physical mechanism of conduction in formed layers and would be relevant to several electronic devices.

The oxidized state of a solid surface occurs so frequently in our environment that an appreciation of its electrical and electrochemical behaviour is often important. The problems involved are challenging, both from a theoretical and experimental point of view but a start has been made in attacking them. We hope that the present article will stimulate discussion and encourage further work in this general field.

References

ALPERT, D., LEE, D. A., LYMAN, E. M., and TOMASCHKE, H. E., 1964, *J. Vac. Sci. Tech.*, **1**, 35–50.

ANDREWS, P. T., 1962, *Proc. Symp. on Nucl. Inst.*, Ed. J. B. Birks, (London: Heywood and Co.), pp. 93–7.

ARGALL, F., 1968, *Solid St. Electron.*, **11**, 535–41.

BARRIAC, C., GIRAUD-HIRAUD, F., PINARD, P., and DAVOINE, F., 1966, *C. R. Acad. Sci., Paris*, **262**, 900–2.

BARRIAC, C., PINARD, P., and DAVOINE, F., 1968, *C. R. Acad. Sci., Paris*, **266**, 423–6.

BARRIAC, C., PINARD, P., and DAVOINE, F., 1969, *Phys. Stat. Solidi*, **34**, 621–33.

BEAM, W. R., and ARMSTRONG, A. L., 1964, *Proc. IEEE*, **52**, 300–1.

BOER, K. W., 1969 a, *Phys. Stat. Solidi*, **34**, 721–32.

BOER, K. W., 1969 b, *Phys. Stat. Solidi*, **34**, 733–40.

BOER, K. W., and OVSHINSKY, S. R., 1970, *J. appl. Phys.*, **41**, 2675–81.

BRODSKY, M. H., and TITLE, R. S., 1969, *Phys. Rev. Lett.*, **23**, 581–5.

BUDENSTEIN, P. B., HAYES, P. J., SMYTH, J. L., and SMITH, W. B., 1969, *J. Vac. Sci. Tech.*, **6**, 289–303.

CHOPRA, K. L., 1963, *Proc. IEEE*, **51**, 941 and 1784–5.

CHOPRA, K. L., 1965, *J. appl. Phys.*, **36**, 184–7.

COHEN, M. H., 1968, *Rev. mod. Phys.*, **40**, 839–44.

COHEN, M. H., FRITZSCHE, H., and OVSHINSKY, S. R., 1969, *Phys. Rev. Lett.*, **22**, 1065–8.

COOK, E. C. L., 1970, *J. appl. Phys.*, **41**, 551–4.

COPE, R. G., and PENN, A. W., 1968, *J. Phys.* D: *Appl. Phys.*, **1**, 161–8.

COX, B., and PHILLIPS, D., 1970, *AECL Report* PR-CMa-10.

DAHLKE, W. E., 1967, *Appl. Phys. Lett.*, **10**, 261–2.

DAVIES, L. W., and COLLINS, R. E., 1969, *Electron. Lett.*, **5**, 462–3.

DEARNALEY, G., 1967, *Phys. Lett.*, **25A**, 760–1.

DEARNALEY, G., 1969, *Thin Solid Films*, **3**, 161–74.

DEARNALEY, G., 1970, *Int. J. Electron.*, **29**, 299.

DEARNALEY, G., MORGAN, D. V., and STONEHAM, A. M., 1970, *J. Non-Cryst. Solids*, **4**, 593–611.

DELORD, J. F., JOHNSON, K. E., COAVERT, R. E., PISTOULET, B., ROUSTAND, G., and ROUZEYRE, M., 1967, *Appl. Phys. Lett.*, **11**, 287–9.

DONNELLY, J. P., FOYT, A. G., LINDLEY, W. T., and ISELER, G. W., 1970, *Solid St. Electron.*, **13**, 755–8.

DRAKE, C. F., SCANLAN, I. F., and ENGEL, A., 1969, *Phys. Stat. Solidi*, **32**, 193–208.

DUKE, C. B., 1969, *Tunnelling in Solids*, (New York: Academic Press).

ECKERTOVÀ, L., 1966, *Phys. Stat. Solidi*, **18**, 3–40.

FIEGNA, A., and WEISGERBER, P., 1968, *J. Electron. Soc.*, **115**, 369–71.

FILARETOV, G. A., and STAFEEV, V. I., 1968, *Sov. Phys.—Semicond.*, **1**, 913–14.

FILARETOV, G. A., STAFEEV, V. I., and BUBNOV, YU. Z., 1967, *Sov. Phys.—Semicond.*, **1**, 1149–52.

FILARETOV, G. A., STAFEEV, V. I., BUBNOV, YU. Z., and CHIZHIK, T. N., 1968, *Sov. Phys.—Semicond.*, **1**, 1242–4.

FISCHER, A. G., 1963, *J. Electrochem. Soc.*, **110**, 733–48.

FISCHER, A. G., 1969, in *Radiative Recombination in Semiconductors* (Paris: Dunod), pp. 266–9.

FISCHER, A. G., and MOSS, H. J., 1963, *J. appl. Phys.*, **34**, 2112–13.

FISHER, J. C., and GIAEVER, I., 1961, *J. appl. Phys.*, **82**, 172–7.

FLINN, I., BEW, G., and BERZ, F., 1967, *Sol. St. Electron.*, **10**, 833–45.
FORLANI, F., and MINNAJA, N., 1964, *Phys. Stat. Solidi*, **4**, 311–24.
FRITSCHE, H., 1969, *IBM J. Res. Dev.*, **13**, 515–21.
FRITSCHE, H., and OVSHINSKY, S. R., 1970, *J. Non-Cryst. Solids*, **4**, 464–79.
FRÖHLICH, H., 1947, *Proc. R. Soc.*, **A188**, 521–42.
FROMHOLD, A. T., and COOK, E. L., 1967 a, *Phys. Rev.*, **158**, 600–12.
FROMHOLD, A. T., and COOK, E. L., 1967 b, *Phys. Rev.*, **163**, 650–64.
GIBBONS, J. F., and BEADLE, W. E., 1964, *Sol. St. Electron.*, **7**, 785–97.
GOETZBERGER, A., and SZE, S. M., 1969, *Appl. Solid-St. Sci.*, **1**, 154–238.
GOULD, R. D., and COLLINS, R. A., 1970, *Appl. Phys. Lett.*, **16**, 393–5.
GRANNEMANN, W. W., BECHTEL, R., and DAVIES, G., 1960, *Univ. of New Mexico Report* PR-13.
GREENE, P. D., BUSH, E. L., and RAWLINGS, I. R., 1968, *Proc. Symp. on Deposited Thin Film Dielectric Materials, Montreal*, Ed. F. Vratny (New York: The Electrochemical Society, 1969), p. 167–85.
GREGOR, L. V., 1966, *Physics of Thin Films*, **3**, 131–62.
GRIGOROVICI, R., and MANAILA, R., 1967, *Thin Solid Films*, **1**, 343–52.
HANNAM, H. J., 1956, *PhD Thesis*, University of Minnesota.
HANNEY, N. B., MACNAIR, D., and WHITE, A. H., 1949, *J. appl. Phys.*, **20**, 669–81.
HARROP, P. J., and WANKLYN, J. N., 1965, *Br. J. appl. Phys.*, **16**, 153–62.
HARROP, P. J., and WANKLYN, J. N., 1967, *J. nucl. Mater.*, **22**, 350–1.
HEINZE, W., and WAGENER, S., 1939, *Z. Tech. Phys.*, **20**, 17.
HIATT, W. R., and HICKMOTT, T. W., 1965, *Appl. Phys. Lett.*, **6**, 106–8.
HICKMOTT, T. W., 1962, *J. appl. Phys.*, **33**, 2669–82.
HICKMOTT, T. W., 1963, *J. appl. Phys.*, **34**, 1569–70.
HICKMOTT, T. W., 1964 a, *J. appl. Phys.*, **35**, 2118–22.
HICKMOTT, T. W., 1964 b, *J. appl. Phys.*, **35**, 2679–89.
HICKMOTT, T. W., 1965 a, *J. appl. Phys.*, **36**, 1569.
HICKMOTT, T. W., 1965 b, *J. appl. Phys.*, **36**, 1885–96.
HICKMOTT, T. W., 1965 c, *Appl. Phys. Lett.*, **6**, 106–8.
HICKMOTT, T. W., 1966 a, *J. Electrochem. Soc.*, **113**, 1223–5.
HICKMOTT, T. W., 1966 b, *J. appl. Phys.*, **37**, 4380–8.
HICKMOTT, T. W., 1969 a, *J. Vac. Sci. Tech.*, **6**, 828–33.
HICKMOTT, T. W., 1969 b, *Thin Solid Films*, **3**, 85–108.
HICKMOTT, T. W., 1971, *J. appl. Phys.*, to be published.
HICKMOTT, T. W., and HIATT, W. R.,
HILL, R. M., 1967, *Thin Solid Films*, **1**, 39–68.
HILL, R. M., 1971, *Thin Solid Films*, to be published.
HILSUM, C., 1969, *Physics Today*, **22**, pp. 89–91.
HOFSTEIN, S. R., 1966, *IEEE Trans. Electron. Dev.*, **ED-13**, 222–37.
HOFSTEIN, S. R., 1967, *Appl. Phys. Lett.*, **10**, 291–3.
HOU, S. L., and MARLEY, J. A., 1970, *Appl. Phys. Lett.*, **16**, 467–9.
IIZIMA, S., SUGI, M., KIKUCHI, M., and TANAKA, K., 1970, *Solid-St. Commun.*, **8**, 153–5.
JAKLEVIC, R. C., DONALD, D. K., LAMBE, J., and VASSELL, W. C., 1963, *Appl. Phys. Lett.*, **2**, 7–9.
JONES, M. D., and MEE, C. H. B., 1964, *Br. J. appl. Phys.*, **15**, 561–5.
JONSCHER, A., 1967, *Thin Solid Films*, **1**, 213–34.
KANTER, H., and FEIBELMAN, W. A., 1962, *J. appl. Phys.*, **33**, 3580–8.
KERR, D. R., LOGAN, J. S., BURKHARDT, P. J., and PLISKIN, W. A., 1964, *IBM J. Res. Dev.*, **8**, 376–84.
KIRCHER, J. F., and BOWMAN, R. E., 1964, *Effects of Irradiation on Materials and Components* (New York: Reinhold).
KLEIN, N., 1969, *J. Electrochem. Soc.*, **116**, 963–72.
KLEIN, N., and GATNI, H., 1966, *IEEE Trans. Elect. Dev.*, **ED-13**, 281–9.
KREYNINA, G. S., 1962, *Radio Eng. Electron.*, **7**, 1949–53.
KREYNINA, G. S., SELIVANOV, L. N., and SHUMSKAYA, T. I., 1960, *Radio Eng. Electron.*, **5**, 219.
LAMPERT, M. A., and MARK, P., 1970, *Current Injection in Solids* (New York: Academic Press).
LAVERTY, S. J., and RYAN, W. D., 1969, *Int. J. Electron.*, **26**, 471–6.
LEE, P. A., SAID, G., and DAVIS, R., 1969, *Solid St. Electron.*, **7**, 1359–61.
LEWOWSKI, T., SENDECKI, S., and SUJAK, B., 1965, *Acta Phys. Polon.*, **28**, 343–48.
LOMAX, R. W., and SIMMONS, J. G., 1968, *Radio Elect. Eng.*, **35**, 265–72.

LOPEZ, A. D., 1966, *J. Electrochem. Soc.*, **113**, 89–90.
MCWHORTER, A. L., 1956, *Semiconductor Surface Physics*, p. 207 (Philadelphia: University of Pennsylvania Press).
MALE, J. C., 1970, *Electron. Lett.*, **6**, 91–2.
MALE, J. C., and WARREN, A. C., 1970, *New Scientist*, **47**, 128–9.
MECKLENBURG, W., 1943, *Z. Phys.*, **120**, 21–30.
MANN, H T, 1964, *J. appl. Phys.*, **35**, 2173–9.
MEAD, C. A., 1961, *J. appl. Phys.*, **32**, 646–52.
MILLER, A., and ABRAHAMS, E., 1960, *Phys. Rev.*, **120**, 745–55.
MOCHNIAK, J., 1966, *Acta phys. polon.*, **30**, 559–66.
MOTT, N. F., 1968 a, *Rev. mod. Phys.*, **40**, 677–83.
MOTT, N. F., 1968 b, *Phil. Mag.*, **17**, 1259–68.
MOTT, N. F., 1968 c, *Notes on Ovonic and Memory Switching Devices*, unpublished.
MOTT, N. F., 1969, *Contemp. Phys.*, **10**, 125–38.
MOTT, N. F., and DAVIES, E. A., 1968, *Phil. Mag.*, **17**, 1269–84.
MOTT, N. F., and GURNEY, R. N., 1948, *Electronic Processes in Ionic Crystals* (Oxford University Press).
NELSON, R. L., and TENCH, A. J., 1964, *J. chem. Phys.*, **40**, 2736–7.
NELSON, R. L., TENCH, A. J., and HARMSWORTH, B. J., 1967, *Trans. Faraday Soc.*, **63**, 1427–46.
NIELSEN, P. H., and BASHARA, N. M., 1964, *IEEE Trans. Electron Dev.*, **ED-11**, 243–4.
O'DWYER, J. J., 1964, *Theory of Dielectric Breakdown* (Oxford University Press).
OHTA, K., WISHIDA, J., and HAYASHI, T., 1968, *Jap. J. appl. Phys.*, **7**, 982.
OVSHINSKY, S. R., 1968, *Phys. Rev. Lett.*, **21**, 1450–3.
PARK, K. C., and BASAVAIAH, S., 1970, *J. Non-Cryst. Solids*, **2**, 284–91.
PAULE, W., 1969, *Physics Today*, **22**, pp. 97–101.
PEARSON, A. D., 1969, *IBM J. Res. Dev.*, **13**, 510–14.
PINTO, R., and SHAHA, B. M., 1968, *Jap. J. appl. Phys.*, **7**, 1542–3.
POLLACK, S. R., FREIBAG, W. O., and MORRIS, C. E., 1963, *Electrochem. Technol.*, **1**, 96–100.
RALPH, J. E., and TANSLEY, T. L., 1970, *J. Phys.* D: *Appl. Phys.*, **3**, 620–3.
RAMASUBRAMANIAN, N., 1968, *Atomic Energy of Canada Ltd. Report* AECL-3082.
REVESZ, A. G., 1967, *Phys. Stat. Solidi*, **24**, 115–26.
REVESZ, A. G., and ZAININGER, H. H., 1968, *RCA Rev.*, **29**, 22–76.
RIDLEY, B. K., 1963, *Proc. Phys. Soc.*, **82**, 945–66.
RUNCIMAN, W. A., 1968, *Solid St. Commun.*, **6**, 537.
SHAW, R. F., 1968, *New Scientist*, **40**, p. 428.
SHAW, R. F., 1968, *New Scientist*, **40**, p. 508.
SHIRVINGTON, P. J., and COX, B., 1970, *J. nucl. Mater*, **35**, 211–22.
SIE, C. H., 1970, *J. Non-Cryst. Solids*, **4**, 548–53.
SIFFERT, P., LAUSTRIAT, G., and COCHE, A., 1964, *IEEE Trans, Nucl Sci.*, **NS-11**, No. 3, 244–8.
SIMMONS, J. G., 1965, *Phys. Rev. Lett.*, **15**, 967–8.
SIMMONS, J. G., 1970, *Contemp. Phys.*, **11**, 21–41.
SIMMONS, J. G., and VERDERBER, R. R., 1966 a, *Proc. Conf. IERE–IEF on 'Applications of Thin Films in Electronic Engineering'* (London, 11th–14th July, 1966).
SIMMONS, J. G., and VERDERBER, R. R., 1966 b, *Erratum*, **17**, 1274.
SIMMONS, J. G., and VERDERBER, R. R., 1967 a, *Proc. R. Soc.*, **A301**, 77–102.
SIMMONS, J. G., and VERDERBER, R. R., 1967 b, *Radio Electron. Eng.*, **34**, 81–9.
SIMMONS, J. G., and VERDERBER, R. R., 1967 c, *Appl. Phys. Lett.*, **10**, 197–9.
SIMMONS, J. G., and VERDERBER, R. R., 1968, *Radio Elect. Eng.*, **35**, 265–72.
SIMMONS, J. G., VERDERBER, R. R., LYTOLLER, J., and LOMAX, R. M., 1966, *Phys. Rev. Lett.*, **17**, 675–7.
SMITH, P. A., and MILES, J. L., 1963, *J. Electrochem. Soc.*, **110**, 1240–5.
SNOW, E. H., GROVE, A. S., DEAL, B. E., and SAH, C. I., 1965, *J. appl. Phys.*, **36**, 1664–73.
SRC, 1970, *The Physics of Amorphous Materials* (London: Science Research Council).
SUGANO, T., HOH, K., KUDO, K., and HISHINUMA, N., 1968, *Jap. J. appl. Phys.*, **7**, 715–30.
TANAKA, K., IIZIMA, S., SUGI, M., and KIKUCHI, M., 1970, *Solid St. Commun.*, **8**, 75.
TANSLEY, T. L., and RALPH, J. E., 1970, *J. Phys.* D: *appl. Phys.*, **3**, 807–11.
TOSSER, A., 1968, *J. Phys., Paris*, **29**, 671–83.
TOSSER, A., and THUREAU, P., 1967, *J. Phys., Paris*, **28**, 642–52.
VARKER, C. J., and JULEFF, E. M., 1967, *Proc. IEEE*, **56**, 728–9.

VERDERBER, R. R., and SIMMONS, J. G., 1967, *Radio Electr. Engr.*, **33**, 347–51.
VERDERBER, R. R., SIMMONS, J. G., and EALES, B., 1967, *Phil. Mag.*, **1**, 1049–68.
VERWEY, J. F., 1969, *Appl. Phys. Lett.*, **15**, 270–2.
VOLKOV, A. F., and KOGAN, S. M., 1969, *Uspekhi S.P.*, **11**, 881–903.
WARREN, A. C., 1969, *Electron. Lett.*, **5**, 461; **5**, 609.
WARREN, A. C., and MALE, J. C., 1970, *Electron. Lett.*, **6**, 567–9.
WEISER, K., and WOODS, J. F., 1965, *Appl. Phys. Lett.*, **7**, 225–8.
WHITEHEAD, S., 1951, *Dielectric Breakdown in Solids* (Oxford University Press).
WORTHING, F. L., 1968, *J. Electrochem. Soc.*, **115**, 88–92.
YEE, J. H., and CONDAS, G. A., 1968, *Solid St. Electron.*, **11**, 419–22.