ELECTRONIC AND OPTICAL PROPERTIES OF AMORPHOUS SEMICONDUCTORS:
THEIR PRINCIPLES AND APPLICATIONS

Alan E Owen PhD, FRSE

PART 2: APPLICATIONS

Thesis submitted for the degree of
Doctor of Science
University of Edinburgh
1994
CONTENTS - PART 2

APPLICATIONS

Index to Cited Publications:
  Applications................................................................. (xv)

Cited Publications: Applications........................................... 341-576

APPENDIX:

APPLICATIONS

(Where there are co-authors, the number in parenthesis indicates my position in the list of authors.)

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>


ELECTRONICALLY-ASSISTED THERMAL BREAKDOWN IN CHALCOGENIDE GLASSES

J. M. ROBERTSON and A. E. OWEN

Department of Electrical Engineering, University of Edinburgh, Edinburgh, Scotland

The threshold characteristics of switches prepared from compositions in the As$_2$Te$_3$ + Si system have been examined. With low resistivity compositions (i.e. low Si content) the characteristics are adequately described by a simple thermal model, irrespective of the geometry of the device. As the resistivity is increased (i.e. increased Si content) the field-dependence of the conductivity becomes significant and has an important effect on the threshold characteristics of thin sandwich devices.

1. Introduction

Chalcogenide glasses have low thermal conductivities and their electrical conductivity increases rapidly with temperature so thermal runaway must always be considered as a possible high-field breakdown mechanism. Alternatively, if switching cannot be completely attributed to thermal runaway, the thermal effects of the current and voltage conditions which prevail at threshold may be significant enough to influence the device characteristics.

The simplest form of thermal breakdown occurs when $\sigma = \sigma(T)^1$). In glasses of higher resistivity the field dependence of conductivity is more important, i.e. $\sigma = \sigma(T, E)$. When this is taken into account, with a suitable expression for $\sigma$, good quantitative prediction can be obtained of switching characteristics$^2,3)$. This may be considered as a form of “electronically-assisted” thermal breakdown. In this paper, the effects of a steady increase in glass resistivity are presented along with some consequences of this type of breakdown.

2. Sample preparation

In the system As$_2$Te$_3$ + Si, glass resistivity increases approximately logarithmically as the proportion of silicon is increased. The range is from 50 ohm m for As$_2$Te$_3$ to $10^8$ ohm m for (As$_2$Te$_3$)$_{60}$Si$_{40}$. On evaporation, dissociation occurs and the films are deficient in silicon with respect to the bulk glass. However, X-ray fluorescence analysis has shown that the As:Te ratio in the thin film is still approximately 0.67 so the general form of the composition does not change.
All samples were deposited on Corning 7059 glass. In most cases there was a simple cross-over geometry with a device area of $2 \times 10^{-7} \text{ m}^2$ and glass film thickness about 1 $\mu$m. Gold electrodes were used.

3. Thermal breakdown

Conductivity may be written in the form:

$$\sigma(E, T) = \sigma_a \exp \left( \frac{\Delta T}{T_1} \right) f(E),$$

where $\sigma_a$ is the conductivity when the temperature deviation from ambient, $\Delta T$, is zero and $T_1$ is a constant, typically 20 K for usual chalcogenide switching glasses; $f(E)$ expresses the field modification of conductivity and must be determined experimentally. This equation is quite accurate for small values of $\Delta T$. The steady state heat dissipation may be expressed by the usual Newtonian expression:

$$VI = \Gamma \Delta T,$$

where $\Gamma$ is the thermal conductance of the film. If eq. (1) is substituted into eq. (2) and differentiated with respect to temperature, the breakdown condition may be established:

$$V_{th}^2 f(E) = \frac{1}{e} R_s \Gamma T_1.$$

With a constant device geometry, $\Gamma$ is approximately constant, so $R_s$ the low field sample resistance is the most controllable parameter.

4. Pulse measurements

When a rectangular voltage pulse is applied to a sample the current rises from an initial value $I_0$ to a final value which is $\Delta I$ greater. This is shown in fig. 1. The capacitive spike is of short duration and the increase in current is characterised by a time constant $\tau$.

For three types of sample: (a) thick (100 $\mu$m) with sandwich electrodes, (b) thin film with coplanar electrodes 100 $\mu$m apart and (c) thin (1 $\mu$m thick) with sandwich electrodes and very low resistivity glass (<60 ohm m), the variation of $I_0$ with voltage was ohmic. Just below $V_{th}$, the magnitudes of $\Delta I$ and $\tau$ were appropriate to values expected of the device were being internally heated so it is reasonable to take $\Delta I$ as the Joule heating component of current. The variation of $(I_0 + \Delta I)$ with voltage followed the usual ohmic
then exponential relation expected for simple thermal breakdown where \( \sigma = \sigma(T) \) only.

As glasses of higher and higher resistivity were used, the breakdown field, \( E_{th} \), increased slowly but the current pulse remained of the same general form as that shown in fig. 1. Characteristics for a 300 ohm film are shown in fig. 2. The \( I_0 - V \) relation can be approximately described by:

\[
\sigma(E) = \sigma_a \exp \left( \frac{V}{V_1} \right). \tag{4}
\]

The variation of \( V_1 \) with thickness is shown in fig. 3. The slope, \( E_1 \), is \( 3.3 \times 10^6 \text{ V m}^{-1} \); \( E_1 \) increases slightly with film resistivity and decreases slowly with increasing ambient temperature. When eq. (4) is substituted into the thermal breakdown eq. (3) the threshold voltage is given by:

\[
V_{th} = V_1 \log \left( \frac{R_T T_1}{e V_{th}^2} \right). \tag{5}
\]

This equation gives a good quantitative prediction of \( V_{th} \) over a wide range of glass film resistivities from the simple ohmic case with \( V_{th} < V_1 \) and \( \rho < 50 \text{ ohm m} \) to the electronically-assisted case with \( \rho < 1000 \text{ ohm m} \) and where \( V_{th} \) is limited to only 3–5 times \( V_1 \). The usual temperature, thickness, and delay time expressions also follow the analysis used to derive eq. (5).

**Fig. 1.** Schematic forms of the voltage and current pulses just below \( V_{th} \); \( \Delta I \) represents the increase in current due to Joule heating.
Fig. 2. Current–voltage characteristics (log-linear) for $I_0$ and $(I_0 + \Delta I)$ of fig. 1 in a 3000 ohm film.

Fig. 3. The constant $V_1$ of eq. (4) as a function of electrode separation.
ELECTRONICALLY-ASSISTED THERMAL BREAKDOWN

5. Localised breakdown

At fields greater than $E_1$, thermal breakdown occurs preferentially at any small region of slightly increased field. Thus the position of a conducting filament may be determined by physical factors such as glass or substrate irregularities or stresses over lower electrode edges.

As glass films of higher resistivity are used, one of the most noticeable effects is that the Joule heating component in the current waveform decreases. This we suggest is due to more localised breakdown as the $\sigma(E)$ term becomes more dominant.

The thermal time constant for a sandwich device, area $A$, is approximately given by:

$$\tau = CAd / \Gamma$$

where $C$ is the heat capacity of the glass. For our devices with $\Gamma = 10 \text{ mW K}^{-1}$ and $C = 2 \times 10^6 \text{J m}^{-3} \text{K}^{-1}$, $\tau = 40 \mu \text{s}$ which is the value measured for samples made from low resistivity glass films. The thermal conductance is determined largely by the substrate and electrodes so any reduction in the effective heated area should result in a corresponding reduction in $\tau$. This occurs when either the contact area is reduced, i.e. $\tau \propto A$, or if glass conductivity is increased i.e. $\tau \propto \sigma$. These results have been combined in fig. 4 but in this case $\tau$ has been plotted against $R/d$.

![Graph](image)

**Fig. 4.** The thermal time constant, measured by a pulse experiment of the kind illustrated in fig. 1, as a function of the ratio of device resistance ($R$) to electrode separation ($d$).
6. Conclusions

As glass film resistivity is increased, the field-dependence of conductivity becomes steadily more important. There is a steady change from the simple thermal breakdown case with an ohmic sample to the electronically-assisted case. At all times, however, the Joule heating contribution to current can be identified and eliminated from any conductivity expression used to calculate breakdown conditions. With high resistivity films local breakdown is significant. Although other effects or processes may be present in these devices, particularly when the voltage approaches $V_{th}$, the thermal breakdown analysis is sufficient to determine the major functional relationships and their numerical values in the devices we have studied.

References

Abstract—Evidence for electronic conduction in the preswitching region of chalcogenide glasses is outlined and discussed in terms of several simple models. Electronic processes likely to contribute to the switching process itself are reviewed and compared, where possible, with the predictions of relevant models from conventional semiconductor physics.

I. INTRODUCTION

If an increasing electric field is applied to a dielectric, an unstable situation will eventually arise where the rate of energy input exceeds the capacity of the system to dissipate it. The results can be catastrophic and after breakdown the dielectric may be left in a conducting (short-circuit breakdown) or nonconducting (open-circuit breakdown) state. In some circumstances breakdown may lead to a conducting state, or ON state, which requires a small holding current and voltage to sustain it. The original high resistance, or OFF state, is recovered if the sustaining voltage is removed. In other cases it is possible to establish a permanent ON state that persists in the absence of any voltage, but the OFF state may be recoverable by the application of a suitable current pulse. These two cases of reversible breakdown are described as threshold switching and memory switching, respectively. In all cases reported so far, switching in chalcogenide glasses is associated, either explicitly or implicitly, with a region of current-controlled negative resistance (CCNR). An essential condition for CCNR is a positive feedback mechanism which, in the region of instability, allows the system to carry the same, or larger currents, with smaller voltages. In the thermal mechanisms of switching discussed by Warren [1], the feedback loop is

\[ \text{high field} \rightarrow \text{increased current} \rightarrow \text{increased power dissipation} \]

\[ \rightarrow \text{greater conductivity} \leftarrow \text{internal temperature rise}. \]

The essential link is the sensitivity of the conductivity to temperature. Other examples, relevant to electronic processes, will be mentioned later (Section IV-B, C). Whether the result is breakdown or switching, the conditions leading to the electrical instability can usually be formally described by an energy-balance equation [2], [3]

\[ A(T, E, \alpha) = B(T, \alpha) \]  

where \( A \) is the rate at which energy is gained from the field \( E \) at temperature \( T \) and \( B \) is the rate at which it is absorbed or dissipated. The parameter \( \alpha \) is introduced to denote any other relevant property of the current carriers pertinent to a particular situation.

The variation of the two sides of this energy-balance equation will generally be as shown in Fig. 1 in which
the abscissa represents some appropriate parameter such as temperature $T$, energy $E$, or injected charge $Q_{inj}$ (see below).

The rate of gain of energy from the field will normally be given by

$$ A = \sigma(T, E)E^2 $$

or some straightforward variant of this; the conductivity $\sigma(T, E)$ is generally a function of temperature and field. With a sufficiently large conductivity that is very temperature dependent, the dielectric will heat up through Joule heating, and if thermal processes are dominant then

$$ B = \left( \frac{\rho C}{K} \right) \frac{\partial T}{\partial t} - \nabla^2 T. \quad (3) $$

The thermal energy-balance condition $A = B$ is then the same as in [1, eq. (1)] (the same notation is used). Thus in Fig. 1 the abscissa represents temperature and the field $E_2$ would be the point at which instability sets in. It is worth noting that for a thermal instability, it is not necessary that the conductivity be a function of field [1]. The isothermal conductivity may be ohmic at all times, but the dynamic conductance will not be a constant because of Joule heating.

Many other conditions of instability leading to breakdown have, of course, been investigated. For intrinsic electronic breakdown in the single electron approximation, for example, the function $A$ is of the form [3]

$$ A = \sigma(\varepsilon)E^2 = \frac{\varepsilon^2 E^2 \tau(\varepsilon)}{m} \quad (4) $$

where $\tau$ is the mean time between collisions and is a function of energy and $N_0$ is the average number of phonons of frequency $\omega$. Thus in Fig. 1 the abscissa represents energy $\varepsilon$, and the field $E_2$ again corresponds to the onset of instability.

If the electron density is high (e.g., at high temperatures) electron-electron collisions may become important and the single-electron theories are no longer valid. With a high collision rate and a rate of exchange of energy due to electron-phonon interactions, the electron energy-distribution function is Maxwellian with a mean temperature $T_{ei}$ above the ambient temperature $T_{ao}$ or the lattice temperature $T$. This means that the energy rate of change terms $A$ and $B$ can be averaged to $\bar{A}$ and $\bar{B}$ having a similar functional dependence so that the general picture of Fig. 1 remains valid at least up to energies $\varepsilon_{inj}$ at the maximum of the $B$ curve in Fig. 1. This "collective" description has a strong analogy with thermal breakdown and the abscissa in Fig. 1 corresponds to the electron temperature $T_{ei}$. In this case, therefore, it is the steady-state electron temperature that rises until, at a critical value, no equilibrium is possible and instability sets in. Near $\varepsilon_{inj}$ however, the electron-phonon energy transfer increases, implying an increase in the internal temperature of the material with consequential thermal effects.

Of the many other breakdown and/or switching mechanisms, at this stage it is worth mentioning double injection [4]. In this case the energy input is either stored in the dielectric as injected charge $Q_{inj}$ or is lost through recombination. An energy-balance equation could be written, therefore, as

$$ \sigma(T, E)E^2 = \frac{Q_{inj}}{D\varepsilon} \frac{dQ_{inj}}{dt} - j_{recombination} \psi_{int} \quad (6) $$

where $\varepsilon$ is the dielectric constant, $D$ is a suitable geometrical factor, and $j_{recombination}$ is the recombination current driven by some appropriate internal potential $\psi_{int}$. In this case the abscissa of Fig. 1 would represent $Q_{inj}$.

In general, therefore, we can describe the condition of electrical instability, whatever its origin, in terms of a diagram such as Fig. 1. This will normally be preceded by a region in which the conductivity is field dependent (nonohmic). What follows, i.e., breakdown or switching, depends upon the properties of the dielectric and on the presence, or absence, of a suitable positive feedback process in the system. For threshold switching the material must be capable of carrying a much increased current, either uniformly or locally (i.e., in a filament), but reversing spontaneously to the original nonconducting OFF state when the sustaining or holding voltage is removed. For memory switching the dielectric material must be capable of changing in some way (e.g., an overall or localized change in the atomic or microscopic structure) into a permanent conducting state. but one
that can be reversed to the off state by a suitable current (energy) pulse. Obviously the system must also be able to absorb the reversing pulse without destructive breakdown.

II. ELECTROTHERMAL PROCESSES

Switching of either the threshold or memory kind has been observed in a wide variety of semiconducting glasses of the chalcogenide [5], [6] and transition metal oxide types [7]. The general form of switching characteristics for chalcogenide glass is shown by Warren [1]. He also gives typical values for threshold voltage $V_{th}$, delay time $t_d$, and describes how these parameters vary with temperature, field strength, and geometry. More detailed treatments of various aspects of switching behavior can be found in the proceedings of recent conferences [8]–[10]. The main point of issue has been to what extent switching can be attributed to thermal or electronic processes and both viewpoints were represented at these three conferences [8]–[10]. There have, in addition, been suggestions that some subtle structural rearrangement akin to a dielectric or ferroelectric transition may be responsible for switching [7], [11]. The purpose of this review is to consider what electronic mechanisms may be appropriate, and mechanisms that rely on structural changes will therefore not be considered.

It is not proposed to argue the pros and cons of thermal versus electronic theories of switching, but it should be recognized from the outset that there must be some convergence of the thermal and nonthermal viewpoints. The properties of semiconducting glasses are such that some thermal effects must be present; the question is—how much? It is now acknowledged, for example, that in chalcogenide switches with electrode separations of 10 $\mu$m or more, the threshold conditions can be quantitatively predicted from a solution of the thermal energy-balance equations (2) and (3) in a one-dimensional form with conductivity expressed only as a function of temperature, i.e., $\sigma = \sigma(T)$ [1], [12], [13].

With smaller electrode separations, however, there are significant changes [1] (e.g., in the temperature and thickness dependence of the threshold voltage) and this has been taken to indicate that electronic effects are predominant [13]. The simple thermal models are certainly no longer adequate, but thermal effects may still be present. Fig. 2(a), for example, shows the current-voltage dependence on a log-log scale for a thin film (1.7 $\mu$m) of glass of nominal composition $As_2Te_3Ge_4$ [14]. The pulse measurements were made at a repetition rate of 100 Hz and the current is that recorded at the end of the pulse. In the dc case and to a lesser extent in the 100-Hz ac measurements, a region of negative resistance was observed before the device switched. Such results are to be expected if self-heating occurs, but note that at voltages below the dc threshold, the characteristics are practically identical.

To establish the conduction mechanism at high fields, it is obviously necessary to measure the $I-V$ characteristics without the complication of self-heating, and it should not necessarily be assumed that pulse measure-
ments will do this—the pulse length has to be shorter than the thermal time constant for the active glass film. Fig. 2(b) shows $I-V$ curves plotted on a log-linear scale with current measured at the leading and trailing edges of a 70-μs pulse (50 ns rise time, 100-Hz repetition rate). The capacitance spike at the beginning of the pulse was very short and it was possible to extrapolate back to zero time to obtain the initial current value. The current rise during the pulse could not be conclusively attributed to Joule heating, but the power dissipated during one pulse was sufficient to give a temperature rise of several degrees, which in turn was sufficient to account for the current increase during the pulse [14]. Furthermore, the time taken for the current to reach a steady value was approximately the same as the estimated thermal time constant for the materials and geometry used. Fig. 2(b) emphasizes that characteristics with and without a Joule heating component of current are of almost identical form. This can make it very difficult to identify effects of Joule heating particularly in higher resistivity films where the heating component of current is smaller and the total current distribution may be nonuniform even in the prebreakdown condition [15]. The straight-line region of a log-linear plot is often very restricted. If the transition from an apparently ohmic to exponential characteristic occurs at a voltage $V_0$, then typically $V_0 + 3V_0$ at room temperature [14], [16]. Thus the voltage range over which experimental observations can be compared with theoretical models is often very limited and the interpretation is correspondingly uncertain. The "leading-edge" current characteristic of Fig. 2(b) shows noticeable nonohmic behavior typical of that seen in many chalcogenide films at fields as low as $10^9 - 10^{10} \text{ V/m}$. A recent paper by de Wit and Crevecoeur [17] suggests indeed that in As$_2$Te$_3$ the conductivity is nonohmic essentially down to zero fields. We thus return to this question of the onset of nonohmic conduction in a later section, but for the moment it is sufficient to note that at high fields an exponential dependence of current (or conductance) on voltage is often observed:

$$I = I_0 \exp \left( \frac{V}{V_0} \right)$$

(7a)

$$G = \frac{I}{V} = G_0 \exp \left( \frac{V}{V_0} \right)$$

(7b)

where $G_0$ is the zero-field conductance. For $V < V_0$ it may be very difficult, experimentally, to distinguish any deviation from ohmic behavior. The constant $V_0$ in (7a) and (7b) will not of course be the same for a given set of data and if they are to be interpreted in terms of some specific model, attention must be given to the appropriate form of plotting the data to arrive at $V_0$.

The establishment of a nonohmic characteristic that is unambiguously free from thermal effects allows two steps to be taken.

1) The conduction mechanism may be identified and possibly linked with an electronic switching mechanism. This is reviewed in Section III.

2) The thermal consequences of this conduction mechanism may be calculated. This approach, with conductivity a function of temperature and field, has been used recently by several authors [13], [15], [18], [19]. Their predictions show reasonable correspondence with experimental data, but while they show that field-assisted thermal runaway is sufficient to explain switching observations, they do not prove that switching is exclusively thermal in origin.

The role of the nonohmic conduction mechanism therefore becomes very important. We take the view that in a purely thermal process the exact mechanism is immaterial and an empirical relation is all that is required to predict the results. The nonohmic conductivity provides a means whereby energy is delivered to the switch at a faster rate than would be possible under ohmic conditions. If, on the other hand, there are features of switching that cannot be accounted for by a complete solution of the thermal energy-balance equation with all the parameters given their correct dependence on temperature, field, etc., then we must examine the conduction mechanism for sufficient conditions to allow some electronic or other nonthermal mechanism to make an intrinsic contribution to the switching process. In the following sections the likely electronic processes are discussed and, wherever possible, compared with experimental information.

The memory type of device will not be discussed explicitly. The formation of a permant on state in a chalcogenide switch involves the precipitation of a crystalline filament from the glassy matrix [20]. This is a problem involving the thermodynamic stability of the various phases involved (glass, liquid, and possible crystalline compounds) and the kinetics of nucleation and crystal growth. Threshold switching always precedes memory switching, however, so any mechanism established for the former is implicitly applicable to the latter.

III. The High-Field OFF State

A. Experimental Data

There is surprisingly little published data on the $I-V$ characteristics of the kind of chalcogenide compositions most often used in switching devices, and hence for the purposes of illustration in the following discussion data obtained by one of the authors [14] will be used. Typical results for a glass of approximate composition (As$_2$Te$_3$)$_{0.2}$Si$_{0.8}$ are shown in Fig. 3(a) for three different thicknesses and a limited temperature range. These data conform approximately to (7). Fig. 3(b), (c) reproduce, for comparison, the data of Walsh et al. [11b] on the composition (weight percent): Te, 49 percent; As, 33 percent; Ge, 6 percent; Si, 3 percent; Ga, 9 percent (thickness 1.2 μm) and Fagen and Fritzsch [21] on the
Fig. 3. (a) Typical $I$-$V$ characteristics without Joule heating illustrating the influence of temperature and electrode separation. Composition: $(\text{AsTe})_x\text{Si}_{1-x}$. (b) Typical $I$-$V$ characteristics from Walsh et al. [16]. (c) Typical $(1/V)$-$V$ characteristics from Fagen and Fritzsche [21].

Composition $\text{As}_4\text{Te}_3\text{Ge}_2$ (thickness 0.9 $\mu$m); both cover a reasonably wide range of temperature. Fagen and Fritzsche's results are plotted in the form $(1/V)$ versus $V$ (7b). Note, in Fig. 3(b) and (c), that a second well-defined exponential $(\log I \propto V)$ region develops, at low temperatures, i.e.,

$$I = I_{OA} \exp (V/V_{OA}) + I_{OB} \exp (V/V_{OB})$$  \hspace{1cm} (8a)

$$G = G_{OA} \exp (V/V_{OA}) + G_{OB} \exp (V/V_{OB}). \hspace{1cm} (8b)$$

There is information available on the high-field behavior of simpler but related amorphous semiconductors such as vitreous Se and As$_2$Se$_3$, and although completely analogous switching behavior has not been demonstrated, the conduction mechanisms probably have a good deal in common with the more complex chalco-
genide glasses. Amorphous or vitreous Se, in particular, has been widely studied and high-field conduction has generally been interpreted in terms of various single-carrier space-charge-limited current models [22]—[26]. It should be noted, however, that except in the case of [25] there has been no systematic investigation of the thickness dependence of current—a crucial factor in space-charge-limited current mechanisms. In contrast, Muller and Muller [27] have recently concluded that the high-field $I-V$ characteristics of amorphous Se are determined by the Poole-Frenkel mechanism [28]. Kolomiets and Lebedev [29] interpreted their results on $\text{As}_2\text{Se}_3$ in terms of space-charge-limited current with the Fermi level in an exponential distribution of traps, but de Wit and Crevecoeur [17] conclude that the current is not space-charge limited. Recent results from the authors' own laboratory (Renouf [30]) on thin evaporated $\text{As}_2\text{Se}_3$ films ($\sim 1 \mu m$) over a wide range of temperature (370–100 K) show similarities with the $I-V$ characteristics of Fig. 3(b), the main difference being that although the curves become steeper as the temperature is lowered, a well-defined second $I \propto V$ region does not develop.

B. Space-Charge-Limited Current

Some of the models proposed to explain threshold switching (see Section III-C) have much in common with double injection and it might be expected therefore that in the preswitching region the current flow is space-charge limited, either single or double carrier [4]. Single-carrier space-charge-limited current flow with the Fermi level in a uniform distribution of traps predicts an $(I'/V)-V$ characteristic exactly of the form of (7b) with

$$
\frac{1}{V_0} = \frac{k_0}{\kappa_0 e N_g k T d^2}
$$

(9)

where $k_0$ is the permittivity of free space, $\kappa_0$ is the dielectric constant, $N_g$ is the trap density, and $d$ is the sample thickness. Hall [31] has suggested that both exponential regions of the data of Walsh et al. [11b] should be interpreted in this way; the increased slope at lower temperatures corresponding to the Fermi level entering a second uniform trap distribution of higher density (e.g., Hall obtains trap densities of $2 \times 10^{22}$ and $6 \times 10^{23}$ m$^{-3}$ eV$^{-1}$ from the two $V_0$ values). Unfortunately, the crucial thickness dependence was not investigated. According to (9), $(1/V_0)$ should be proportional to $(1/T)$ and $(1/d^2)$. Some of Robertson's [14] results for three As–Te–Si glasses are plotted to illustrate these relationships in Fig. 4(a)–(c), and although the data show the expected functional dependence on temperature and thickness, none of the lines in Fig. 4(b) and (c) pass through the origin. This is confirmed by a comparison with some results of Fagen and Fritzsche, also shown in Fig. 4(b). Robertson [14] obtained trap densities in the region $2 \times 10^{22}-2 \times 10^{24}$ m$^{-3}$ eV$^{-1}$ for a range of compositions, but there was no systematic variation with composition.
These trap densities are very similar to those derived by the same methods for amorphous Se [23] and Se–Ge glasses [32], but, as Hall [31] points out, are much less than those obtained by other electrical techniques for the more complex chalcogenide glasses [33]. The latter point is not, perhaps, too serious, as there is some doubt about the large trap densities calculated from some electrical measurements, but the nonzero intercepts of Fig. 4(h), (c) must cast considerable doubt on the applicability of the simple form of the space-charge-limited model.

It was mentioned in the previous paragraph that Kolomietz and Lebedev [29] reported space-charge-limited current in vitreous As2Se3, but again the thickness dependence does not support this.

According to the recent measurements by de Wit and Crevecoeur [17], $V_o$ is proportional to $d$, i.e., the field parameter $E_o$ is constant at, in this case, a value of about $1.6-1.8 \times 10^7 \text{ V/m}$. Renouf’s data are also consistent with this [30].

C. Hopping Conduction

In an amorphous semiconductor with a high density of localized states, a large fraction of the carriers is likely to be trapped at any instant. Hopping conduction, either directly between traps or via the conduction (or valence) band, is therefore a possible means of transport and, as Bagley [34] has pointed out by analogy with well-known treatments of ionic conduction, the $I$–$V$ relationship should be

$$I = I_o \sinh \left( \frac{V}{V_o} \right)$$

where now

$$V_o = \frac{2kTd}{e\alpha}$$

where $\alpha$ is the “jump” distance.

The preexponential constant $I_o$ is given by

$$I_o = A \text{area} \exp \left(-\frac{\phi}{kT} \right)$$

with $n$ the trapped carrier concentration, $\phi$ the height of the potential barrier, $\nu$ a phonon frequency, and $A$ the area. Equation (10) leads to an exponential $I$–$V$ characteristic (7a) for $V > V_o$. From (10) the low-field ohmic conductance ($V \to 0$) is given by

$$G_0 = 2(I_o/V_o).$$

Superficially, at any rate, the data of Fig. 3(a) and for other samples, conform approximately to (11) as shown in Fig. 5, in which the straight line corresponds to a constant field $E_o = 3.3 \times 10^7 \text{ V/m}$. At room temperature this gives a value for the jump distance $\alpha$ of about 15 nm. This is a rather larger value than expected, but seems to be typical (other values are summarized later).

The temperature and thickness dependence of the “constants” $V_o$ and $I_o$ appear to be anomalous. In Fig. 3(a), which is representative of Robertson’s [14] data, an increase in ambient temperature results in an increase in $I_o$ and a decrease in $V_o$; increased thickness has the opposite effect. According to (11) and (12), $V_o$ should increase with $T$ and $I_o$ should be independent of film thickness. Experimentally it is found that both $V_o$ and $I_o$ vary exponentially with $(1/T)$, but in opposite senses. Thus activation energies can be derived and typical values are listed in Table I, along with the low-field activation energy for the “ohmic” conductance $G_0$.

The $I_o$ activation energies correspond, of course, to $\phi$ in (12) and the difference in the values for $V_o$ and $I_o$ does roughly agree with that for $G_0$ as expected from (13). The apparent “activation” of $V_o$ is puzzling; the only suitable term in (11) that could be temperature activated is the jump distance $\alpha$, and the implication is, therefore, that

$$\alpha = a_o \exp \left(-\frac{\Psi}{kT} \right)$$

with $\Psi = 0.06 - 0.14 \text{ eV}$ (Table I) and $a_o$ is in the range 0.2–4 $\mu$m. From the calculated values of $\phi$ and $\alpha$ and a typical value for $\nu$ ($10^14 \text{ s}^{-1}$), an estimate of the carrier density $n$ can be made, and hence also of the mobility $\mu$, using the ohmic conductivity. Some representative results are given in Table II.

The spread in the possible values of $\phi$ introduces an uncertainty of about one order of magnitude in $\mu$ and $\alpha$. Thus although superficially the simple hopping-type model for conduction fits the functional form of the $I$–$V$ characteristics, a detailed analysis of the parameters involved leads to inconsistencies. Abnormally large jump distances seem to be a typical consequence, the apparently “temperature-activated” behavior of $V_o$ (and $\alpha$) is difficult to reconcile with the simple model, and $V_o$ and $I_o$ do not vary with temperature in the required way.

![Fig. 5. Variation of $V_o$ in (12) with electrode separation. Composition same as in Fig. 3(a).](image-url)
TABLE II

<table>
<thead>
<tr>
<th>Unit</th>
<th>Thickness (µm)</th>
<th>Area (x 10^-7 m²)</th>
<th>G (x 10^-7 S m^-1)</th>
<th>V (Volts)</th>
<th>I (mA)</th>
<th>n (x 10^22 m^-3)</th>
<th>V (x 10^3 V s^-1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.1</td>
<td>3.6</td>
<td>1.8</td>
<td>3.7</td>
<td>0.1</td>
<td>15</td>
<td>1.4</td>
</tr>
<tr>
<td>2</td>
<td>0.8</td>
<td>3.4</td>
<td>3.3</td>
<td>2.3</td>
<td>0.18</td>
<td>17</td>
<td>2.4</td>
</tr>
<tr>
<td>3</td>
<td>0.45</td>
<td>4.8</td>
<td>3.6</td>
<td>1.9</td>
<td>0.5</td>
<td>12</td>
<td>6.7</td>
</tr>
<tr>
<td>4</td>
<td>0.45</td>
<td>2.2</td>
<td>12.0</td>
<td>1.5</td>
<td>0.52</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>0.75</td>
<td>3.0</td>
<td>9.1</td>
<td>1.7</td>
<td>0.3</td>
<td>22</td>
<td>3.5</td>
</tr>
<tr>
<td>5</td>
<td>0.2</td>
<td>2.6</td>
<td>5.9</td>
<td>0.9</td>
<td>0.45</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>

Note: All samples have nominal film composition (As,Te)x, Si, Units 1-4 are the same as those listed in Table I.

D. Field-Assisted Emission from Traps: The Poole-Frenkel Effect

The Poole-Frenkel effect has been widely used to interpret the current-voltage characteristics of many amorphous oxide and polymeric dielectrics [28], [35]. It describes the effect of an applied field on the attractive Coulombic potential between a carrier and an oppositely charged trap from which the carrier has been released. In this case, therefore, the traps are neutral when occupied and correspond to donors (for electrons) or acceptors (for holes). For one-dimensional current flow, the Poole-Frenkel effect leads to a field dependent conductivity of the form [28], [35]

$$
\sigma = \sigma_0 \exp \left( \frac{\beta E^{1/2}}{kT} \right)
$$

(14)

where

$$
\beta^2 = \left( \frac{e^2}{\pi \epsilon_0 \epsilon} \right)
$$

(15)

The precise value of $\beta$ depends upon the dimensionality of the problem [36] and on the position of the Fermi level relative to the various charged and uncharged traps [37].

The overlapping of the Coulombic potential between adjacent charged sites must also be considered and this radically affects the form of the $I-V$ dependence [28]. The maximum in the combined potential occurs at the midpoint between traps and it stays essentially fixed, independently of field, until some critical field $E_c$ is reached. Only above $E_c$ is the position of the Coulombic barrier affected by the field, and this is a necessary condition for (14). Below $E_c$ the position of the potential barrier is not changed by the field. This is identical with the situation in “hopping conduction” (Section III-C) and hence, below $E_c$, equations of exactly the same form as (10) and (11) are obtained [28]. As before, the “sinh” form of (10) goes over to exponential (7a) for large voltages. The value of the critical field can be estimated and it is shown in Fig. 6 as a function of the density of charged sites for $\beta = 3 \times 10^{-5}$ eV/V m [12]. Below $E_c$, the potential maximum is “fixed” midway between traps; only above $E_c$ is the position of the maximum affected by the field. See text for explanation (Section III-D).

![Fig. 6. Critical field $E_c$ as a function of trap density in the Poole-Frenkel effect calculated for $\beta = 3 \times 10^{-5}$ eV/V m. Below $E_c$, the potential maximum at $V=0$ is fixed midway between traps; only above $E_c$ is the position of the maximum affected by the field. See text for explanation (Section III-D).](image)

Typical experimental data for a chalcogenide glass are shown in Fig. 7(a) in the form of log $G$ versus $V^{1/2}$, and the general behavior agrees with that expected for the Poole-Frenkel mechanism. At constant temperature, lines representing films of different thickness extrapolate to a constant value at $V=0$ (allowing for small errors due to slightly different electrode areas), and the slope decreases as thickness and temperature increase. A slightly more detailed comparison is less satisfactory, however, as demonstrated in Fig. 7(b) and (c), which show the variation of slope with temperature and thickness, respectively. Neither line passes through the origin, and this cannot be accounted for by experimental error.

The data of Fig. 7(a) also fit a log $G$ versus $V$ plot equally well, however, as noted above this is consistent with field-assisted emission in circumstances where the overlapping of the Coulombic
barriers is significant. Equation (7a) applies with the values of Section III-B and \( V_0 \) given by (11). The jump distances derived in Section III-B will therefore be applicable and taking 15 nm as a typical value gives the concentration of charged traps as \( 3 \times 10^{23} \text{ m}^{-3} \). This is probably not an unreasonable figure for chalcogenide glasses, although rather higher values have sometimes been quoted [21]. According to Fig. 6, the \( E^{1/2} \) Poole–Frenkel behavior would not be expected to develop until fields approaching \( 10^7 \text{ V/m} \) were reached. This is the magnitude of the threshold field for typical chalcogenide glass switches of thickness less than 10 \( \mu \text{m} \).

The overlapping Coulombic potential modification to the conventional Poole–Frenkel effect does, therefore, offer an alternative interpretation that overcomes one of the objections to the simple hopping model (Section III-B), but the problem of the apparently temperature-activated jump distance remains. It should also be recognized that at the charge densities involved, the Coulomb potentials will probably be considerably modified by screening (overall charge neutrality is required) and this will affect the shape of the barriers, and hence the
predicted $I-V$ characteristics. Yet another difficulty, and one which raises more general problems, is that the $I-V$, or $(I/V)-V$, characteristics do not show the expected change from exponential to sinh behavior when $(V/V_0) < 1$. This is discussed more fully in the following section.

**E. Comments and Discussion**

Electronic processes certainly make an important contribution to the nonohmic conductivity of chalcogenide glasses in the preswitching region, but none of the common models give an adequate description, at least in their elementary forms.

The principal experimental features seem to be two exponential $I-V$ or $G-V$ regions [i.e., equations (8a) or (8b)], the second region at higher fields having a higher slope (lower $V_0$ value). In the more conducting complex compositions used for switching [Te, 49 percent; As, 33 percent; Ge, 6 percent; Si, 3 percent; Ga, 9 percent (all in weight percent)] [11b]; (As$_{15}$Te$_{34}$Ge$_2$ [21]; Ge$_{15}$Te$_{4}$[38]), the high-field region only becomes apparent at low temperatures ($\leq 200$ K), but in As$_8$Se$_6$ [17], [30], for example, it is observed at room temperature and above. In all cases the break from the low to high slope region occurs at a field value of about $1-5 \times 10^6$ $V \cdot m^{-1}$ with only small variations with temperature.

The major problem in explaining the lower field region is that it extends to smaller $(V/V_0)$ or $(E/E_0)$ ratios than expected for any of the simple variants of hopping conduction. This is illustrated in Fig. 8 where $(\sigma/\sigma_0)$ is plotted versus $(E/E_0)$ for a variety of data. The exponential field dependence is followed down to $(E/E_0)$ ratios at least 0.05 and probably 0.01. Beyond this it becomes an extremely difficult matter to distinguish experimentally between ohmic and nonohmic conduction and it is a moot point whether or not an ohmic characteristic is ever reached. De Wit and Crevecoeur [17] imply that it is not. This behavior, although apparent in some of the published data, does not appear to have been explicitly recognized until the recent work of de Wit and Crevecoeur [17] and Marshall and Owen [38]. The former authors briefly speculate that their results
on As$_2$Se$_3$ may be explained by potential fluctuations with a wavelength greater than some characteristic "jump distance" related to $E_c$. Marshall and Owen also find that the same exponential field dependence is observed in the trap-limited hole drift mobility in vitreous Se down to surprisingly low fields [38] and typical results are included in Fig. 8. It would be premature to ascribe this seemingly general behavior to a common mechanism. The Ge$_3$Te$_6$ and (As$_2$Te$_3$)$_{6.5}$Si$_{1.5}$ glasses are probably phase separated, for example, and this may affect their field-dependent properties. The circumstantial evidence is strong, however, and Marshall and Owen discuss the consequences in qualitative terms. The same field dependence of the conductivity and trap-limited mobility means that the field is not affecting the free carrier concentration $n$. The other relevant parameters that may be affected are the trapping time $\tau_t$, involving the capture cross section, and the free mobility $\mu_o$. Marshall and Owen eliminate the first two and conclude, therefore, that the field dependence of both conductivity and trap-limited mobility enters through $\mu_o$, although as yet there is no direct experimental proof of this. They suggest, tentatively, that the field increases $\mu_o$ by increasing the percolation probability of carriers in the "diffusive" extended states at the bandedge and also, possibly, by increasing the overlap of the wave functions of such states.

The apparent similarity in behavior of conductivity and mobility extends to the breakaway to an enhanced field dependence at fields in the region of $1-5 \times 10^7$ V/m, and this is also indicated in Fig. 8 [the breakaway occurs in this diagram, at different (E/E$_c$) values because $E_c$ varies]. This higher field region can only be observed over a limited range as it is quickly followed by switching or breakdown. Marshall and Owen argue that this probably represents the onset of hot carrier phenomena which, they suggest, will start at fields of the order of

$$E \approx \left(\frac{kT\nu}{2\mu_o}\right)^{1/2}$$

(16)

where $\nu$ is the frequency of the energy loss process. A similar equation has been proposed by Mott [13]. The free mobility $\mu_o$ is typically $3 \times 10^{-4}$ m$^2$ V$^{-1}$ s$^{-1}$ in amorphous chalcogenide semiconductors [38] and substituting this value with $\nu$ gives $E \approx 5 \times 10^7$ V/m. Clearly there is a need for a much more thorough investigation and understanding of high-field phenomena, conductivity, and mobility in chalcogenide glasses.

Any of the mechanisms discussed in the previous paragraphs could be very significant in electronic switching processes and hot carrier phenomena especially, leading probably to impact ionization and avalanching, could promote electrical instabilities. It must be remembered, however, that in the compositions typical used for switching devices the latter effects only come into play at temperatures below about 250 K; at higher temperatures switching occurs before the appropriate fields are reached [11b], [21].

At the present time, models and theories of high-field phenomena are not developed well enough to apply confidently to the even more complex switching process. All that can be done is to speculate on the likely connection between the behavior in the preswitching region and the actual threshold conditions.

IV. Threshold Conditions

A. Experimental Features

As mentioned previously, the experimental data on switching are reviewed by Warren [1] and relevant references can also be found there. For the purposes of the present discussion, however, the main experimental facts concerning the threshold conditions are summarized below. It must be remembered at all times that the concern here is with devices less than about 10 $\mu$m thick. In thicker films, and particularly for thicknesses greater than 100 $\mu$m, the switching process is almost certainly thermally induced and some of the features listed below do not apply. The controversy is really about the apparently different behavior of "thin" (< 10 $\mu$m) devices.

1) The threshold voltage $V_{th}$ is a slowly decreasing function of ambient temperature. Very approximately $V_{th} \propto (1/T)$, $V_{th} \propto (-T)$, or $V_{th} = V_{the} (1-bT)$, where $b$ is a constant.

2) The threshold voltage is proportional to thickness $d$, i.e., the switching field is essentially independent of thickness for $d \leq 10$ $\mu$m. The magnitude of threshold field is typically $8 \times 10^7-2 \times 10^8$ V/m.

3) The delay time $\tau_d$ decreases exponentially with pulse amplitude $V$, i.e., $\tau_d \propto \exp (-V)$, but tends to infinity as $V \rightarrow V_{th}$. For overvoltages $(V-V_{th})$ of a few volts, $\tau_d$ is typically 0.1-1 $\mu$s. If the applied voltage is removed during the delay time the switch returns to the off state.

4) The delay time is followed by the instability leading to the on state. The switching time $\tau_s$ is of the order of 100 ps.

5) In the on state the evidence is that the current flows in a narrow channel or filament. For a threshold switch a minimum holding voltage $V_h$ and current $I_h$ must be maintained, otherwise the device returns to the off state. The on state will be discussed in Section V.

A number of electronic processes have been suggested as applicable to the threshold behavior, and the on state, of chalcogenide glasses, but none has been developed to the stage where quantitative calculations and predictions can be made [13], [39]–[41]. Generally speaking, however, the suggested models make few, if any, concessions to the amorphous nature of the material, and they usually draw heavily on concepts familiar in conventional semiconductor physics. The chalcogenide glass
is often regarded as a typical semiconductor modified only by the presence of a larger than normal density of traps and/or recombination centers—a view which, in the present circumstances and context, may be quite justifiable. The most fruitful approach at the moment, therefore, is to review the suggested mechanisms and, where it is relevant, to appeal to related data and theories developed for crystalline semiconductors.

B. Carrier Injection Mechanisms

The electronic process so far having most appeal—probably because it is basically simple to visualize—is double injection of both carriers. Excess electrons are injected from the cathode, holes from the anode, forming a negative and positive space charge at the cathode and anode, respectively. The space-charge clouds build up from the electrodes and when they overlap neutralize each other, causing the field in the interior to collapse and the same, or larger, current can flow with a lower applied voltage, i.e., there is a negative-resistance characteristic. A qualitative explanation along the above lines was first explicitly suggested by Henisch [39] and Henisch et al. [40]. A variation has recently been proposed, in slightly more quantitative terms, by Lucas [41]. Mott [42] has also suggested a related process to describe the situation in the ox state; this will be mentioned in Section V.

A schematic representation of Lucas' [41] model, freely interpreted, is shown in Fig. 9. The vertical axis in this diagram is not drawn to scale, nor is the scale intended to be consistent in the various parts. The assumption is that the "tails" of localized states extending from the conduction and valence bands overlap somewhere near the center of the energy gap (or mobility gap). This provides a completely compensated set of positively and negatively charged states above and below the Fermi level, respectively, with the Fermi level pinned at or near the center of the gap. These features are inherent in the Cohen–Fritzsche–Ovshinsky (CFO) [43] model of the electronic-band structure of a chalcogenide glass.

For simplicity in Fig. 9, the charged states on either side of the Fermi level are denoted by a single row of positive and negative signs. The injected electrons and holes will recombine with and neutralize the positively and negatively charged states setting up a negative and positive space charge, Fig. 9(a), (b), adjacent to the anode and cathode, respectively. Charge will also be trapped in the neutral localized states, of course (not shown in Fig. 9), but Lucas ignores this. Certainly one would expect that recombination via the charged states would be the dominant process in establishing the space charge, at least initially. The regions of space charge will limit the current flow in the vicinity of the electrodes and the field will be redistributed—decreasing near the electrodes and increasing in the center—Fig. 9(b). As the applied voltage is increased more charge is injected and the space-charge regions grow until eventually they meet and overlap—Fig. 9(c). The situation then rapidly becomes unstable. With recombination centers filled, the carrier lifetimes, and hence the electron and hole conductivities, increase. In the region of overlapping space charge the material is neutral; the conductivity in the center increases and the field decreases while the field near the electrodes increases—Fig. 9(d). Electrons and holes are accelerated rapidly across the neutral region and because of the increased field, injection of carriers at the electrodes increases. Both effects increase the rate at which space-charge overlap occurs, i.e., there is positive feedback and hence an unstable situation. The stable state, corresponding to the ox state, is shown in Fig. 9(e). Here, the space charge has spread right across the diode and the bands are practically flat. Shottky-type barriers are established at the electrodes, but because of the high density of traps (the CFO model [43]) they are thin (about 1 nm) and electrons can easily tunnel through from the Fermi level of the metal into the conduction band of the chalcogenide glass and, similarly, holes from the anode tunnel into the valence band. Because the space charge has been neutralized and the traps filled, the conductivity in the device is now high (the mobility is no longer limited by trapping). The minimum voltage required to keep the ox state stable (i.e., the holding voltage $V_h$) is approximately equal to, or a little greater than, the mobility gap of the chalcogenide glass. Further consideration will be given to this point in Section V.

The model of Henisch [39] and Henisch et al. [40] is similar except that the charge is trapped in the neutral localized states of the "tail" adjacent to the electrodes at which they are injected. Thus a negative space charge again develops at the cathode and a positive space charge at the anode.

Apart from the magnitude of the holding voltage $V_h$, no quantitative predictions have been made on the basis of the Henisch–Fagen–Ovshinsky model, which can be
compared with the basic experimental features outlined in Section IV-A, Lucas [41] has made a semiquantitative analysis of the model outlined in Fig. 9, which she describes as a recombination instability, and finds that the critical current \( j_r \) at which the high-resistance state collapses [i.e., corresponding to Fig. 9(c)] is given by

\[
j_r = \frac{2L_e eN}{\tau_{rel}}
\]

(17)

where \( L_e \) is the diffusion length (a function of carrier lifetime and hence of injection level), \( N \) is the density of recombination centers, and \( \tau_{rel} \) is the relaxation time of excess charge.\(^1\) A complete quantitative analysis of these charge injection models of switching, relevant in all its particulars to chalcogenide glasses, would of course be a formidable task, but there are similarities with the established ideas of double-injection space-charge-current problems in conventional semiconductors [44] and certain comparisons may be profitable. Even here, however, the only case leading to negative resistance (switching) that has been analyzed in any detail is that of a semiconductor with a single set of deep traps (recombination centers). One feature that should nevertheless be generally applicable is that the threshold voltage is a function of the square of thickness (electrode separation) \( d \). This is because in all space-charge-dominated transport processes (single or double carrier), the transition from one regime to another is determined by the transit time of the appropriate carrier becoming equal, at say, \( V_{th} \), to its lifetime \( \tau \), i.e., [44],

\[
\frac{d^2}{V_{th}^2} = \tau \rightarrow V_{th} = \frac{d^2}{\mu \tau}
\]

(18)

where \( \mu \) is the effective carrier mobility. Thus in this simple case, the threshold field increases with thickness, which is not the relation observed in switching in chalcogenide glasses (Section IV-A). One could conjecture about the temperature dependence of the threshold voltage given by (18), but this could not be realistically applied in the present context since the equation is significantly modified by, for example, incomplete occupation of the recombination centers at thermal equilibrium (leading to a space-charge limitation on the prebreakdown current [45a]) and by the effects of shallow trapping [45b]. The thickness dependence remains unaffected, however.

The most directly comparable situation to have received attention in conventional semiconductors seems to be the \( p^+ - i - n^+ \) structure in silicon containing deep-donor (hole recombination centers) and deep-acceptor (electron recombination) levels with densities \( N_A \) and \( N_D \), respectively [46]. This is exactly equivalent to Lucas' [41] model for the chalcogenide glasses, described earlier. Wright and Ibrahim [46] show by a very simple and approximate analysis, which nevertheless gives good agreement with experiment (on their Si \( p^+ - i - n^+ \) structure), that

\[
V_{th} = \frac{d \varepsilon N_A N_D}{2k(N_A + N_D)}
\]

(19)

The \( d^2 \) dependence of \( V_{th} \) still obtains, as one would expect for any model of this kind. The donor and acceptor concentrations would be approximately the same; in fact, in the present context [41] they will be equal: \( N_A = N_D = N \). Using a value of \( N = 10^{26} \text{m}^{-3} \) [43], inserting \( d = 1 \) \( \mu \text{m} \) as a typical device thickness, and taking the relative dielectric constant to be 10 gives \( V_{th} = 4 \times 10^4 \text{V} \), i.e., \( \sim 4 \times 10^4 \text{V/m} \) for the threshold field! This is two orders of magnitude, or more, greater than threshold voltages observed in chalcogenide glasses (Section IV-A), but clearly could be made to agree by taking a smaller value for \( N \left( \sim 10^2 \text{m}^{-3} \right) \). A smaller \( N \) also leads to difficulties, however [41], and one is still left with the prediction of a threshold field that increases with thickness in contrast to the experimentally observed constant field.

In its simplest interpretation the delay time \( \tau \) is the time required to fill (or half fill) all the traps and/or recombination centers, i.e., to reach the condition represented by Fig. 9(c) where the space-charge regions are just beginning to overlap. For overvoltages (\( V - V_{th} \)) of a few volts, \( \tau \) is about 1 \( \mu \text{s} \). Again using a value of \( 10^{26} \text{m}^{-3} \) for the density of recombination centers, the current required to inject the equivalent amount of charge during \( \tau \) (1 \( \mu \text{s} \)) is approximately \( 10^3 \text{A/m}^2 \). This should correspond to Lucas’ critical current \( j_r \) and again a discrepancy of several orders of magnitude is found. Lucas [41] remarks on the same problem but in slightly different terms. One could, of course, conclude that \( N \) was much less, but as Lucas notes this would adversely affect other features of the model, viz., short-screening lengths at the electrodes. Incomplete filling of the traps and/or recombination centers could be assumed, but carried over into the on state this would imply a trap-limited mobility and probably also a space-charge-limited on-state current, contrary to observations (Section V). A third alternative, mentioned by Lucas, is that the current-carrying area is contracting (i.e., filament formation) well before space-charge overlap occurs. There seems to be little evidence at the moment for or against filament formation in the prebreakdown region, but it is relevant to note that according to Vogel and Walsh [47] the capacitance of a chalcogenide glass switch retains its geometric value up to \( (V/V_{th}) \sim 0.8 \). In As\(_2\)Se\(_3\), deviations (negative) from the geometric capacitance also occur, but even closer to the breakdown voltage, e.g., \( (V/V_{th}) \sim 0.95 \) [30]. Furthermore, above some critical frequency the small signal capacitance remains constant as a function of bias virtually to \( V_{th} \) [47]. [48].

\(^1\) Taking \( N = 10^{26} \text{m}^{-3} \) as a typical value [43] and \( \tau_{rel} = 10^4 \text{s} \), Lucas obtains \( j_r \sim 10^5 \text{A/m}^2 \), which compares reasonably with experimental values. Her choice of \( \tau_{rel} \) does not seem physically reasonable, however; it is likely to be one or two orders of ten smaller. In addition, a low value of \( L_e \) appears to have been used. As a result \( j_r \) estimated from (17) is likely to be several orders higher than the value quoted by Lucas, still taking \( N = 10^{26} \text{m}^{-3} \).
The most cogent evidence against charge injection models for switching is the absence of polarity effects, and this has been widely debated at conferences and elsewhere. Shanks [49] and Balberg [30] have reported that the delay time $t_d$ is unaffected by reversals of the polarity of the biasing pulse. Henisch and Pryor [51] have recently carefully repeated this type of experiment and they do find some polarity effects, particularly at low temperatures ($\sim 200$ K). The effects are not of sufficient magnitude, however, to suggest that the whole of the space-charge pattern has to be nullified and reestablished in the opposite direction on reversal of polarity; i.e., if the polarity is reversed at, say, $0.5 t_d$, the resultant total delay time is greater than the original $t_d$ but is rather less than $1.5 t_d$. Nevertheless, these experiments do suggest that at low temperatures some charge-injection processes may have a more dominant role. This could be connected with the increased field dependence of the preswitching current observed in the “high-field” region and described in Section III-A, E; in typical switching compositions this also becomes apparent only at temperatures $\leq 200$ K [11], [21].

C. Hot-Electron Effects

A variety of effects related to hot-electron phenomena and dielectric breakdown could lead to the initiation of an instability, negative resistance, and switching of the form observed in chalcogenide glasses. Once more, however, there are no well-developed theories readily applicable in the present context and it is possible only to speculate rather qualitatively. This section will, therefore, be brief and limited to a discussion of two manifestations of hot-electron effects that are well understood in other contexts and provide some experimental and/or theoretical data for comparison. It must be emphasized, however, that such evidence as there is (see Section III and [38]) suggests that the onset of hot-electron effects in chalcogenide glasses occurs at fields $>10^6$ V/m, probably at least $3-5 \times 10^6$ V/m. These figures agree with Hindley's theoretical estimates based on the random-phase approximation [52a]. In typical chalcogenide glass switches, the threshold field at room temperature is usually $\leq 10^5$ V/m. Thus hot-electron effects are unlikely to be of first-order importance except in pulse experiments at substantial overvoltages, or at low temperatures.

1) Collective Breakdown: When the electron density and the electron-electron collision rate is sufficiently high, the rate of energy exchange due to electron-electron collisions is greater than that due to electron-phonon interactions. The electron energy-distribution function is Maxwellian with a mean temperature $T$ above the ambient temperature. The situation has a strong analogy with thermal breakdown, but in this case it is the steady-state electron temperature that rises until at a critical value $T^c$.

Equilibrium is no longer possible and breakdown occurs. There is always some increase in the electron-phonon energy transfer near the energy maximum, however, and this implies some internal temperature increase. Thus thermal effects are usually also linked with collective breakdown.

Stratton [3] has calculated that the electron density necessary for collective breakdown is $10^9$ to $10^{10}$ times greater than the low-field electron density. In chalcogenide glasses the preswitching $I-V$ characteristics imply that the maximum electron density is at the most $10^6$ to $10^7$ times the low-field value, so collective breakdown is unlikely to be important in the initiation of switching.

2) Impact Ionization and Avalanche Breakdown: If the electron density is low, a high-energy electron may collide with an atom, instead of another electron, and ionize it producing a hole and two low-energy electrons. The two electrons will, in turn, be accelerated to high energies and ionize more atoms, thus the process can cascade to cause an electron avalanche. This process is, of course, well known in dielectric breakdown and has been discussed in detail by many authors (see, for instance, [2] and [53]). Impact ionization, or any other hot-electron effect, is not in itself sufficient, however, to produce negative resistance. As mentioned before, a positive feedback mechanism is required. Crandall [53a] has investigated the conditions for CCNR and shows that a possible feedback mechanism is the increased screening of the scattering potentials due to the increased carrier density. The electron scattering decreases, leading to an increase in the average electron energy. The hotter electrons cause more impact ionization, generating more carriers and making the distribution still hotter. Hence, a higher rate of impact ionization can be sustained at a value of the electric field that is lower than that required for impact ionization at low-carrier density.

One of the most important applications of avalanche breakdown is to p-n junctions. It provides a limit to diode reverse bias voltages or to transistor collector voltages and has, therefore, been studied extensively. There are several important characteristics of avalanche breakdown in p-n junctions that are probably general features of the phenomenon.

a) The breakdown is localized and the small regions of avalanche are described as microplasmas. Light is emitted due to electron-hole recombination so the spatial location and density of the microplasmas can be studied, in p-n junctions at least [53b]. The microplasmas are thought to form at high-field spots associated, for example, with crystalline defects, inclusions, non-uniformities in impurity concentration, or other inhomogeneities [54].

b) Rose [55] estimated that the temperature rise in a microplasma was in the range of 15-40 K. A more recent treatment by Martirosov [56] puts it at ~100 K depending on current density.

c) The formation of microplasmas is a statistical process and this has two important consequences.
i) If a high-voltage is applied, there is a time delay before any avalanche occurs, but this time lag appears to be entirely statistical and not a parameter unique to a given set of experimental conditions. Thus in avalanche breakdown, the delay time can only be defined statistically [57].

This contrasts with the situation as normally reported in chalcogenide switches where it is usually implied that, except possibly at small overvoltages, the delay time is a uniquely defined parameter. Very recently, however, Lee and Henisch have reported on significant statistical variations of threshold switching in chalcogenide glasses [58]. According to Nield and Leck's [57] experiments, the delay time for avalanche breakdown in silicon decreases approximately inversely with voltage. This is rather slower than the variation observed in chalcophenoxide glass switches (Section IV-A).

ii) The on-off fluctuations of microplasmas at fields below the overall breakdown field result in a very characteristic form of high-level current noise [59]. Current noise has been measured in As$_2$Se$_3$ glass at fields of up to $0.5 \times 10^5$ V/m with no evidence of any avalanche microplasma features [60], although preliminary measurements on a more conducting glass composed of As-Te-Ge-Si have indicated that the current noise does increase considerably at voltages close to $V_{th}$. This could be due to localized avalanche breakdown in regions of the sample where the field is much greater than the average value, but there are also other possible causes.

d) An increase in ambient temperature reduces the mean free path between electron-phonon collisions; electrons lose more energy to the lattice and a higher field is required to produce an avalanche. This trend has been confirmed experimentally for avalanche breakdown in abrupt silicon junctions where it is found that $V_{th}$ increases approximately linearly with temperature [61]. This contrasts with the decreasing threshold voltage in chalcogenide glass switches.

The conclusion is, therefore, that experimental and theoretical evidence exists for hot-electron effects, most likely electron avalanching, at voltages $\geq V_{th}$ in chalcogenide glasses, and such effects may be present especially in pulse experiments with amplitudes somewhat greater than $V_{th}$. Comparisons with well-known examples suggest, however, that they are not the dominant mechanism giving rise to switching. Electron avalanching at the electrodes could also be significant in the initiation process if appreciable internal self-heating occurs, causing enhancement of the fields at the electrodes [13], [14]. It is, of course, debatable to what extent such comparisons, e.g., with avalanche phenomena in p-n junctions, are valid, but in the absence of definitive theory specific to transport in chalcogenide glasses a comparative approach does at least offer some guide to speculation and conjecture. It is also in the spirit of the current models of the electronic band structure of chalcogenide glasses which, so far as switching is concerned, are usually regarded as "typical" semiconductors with a more or less large density of traps and/or recombination centers [39]–[41].

V. THE ON STATE

A. Experimental Features of the ON State

There is a real lack of hard experimental data on the ON-state parameters. Much of what has been published is essentially concerned with the same type of device using basically the same composition, viz., the Ovonic threshold switch (OTS) of Energy Conversion Devices, Inc., in a "DO-7 package" (see [6], [40], [49], [51], [62a], [62b]). It is difficult to be sure, therefore, whether the data are typical or peculiar to that particular form of device. Neale's paper [62b] does give some information on switches in the form of a simple sandwich of metal–chalcogenide–metal with crossover electrodes and also on devices fabricated in the "pore-type" structure. Recent work often utilizes one or other of these structures [14], [63]–[68], but again similar compositions are usually investigated. In general, however, the main experimental facts are as follows.

1) In the ON state the dynamic resistance is essentially zero (the $I-V$ characteristic is practically vertical), but sometimes a small region of negative resistance is observed as the device switches off. The conductance at the critical point ($V_m, I_m$) is usually about $10^{-4} \Omega^{-1}$.

2) The holding voltage $V_h$ is substantially independent of temperature and thickness. The latter is taken as evidence that most of the holding voltage is dropped near the electrodes and that there is only a small field across the bulk of the device.

3) The magnitude of $V_h$ is often in the region of 1–2 V, but there are exceptions. Table III lists results from a variety of sources showing that holding voltages as high as $20 \text{ V}$ have been reported [64]. This seems to depend on electrode material, but there have been few, if any, systematic studies on this point. It should be noted that all of the results in Table III, except the last entry, refer to thin-film devices ($<10 \mu\text{m}$).

4) The holding current $I_h$ is relatively insensitive to temperature, usually decreasing slightly as temperature increases. It is also independent of the electrode area. These two points are taken to show that in the ON state, the current is carried in a filament of a semimetallic or degenerate-semiconductor nature.

Pearson and Miller [70] and Weirauch [71] have reported evidence for conducting filaments on the ON state from consequential thermal effects observable on the electrodes (i.e., after passing high currents). Uttecht et al. [72] and Sie [73] have observed filament formation.
TABLE III

<table>
<thead>
<tr>
<th>( V_h ) (Volts)</th>
<th>( I_h ) (mA)</th>
<th>( V_I ) (Volts)</th>
<th>Electrodes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.3</td>
<td>0.6</td>
<td>0.8</td>
<td>C</td>
<td>(49)</td>
</tr>
<tr>
<td>3</td>
<td>0.3</td>
<td>0.9</td>
<td>Mo</td>
<td>(62)</td>
</tr>
<tr>
<td>12</td>
<td>0.1</td>
<td>1.2</td>
<td>Mo</td>
<td>(62)</td>
</tr>
<tr>
<td>2.5</td>
<td>0.4</td>
<td>1</td>
<td>Au</td>
<td>(69)</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>Au</td>
<td>(69)</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>Au</td>
<td>(69)</td>
</tr>
<tr>
<td>7</td>
<td>0.3</td>
<td>-2</td>
<td>Sn, Pt</td>
<td>(66)</td>
</tr>
<tr>
<td>2.8</td>
<td>0.4</td>
<td>1.1</td>
<td>Cu</td>
<td>(67)</td>
</tr>
<tr>
<td>3.0</td>
<td>0.4</td>
<td>1.2</td>
<td>Brass</td>
<td>(67)</td>
</tr>
<tr>
<td>1.8</td>
<td>0.3</td>
<td>0.6</td>
<td>Y</td>
<td>(67)</td>
</tr>
<tr>
<td>1.3</td>
<td>1.9</td>
<td>2.5</td>
<td>C</td>
<td>(67)</td>
</tr>
<tr>
<td>20</td>
<td>0.06</td>
<td>1.2</td>
<td>Si</td>
<td>(67)</td>
</tr>
<tr>
<td>1-3</td>
<td>1</td>
<td>1-2</td>
<td>Au, Mo</td>
<td>(14)</td>
</tr>
<tr>
<td>10 a</td>
<td>0.1 a</td>
<td>1 a</td>
<td>Ag a</td>
<td>(14) a</td>
</tr>
</tbody>
</table>

* These figures are for a typical "thick" device, i.e., thickness \( \sim 100 \mu m \).

Note: The value of \( V_h \) has been taken at \( I_h \).

The holding current \( I_h \) is typically 0.1–1 mA (see Table III). The ON-state current can increase, at virtually constant voltage, to several tens of milliamperes; under pulsed conditions, ON-state currents of up to 1 A have been reported [49].

5) The holding current \( I_h \) is typically 0.1–1 mA (see Table III). The ON-state current can increase, at virtually constant voltage, to several tens of milliamperes; under pulsed conditions, ON-state currents of up to 1 A have been reported [49].

6) The most useful technique so far established for examining ON-state behavior is that described by Prior and Henisch [62a]. They applied a voltage pulse sufficiently large to switch a device and then superimposed a smaller pulse in the ON-state region to measure the transient ON characteristic (TONC) Fig. 10(a). Provided that \( t_h \) was short (\( \sim 0.1 \mu s \)), an extended ON-state characteristic could be measured Fig. 10(b). This very fast transient response was explained in terms of the capture of free carriers.

When \( t_h \) is increased beyond 1 \( \mu s \), gradual recovery of the off-state features (\( V_{on}, t_h, R_{on} \)) occurs. This is the more familiar recovery from a switching event and may be expected to occur when the recombination rate of carriers exceeds their injection rate. On Prior and Henisch's model this recovery is characterized by the recombination time of nonequilibrium carriers in traps.

B. Discussion

In the injection mechanisms considered in Section IV-B, the ON state corresponds to Fig. 9(c). This diagram is appropriate either to Lucas' [41] model or to that of Henisch, Fagen, and Ovshinsky [39], [40]. Mott [13], [42], [75] has also proposed the same band model for the ON-state without necessarily requiring that it is established by the same sort of injection process. Electrons and holes tunnel through the thin (e.g., about 5 nm) potential barriers at the electrodes directly into the conduction and valence bands and it is assumed that all traps and/or recombination centers in the mobility gap are full. As Mott [13] points out, the presence of
deep traps is necessary in order to make the barriers thin enough to be compatible with tunnelling (also see Rose [76]). Except near the barriers [13], however, all traps are full and hence the on-state current is not trap or space-charge limited [13] [see Section V-A4] and the quasi-Fermi levels for electrons and holes are above and below, respectively, the edges of the mobility gap. Hence the on-state conductivity is high and there is only a small potential dropped across the bulk of the device. The holding voltage \( V_h \) is, therefore, the potential required to sustain the quasi-Fermi levels at the positions shown in Fig. 9(e), i.e., \( V_h \) is approximately equal to or a little greater than the mobility gap. This is the only quantitative prediction made by the present electronic models of switching. The mobility gap in the type of chalcogenide glass used for threshold switches is typically 1–1.5 eV and \( V_h \) is often close to this but, as shown in Table III, there are exceptions.

An estimate of the filament diameter can be obtained on the assumption that Fig. 9(e) does represent the on-state. Following Mott [13] we assume a carrier density of \( 10^7 \text{ m}^{-2} \) and a mobility of \( 10^{-1} \text{ m}^2 \text{ V} \cdot \text{s}^{-1} \), giving an on-state conductivity \( \sigma = ne\mu \) of approximately \( 10^4 \Omega^{-1} \text{ m}^{-1} \). Also assume [13] that the voltage across the bulk of the device is about 0.1 V, i.e., a field of \( 10^9 \text{ V/m} \) taking 1 \( \mu \text{m} \) as a typical device thickness. The holding current is generally about 1 mA and hence, from the conductivity, the filament radius is in the region of 0.3–0.4 \( \mu \text{m} \). There is evidence that the current density in the on-state is constant [51], [62a]. With a current of 50 mA the filament diameter would therefore increase to about 3 \( \mu \text{m} \); the maximum on-state current quoted for pulsed operation is about 1 A and, under the same assumptions, this would correspond to a filament diameter of about 11 \( \mu \text{m} \). These figures are to be compared with the diameter of the total active area. For switches constructed in the "DO-7" package, Neale's [62] figures give a maximum diameter of 8–10 \( \mu \text{m} \); "pore-type" structures usually have an active diameter of about 20–50 \( \mu \text{m} \).

One problem always associated with small filament size is the implication of very large power densities. For a filament 1 \( \mu \text{m} \) long and 0.3–0.4 \( \mu \text{m} \) diam, a minimum holding condition of 1 V, 1 mA gives a power density of about \( 10^6 \text{ W} \cdot \text{m}^{-2} \). This power may only be applied for a short time period (e.g., 1 \( \mu \text{s} \)) but even if 90 percent of the energy is dissipated out with the filament, the remainder is sufficient to raise the filament temperature by several hundred degrees. This emphasizes one of the biggest problems of interpretation of experimental data: in addition to electronic relaxation processes during recovery from a switching event, there are also likely to be cooling processes with short (several microseconds) thermal time constants. Indeed, the relatively constant values of minimum holding power \( (P_h = V_h I_h) \) shown in Table III suggest that the minimum holding conditions may be mainly controlled by thermal effects (though this need not apply to the rest of the on state).

VI. Summary and Conclusions

There is no doubt that there are significant electronic effects causing nonohmic conduction in the preswitching region of chalcogenide glass switches. The \( I-V \) characteristics in the preswitching region can be fitted, superficially, to simple models of space-charge limited current, hopping conduction, or the Poole–Frenkel effect, but on the present evidence, at any rate, a closer examination reveals inconsistencies in each case (Section III-B-D). There is, in fact, growing evidence that the nonohmic effects in chalcogenide glasses could extend down to essentially zero fields and that a fundamentally new approach may be required (Section III-E). There is also evidence of some undetermined hot-electron effects contributing to the nonohmic behavior at "high" fields, but these are unlikely to be significant in the chalcogenide glasses most commonly used for switching devices except perhaps in pulse experiments well above the threshold voltage or in cases where there is a very nonuniform field distribution across the device [Section III-E and Section IV-C2)]. Thus at the present time it is not possible to reach any firm conclusions about the electronic mechanism(s) of the preswitching current.

It is hardly surprising, therefore, that attempts to account for the actual threshold conditions and switching process in purely electronic terms are bound to be qualitative and speculative. This is not to question the fact that there is plenty of scope, in the chalcogenide glasses, for electronic instabilities, but present theories are such that only the crudest quantitative comparisons can be made and these are only partially successful (Section IV-B, C). In the present circumstances the most useful approach seems to be to compare, wherever possible, the functional dependence of parameters such as the threshold voltage and delay time on, for example, temperature, field, and geometry with the predictions of related models for crystalline semiconductors where such comparisons are possible that they suggest that of the most likely electronic mechanisms, none are actually dominant in the switching process (Section IV-B, C).

To conclude, we draw attention again to Section II and to the final comment of Section V-B. These emphasize the difficulty, in materials like the chalcogenide glasses, of disentangling electronic and thermal mechanisms and the inevitability of some thermal effects. As mentioned earlier, our view is that there must be some convergence of the two approaches. Indeed, even in purely electronic terms there is no reason why there should not be more than one contributory mechanism with different processes being dominant in different circumstances. It is only necessary to cite the closely related subject of dielectric breakdown [2], [3] to illustrate a phenomenon in which several mechanisms
can compete and where the dominant process can change drastically as a function of, for example, temperature and/or geometry. In the present subject there was, at one time, a tendency for unique causes exclusive of all others to be propounded. Happily, that phase seems to have passed.

ACKNOWLEDGMENT

The authors wish to thank their colleague, Dr. J. M. Marshall, for his many discussions, particularly on high-field effects (Section III), and for some of the information in Section III-E and Fig. 8.

REFERENCES

[64] U. Uehtz et al., in [18, p. 388.
VARIOUS DEVICES.

THE CHARACTERISATION OF METAL-THIN INSULATOR-n-p\textsuperscript* SILICON SWITCHING DEVICES

J. BUOX, A. E. OWEN (*), G. SARRABAYROUSE and J. P. SEBAA

Laboratoire d'Automatique et d'Analyse des Systèmes du C.N.R.S., 7, avenue du Colonel-Roche, 31400 Toulouse, France

Résumé. — Les auteurs présentent des résultats expérimentaux sur les caractéristiques de commutation des dispositifs MISS comportant une couche mince d'oxyde (< 50 Å); les résultats portent notamment sur les effets thermiques, la modulation apportée par un courant de base et enfin les effets dynamiques. L'analyse physique est conduite à l'aide d'un modèle à contre-réaction dont les éléments sont décrits. Un bon accord théorie-expérience est obtenu.

Abstract. — Experiments are reported on the switching characteristics of MISS devices incorporating a thin (< 50 Å) oxide layer, including the influence of a modulating base current, the effect of temperature and the dynamic performance. A quantitative analysis of a regenerative model of switching is briefly described and shown to give a good account of the experimental results.

1. Introduction. — There have been several recent reports of potentially useful bistable switching properties in metal-insulator-n-p\textsuperscript* (or p-n\textsuperscript*) structures on silicon (i.e. MISS devices) [1-6]. It is necessary for the insulator to be either thin enough (e.g. < 50 Å) to pass substantial tunnelling currents, or for it to be otherwise semi-insulating [3], [5]. Switching occurs, at a critical voltage (the switching voltage \( V_s \)), when the p\textsuperscript*n- (or n\textsuperscript*p-) junction is biased in the forward direction; with the opposite polarity the characteristics are similar to those of a reverse-biased p-n junction.

The switching voltage can be modulated through the influence of an additional (third) contact to the intermediate Si layer [3], [4], and Yamamoto et al. [4] have shown that by utilizing this feature, MISS devices could be applied to logic systems such as shift registers. Yamamoto and Morimoto [1] have also suggested that MISS devices could be light sensitive and hence used, for example, as optically-modulated switches.

For the two terminal device, Simmons and El-Badry [5] have developed two simple electrostatic models based on punch-through and avalanche mechanisms appropriate, respectively, to lightly (< 10\textsuperscript{18} cm\textsuperscript{−2}) and heavily (> 10\textsuperscript{17} cm\textsuperscript{−2}) doped intermediate Si layers. The equations predicted by these models are, in fact, in reasonable agreement with the results of El-Badry and Simmons [6]. However the experiments described in [1-4] have been carried out with Si of medium doping levels (\( \geq 10\textsuperscript{15} \) cm\textsuperscript{−2}) and neither these results, nor our own, fit the models of Simmons and El-Badry [5].

This paper presents more detailed results than hitherto on the essential characteristics of MISS switches, particularly on their temperature dependence, the influence of a modulating base current (through a third contact) and on the dynamic performance. We also describe the basic features of an alternative theory which gives good agreement with experiment and is capable of being extended to give a complete description of both the static and dynamic switching characteristics, including the ON-state and the negative resistance region.

2. Device fabrication. — MISS switches were fabricated from Si n-epitaxial layers (doping level = 2 x 10\textsuperscript{15} cm\textsuperscript{−3} and 7 \textmu m thick) on p\textsuperscript* substrates with a thin oxide layer thermally grown on the n-type Si substrate.

* Permanent address : Department of Electrical Engineering, University of Edinburgh, King's Buildings, Edinburgh EH9 3JL, Scotland, Grande-Bretagne.

(*) Permanent address : Department of Electrical Engineering, University of Edinburgh, King's Buildings, Edinburgh EH9 3JL, Scotland, Grande-Bretagne.
3. Experimental results. — 3.1 STATIC MEASUREMENTS.

- Figure 1b is a schematic illustration of the I-V curve as obtained on a typical transistor curve-tracer; similar results would be expected from a slow-ramp or a d.c. experiment. For the present purposes the p⁺-substrate is called the emitter (E), the contact to the thin oxide the collector (C) and the n⁺-contact to the intermediate Si layer the base (B). Figures 1a and b then define the current and voltage directions with the arrows in figure 1a indicating conventional current. The device switches at the voltage \( V_S \) with a positive potential \( V_E \), applied between emitter and collector. Typical experimental results, taken from oscilloscope traces, are shown in figure 1c, which also illustrates the influence of a base current \( I_B \) with \( I_E = 0.05 \text{ mA} \) for both directions from \( I_E = 0 \). Note that the low-impedance ON-state characteristic remains unchanged.

Measurements have also been made of the static characteristics as a function of temperature with the results, for another device, shown in figure 2. The curves labelled (1) to (7) correspond to temperatures ranging from \(-193.5^\circ C\) to \(40^\circ C\) (see figure caption for details). There is a significant change with temperature, particularly in the switching voltage \( V_S \) and this is shown explicitly for another device, in figure 3. According to the present results \( V_S \) increases as temperature decreases and seems to saturate at about \(-200^\circ C\); so far, however, experiments have not been taken to lower temperature to determine whether \( V_S \) remains constant or goes through a maximum. These appreciable effects of temperature are in marked contrast to the results of El-Badry and Simmons [6]; they found essentially no change in \( V_S \), although they did observe a small effect on the ON-state characteristic but in the opposite sense to that shown in figure 2.

3.2 DYNAMIC MEASUREMENTS. — When a voltage pulse of magnitude greater than \( V_S \) is applied to an MISS switch, the current through the device remains low (essentially the OFF-state current) for a period \( t_0 \), after which it rises more or less instantaneously to the ON-state current, i.e. there is a delay time \( (t_0) \) before switching takes place. Figure 4 illustrates this for three pulses of different magnitude: as the applied pulse height increases (with respect to \( V_S \)) the delay time decreases. The shape of the current response appears to be controlled by the internal switching mechanisms of the device and is independent from the external circuit used for the measurements. In figure 4 the current seems to rise from its OFF to ON-state values in a two-stage process; this was a common but not universal feature of our measurements: for the range of pulse heights used, \( t_0 \) is of the order \( 10^{-9} \text{ to } 10^{-8} \text{ seconds} \). Kroger and Wegener [2] report delay times as short as a few nanoseconds or less, but they applied pulses rather greater in magnitude than those used for the present studies.

The base current also influences the delay time and figure 5 illustrates the combined effects of \( I_B \) and pulse height on \( t_0 \). In figure 5a \( I_B \) is plotted as a function of
CHARACTERISATION OF METAL-THIN INSULATOR-n-p' SILICON SWITCHING DEVICES

4. Discussion. — It seems clear that the punch-through and avalanche models of MISS switching, developed by Simmons and El-Dadry [5], do not apply to our own results nor, most likely, to the results of other workers [1-4]. Yamamoto and Morimoto [1] have already tentatively suggested the possibility of a regenerative mechanism of switching related to the build-up of an inverted region in the Si at the Si-SiO₂ interface. In the following we give a brief description of the quantitative development of such a model. The essential features of the model are given in figure 6 which shows the band structure of the OFF- and ON-states; it also defines the parameters used in the following discussion. It is necessary to solve the continuity equations for both the majority \( J_m \) and minority carrier current \( J_m \), taking into account the link between \( J_m \) and \( J_m \) imposed by the presence of the p' n junction. The numerical solution of these equations has been carried out and a continuous description of the static I-V characteristic is obtained, including the negative resistance region. Preliminary calculations also show that the regenerative model is capable of describing pulsed operation (dynamic performance) but for reasons of space the present discussion is restricted to considering the switching voltage \( V_s \) and the ON-state characteristic.

Throughout the OFF-state characteristic the voltage \( V_{sc} \) is practically confined to the MIS part of the structure. The continuity equation for the transport

\( I_d \) (positive and negative) with pulse height as a parameter. For completeness figure 5a shows the variation of switching voltage \( V_s \) as a function of base current; \( I_d \) can be regarded as influencing \( V_s \) through its effect on \( V_s \).
of minority carriers has been discussed previously [7], [8], and for the present case it is written as follows:

\[ J_s = \alpha^* a_s T^2 \exp \left[ - \frac{(\phi_w - \chi - V_s)}{kT} \right] \times \left[ 1 - \exp(-U_{4i}) \right] \]

\[ = A^* a_s T^2 \exp(-U_{4i}) \left[ 1 - \exp(-U_{4i}) \right] + J_4 \]

(1)

where \( \gamma \) is the \( p^+ n \) junction efficiency [9], \( J_s \) the fraction of the diffusion current that tunnels via the acceptor state density [10], \( \alpha \) and \( \alpha_s \) are tunnelling attenuation factors [7]. \( A^* \) is the Richardson constant, \( T \) the temperature and \( k \) the Boltzmann constant.

The other parameters are as defined in figure 6. Equation (1), together with the expression which relates the voltage across the oxide layer, \( V_s \), to the applied voltage \( V_4 = (kT/q) U_{4i} \) [8], gives a description of the OFF-state characteristic. As soon as the inversion situation occurs, \( U_{4i} = U_{4i} + U_4 \) (where \( U_{4i} \) is the saturation value of \( U_{4i} \) at inversion), the majority carrier driving current of the regenerative system, \( J_{4i} \), will tend to increase rapidly because the inversion layer imposes larger field values across the oxide, thereby making \( J_s \) decrease. The minority carrier current given by equation (1) will then have to adapt to the condition \( U_{4i} = U_{4i} + U_4 \), which remains unchanged once the inversion situation is reached. As a result, the control of \( J_{4i} \) is obtained with much lower values of \( U_4 \) in order to satisfy equation (1), i.e. the system switches to the ON-state. By setting \( U_{4i} = U_{4i} + U_4 \) in equation (1) we can derive the voltage \( V_s = (kT/q) U_{4i} \) at which the inversion situation is first established, i.e. the switching voltage. We obtain:

\[ V_s = \frac{\gamma N_C N_A}{2 q N_D N_A} \times \left[ \frac{N_C}{N_A} \right]^{1/2} \exp \left[ \frac{kT \ln \left( \frac{N_C}{N_A} \right)}{N_A \left[ 1 + (J_{4i} A^* T^2 a_s) \right]} \right] \]

\[ + kT \ln \left( \frac{N_C}{N_A} \right) \left[ \frac{N_C}{N_A} \right] \left( N_A - (\phi_w - \chi) \right) \]

(2)

where \( N_C \) and \( N_A \) are the densities of state in the valence and conduction bands respectively, \( N_D \) is the doping level, and \( \varepsilon_0 \), \( \varepsilon_r \), \( \varepsilon_0 \) are the insulator, silicon and vacuum permittivities.

\[ J_s = \alpha^* a_s T^2 \exp(-U_{4i}) \left[ 1 - \exp(-U_{4i}) \right] + J_4 \]

where \( N_A^0 \) is the acceptor state density, \( \alpha \) the acceptor capture cross section and \( \varepsilon_r \) the thermal velocity. [10]. With reasonable values of the parameters involved, equation (2) gives a good approximation to the magnitude of \( V_s \) and its temperature dependence as can be seen from figure 3. The dashed line in figure 3 was calculated from equation (2), with the variables set at the values given in the caption. It is possible to obtain an even better fit between theory and experiment by including, for example, the variation in the width of the space-charge region in the Si at the Si-SiO\(_2\) interface, but the more cumbersome computation involved has not yet been completed. For the present purposes it is sufficient to note that the model leading to equation (2) provides an adequate description of the experimental data on \( V_s \). It could also be extended to include the influence of base current \( I_b \) on \( V_s \).

The ON-state can be analysed but a new boundary condition for the diffusing minority carriers must be introduced because of the collapse of the barrier at the oxide-semiconductor interface. The appropriate equation is:

\[ \gamma A^* a_s T^2 \exp \left[ - \frac{(\phi_w - \chi - V_s)}{kT} \right] \times \left[ 1 - \exp(-U_{4i}) \right] \left( \frac{qP_w}{L} \right) \]

\[ = A^* a_s T^2 \exp(-U_{4i}) \left[ 1 - \exp(-U_{4i}) \right] + J_4 \]

(3)

where \( P_w \) is the diffusion coefficient of the minority carriers, \( L \) the width of the intermediate \( (n) \) Si layer and \( \rho_w \) the minority carrier density at the edge of the MIS space-charge region.

A more complete account of the regenerative model, including its application to the effect of a base current, the negative resistance region and dynamic operation, will be published elsewhere. Clearly however, it is capable of providing a good approximation to our results and, possibly, to those of other workers. Since the models of Simmons and El-Badry also fit reasonably well to their particular results [5], [6], one must conclude, for the time being, that there are at least three operating mechanisms of MIS switches: punch-through, regeneration and avalanche mechanisms. Which one is dominant presumably depends on device parameters such as the doping level and width of the intermediate Si layer.

References

THE THRESHOLD CHARACTERISTICS OF CHALCOGENIDE-GLASS MEMORY SWITCHES

A.E. OWEN *
Laboratoire d'Automatique et d'Analyse des Systèmes, CNRS, 7 avenue du Colonel-Roche,
31400 Toulouse, France

J.M. ROBERTSON and C. MAIN **
Department of Electrical Engineering, University of Edinburgh, King's Building, Edinburgh
EH9 3JL, UK

Received 18 July 1978

A detailed study is reported of the threshold characteristics of memory switches made from a chalcogenide glass of composition Ge$_{1.5}$Te$_8$S$_2$Sb$_2$. Most of the devices were fabricated by microelectronic processing techniques compatible with integrated-circuit technology. The main experimental feature is the systematic variation of the factors which influence the thermal parameters of the devices, i.e. temperature, geometry (thickness and radius of the active region of the switch) and substrate material. Measurements are also reported of the temperature and field dependence of the isothermal conductivity of the chalcogenide glass. Analysis shows that the results can be described adequately by an electrothermal model which takes into account the full field dependence of the isothermal conductivity and of the thermal fringing in the substrate.

1. Introduction

The past 10-15 years has seen the field of "amorphous semiconductors" develop into a distinct field of solid-state electronics and physics. Much of the initial impetus arose from the great interest caused by electrical switching phenomena in chalcogenide glasses, although reports of a form of electrical switching in variety of materials, amorphous and crystalline, have appeared in the literature for at least several decades. The first references to a relatively crude type of switching in chalcogenide glasses were reported nearly 20 years ago [1-3], but Ovshinsky's paper of 1968 [4], which described the prototype of a potentially reliable device,
probably marked the point at which the subject became of serious interest in solid-state electronics.

Chalcogenide glass switches are normally classified into "threshold" (or mono-stable) and "memory" (or bi-stable) devices. Their general characteristics are well-known, but for convenience the familiar schematic $I-V$ curves are reproduced in fig. 1a for the threshold device, and in fig. 1b for the memory switch; $V_{TH}$ is the threshold voltage, curves 1 and 2 indicate the OFF- and ON-states respectively and, in fig. 1a $I_h$ is the holding current and $V_h$ the holding voltage for the threshold ON-state.

Whether threshold or memory characteristics are observed is mainly a matter of the composition of the chalcogenide glass. A typical threshold composition is $\text{As}_{30}\text{Te}_{48}\text{Si}_{12}\text{Ge}_{10}$ (the so-called STAG glass) [4], whereas memory materials often have a composition close to $\text{Ge}_{15}\text{Te}_{81}\text{X}_4$, where $X$ can be a variety of elements such as As, Sb, S etc. or even a mixture of these [5,6]. The reason why different compositions are usually more suitable, either for threshold devices or for memory switches, can be understood qualitatively in terms of their glass-forming tendencies. The prismatic diagram of fig. 2 shows, for example, the approximate glass-forming region in the four-component system Ge—Si—Te—As.

The glass-forming regions in the two ternary systems Ge—Te—As and Si—Te—As have been determined experimentally [7,8], but in the four-component system it has been drawn in by a reasonable interpolation *.

The STAG threshold glass corresponds to point A in fig. 2; it is near the centre of a large glass-forming region and hence is unlikely to devitrify very readily, i.e. the STAG glass is expected to be relatively stable to crystallization. The memory glass $\text{Ge}_{15}\text{Te}_{81}\text{As}_4$ corresponds to point B; it is at the edge of a glass-forming region (as defined here — see footnote) and hence on the one hand it will readily crystallize at temperatures close to its glass-transformation range, but, on the other hand, it can be quenched to form a glass without too much difficulty. It is not the purpose of this paper to discuss memory-switching, but it should be noted that bi-stable operation (fig. 1b) first involves switching to the threshold ON-state followed by the precipitation of a crystalline filament in the hot ($\geq T_g$) ON-state channel. Thus, SETTING to the permanent ON-state requires a current pulse of several milliamps for perhaps a millisecond or longer; to return to the OFF-state (RESET) requires a much larger but shorter (1–10 µs) and rapidly terminated pulse, in order to remelt the crystallized material and quench it into the glassy state [5]. Thus the choice of a material for memory-switching calls for a fine balance between the ease with which it crystallizes and the tendency to form a glass. A "good" threshold material is one which, other things being equal, is as "stable" as possible in its vitreous form.

* It must be recognized, of course, that 'glass-forming ability' is not well-defined notion. The boundaries of fig. 2 correspond to the preparation of at least several grams of a glass by normal quenching from the melt. Vapour deposition or special techniques for rapidly quenching small droplets of a melt would considerably expand the glass-forming region.
For example, another material used for threshold switching is the five-component composition $\text{As}_{36}\text{Te}_{39}\text{Si}_{17}\text{Ge}_{7}\text{P}_{1}$ [9]. It is difficult to relate this to the four-component diagram of fig. 2, but it is worth noting that the addition of extra components in order to stabilize a glass towards crystallization is a well-established trick of the traditional glass-maker's art and was probably discovered empirically long before such esoteric concepts as "configurational entropy" came on the scene. Ultimately, however, all glasses are thermodynamically unstable with respect to the crystalline state (or, at the most, metastable), and the distinction between threshold and memory materials is one of degree rather than of kind. In our experience, a memory switch can always be operated satisfactorily as a threshold device, provided the ON-state current is kept low, and a threshold switch can be made to function
bi-stably, albeit with a limited number of cycles between the OFF-state and the permanent ON-state (SET and RESET).

The qualitative “explanation” of the difference between threshold and memory-switching materials in terms of their glass-forming tendencies does, of course, beg the important question — i.e. why are some materials much more easily prepared in a glassy form, and much less readily crystallized, than others? This is a complex problem involving structural, kinetic and thermodynamic factors and is beyond the scope of this paper. Nevertheless, what is known about glass-forming tendencies does provide a useful empirical correlation with switching characteristics.

2. Device structures and scope of the present work

The work to be described here was part of a project concerned primarily with memory-switching, and hence the composition $\text{Ge}_{15}\text{Te}_{38}\text{Se}_{2}\text{Sb}_{2}$ was used for the majority of measurements (some related glasses in the Ge–Te–As system have also been studied but will not be reported on here). In most cases the switching devices were fabricated by standard microelectronic processing techniques in the so-called “pore” structure [10]. The active area of the switch was defined by etching a small pore (radii varying from 5 to 75 $\mu$m in the present work) through a suitable insulator. A few devices of larger area were also defined by the cross-over of ortho-
gonally deposited upper and lower metal strip electrodes. The chalcogenide glass was deposited by rf sputtering, as also were the Mo electrodes. Evaporated Al contacts were provided between the lower Mo electrode and the substrate, and on top of the upper Mo electrode. The substrates were either borosilicate glass slides (Corning 7059) or silicon slices with a thin (0.4 μm) thermally grown oxide layer for electrical isolation.

As mentioned in section 1, the establishment of a permanent ON-state in a memory switch is preceded by the threshold action, and it is this aspect of the work which is the subject of this paper. The threshold voltage has been measured as a function of temperature and the delay time \( t_D \) (under pulsed operation) as a function of voltage, with systematic variations of thermal parameters such as substrate thermal conductivity and device geometry. From the interpretative viewpoint the aim was to determine to what extent the threshold characteristics could be quantitatively described by simple thermal or electrothermal models which take into account the thermal influence of the device geometry and substrate. Sections 3 and 4 give a brief account of, respectively, one-dimensional thermal and electro-thermal theories, and of the appropriate thermal conductance calculations. The experimental results and analyses are considered in subsequent sections.

3. Thermal and electrothermal switching

Figure 3 shows a typical \( I-V \) curve obtained from a pore device using a slow current ramp. The turnover and current-controlled negative differential resistance region (CCNDR) were observed, reproducibly, in all devices. A thermal theory attempts to explain this in terms of Joule heating, with switching following as an instability associated with thermal runaway. The cylindrically symmetric geometry analysed is shown schematically in fig. 4. A thin film of chalcogenide, thickness \( d_f \), thermal conductivity \( K_f \) and specific heat capacity \( C_f \) is deposited on a substrate, with the parameters \( d_s, K_s \) and \( c_s \), respectively. The sandwich electrodes are circular, radius \( a \), and the substrate lower surface is in contact with a heat sink at the ambient temperature \( T_a \). The heat transport equation is

\[
\nabla^2 (K T) + \int_0^a \rho = c(\partial T/\partial t),
\]

where \( \rho \) is the electrical current density.

The following simplifying and self-consistent assumptions are made:

(1) The temperature distribution in the active region of the device is uniform, i.e.

\[
\begin{align*}
(\partial T/\partial r) &= 0 \\
(\partial T/\partial z) &= 0 \\
(0 < z < d_f) \\
(0 < r < a)
\end{align*}
\]

(2) The upper surface, \( z = 0 \), is adiabatic

\[
(\partial T/\partial z) = 0 \quad (z = 0).
\]
Fig. 3. A typical $I-V$ curve for a Ge$_{1.5}$Te$_{81.5}$S$_2$Sb$_2$ pore device on a borosilicate glass substrate measured using a slow-current ramp. The turnover voltage is indicated by $V_t$ and the switching voltage by $V_{SW}$ (see text).

(3) Current and heat flow in the device are parallel and uniform.

(4) The electric field is uniform through the device.

(5) The electrodes are relatively poor heat sinks.

While these assumptions seem reasonable for thin-film structures, their validity can in most cases be checked experimentally, or theoretically, as will be shown later. The first and third assumptions imply no current filamentation, which is probably the case prior to turnover, while the fifth is required to provide CCNR, as electrodes held at ambient preclude turnover [11]. In the steady state, eq. (1) reduces to the heat balance equation:

$$(T - T_a) = \frac{VI}{F} = g(V, I),$$

where $T$ is the temperature of the active region, $T_a$ the ambient temperature, $F$ the overall thermal dissipation constant and $g$ represents some function of $V$ and $I$. If $G$ is the electrical conductance the device current $I$, is

$$I = VG = f(V, T).$$

Turnover occurs when

$$\left(\frac{\partial f}{\partial T}\right)\left(\frac{\partial g}{\partial V}\right) = 1.$$

Thermal runaway will occur at this point at a voltage $V_t$, for a constant voltage (zero resistance) source. This is a limiting case of a load-line instability.

The simplest case of field-independent activated conductivity gives the well-
known thermistor results. Conductivity:

\[ G = (\pi a^2/d_f) \sigma_0 \exp(-\Delta E/kT). \] (3)

Turnover temperature:

\[ T_t \approx T_a + (T_a^2 k/\Delta E). \] (4)

Turnover voltage:

\[ V_t \approx 0.6 T_a \left( \frac{kT d_f}{\Delta E \pi a^2 \sigma_0} \right)^{1/2} \exp(\Delta E/2kT_a). \] (5)

Note that, assuming \( \Gamma \) is constant, eq. (5) gives \( V_t \propto d_f^{1/2} \). While these results may apply to thick films and bulk samples \( (d_f \geq 10 \mu m) \), in general they do not apply to thin films \( [12] \), even qualitatively.

An electrothermal analysis takes into account the field-dependent conductivity of the active material. A number of functional forms for this dependence have been reported in the literature, but in many chalcogenides it appears to be given by \( [13] \).

\[ \sigma(T, \epsilon) = \sigma_0 \exp[-(\Delta E - c_a(T)\epsilon)/kT], \] (6)

where the field parameter \( c_a(T) \) has the dimensions of length and is temperature dependent. That eq. (6) describes a bulk effect has been verified by using a wide range of interelectrode separations \( (0.3-25 \mu m) \). At high fields \( (\epsilon > \epsilon_1 = 2.5-3 \times 10^5 \text{ V/cm}) \) a second, steeper field dependence becomes dominant in some chalcogenides, i.e.

\[ \sigma(T, \epsilon) = \sigma_0 \exp[-(\Delta E - c_a(T)\epsilon_1)/kT] \exp[h(\epsilon - \epsilon_1)/kT], \] (7)

where \( h \) is also a constant with the dimensions of length. The mechanisms responsible for the field-dependent conductivity have not yet been clearly identified, but such detailed knowledge is not required for an electrothermal analysis.

Substituting eq. (6) into the heat balance equation gives, at turnover, turnover
temperature:

\[ T_t \approx T_a + (T_a^2 k/\Delta E'), \]  

\[ \Delta E' = \Delta E - (a_0 V_t e/d_f). \]  

Turnover voltage:

\[ V_t = \frac{2d_f k T_a}{e(a_0 - d T_a)} \left[ \frac{\Delta E}{2k T_a} + \log \left( \frac{0.6T_a}{V_t} \left( \frac{k d_f \Gamma}{\Delta E a_0 \pi a} \right)^{1/2} \left( 1 - \frac{V_t e a_0}{2d_f \Delta E} \right)^{-1} \right) \right]. \]  

Note that \( V_t \) is now much less temperature sensitive, as

\[ (\partial V_t/\partial T_a) < [d_f \Delta E/eT_a(a_0 - d T_a)] \]

and

\[ V_t \propto d_f \quad (\Gamma \text{ constant}). \]

A similar expression is obtained for the “second field regime” [eq. (7)], giving an even smaller temperature dependence. The behaviour predicted for \( V_t \) is in qualitative agreement with a number of observations on thin chalcogenide devices, but a quantitative prediction of the behaviour depends upon the value of \( \Gamma \) and its dependence upon device parameters, and this is considered next.

4. Thermal conductance calculations

The overall thermal conductance of the device of fig. 4 consists of the series combination of film–substrate–heat-sink conductances, \( \Gamma_f \) and \( \Gamma_s \), respectively. The smaller of the two terms will control the overall dissipation, and hence \( V_t \).

For uniform device heating, and heat flux, the minimum value of \( \Gamma_f \) is obtained:

\[ \Gamma_f(\text{min}) = (2K_f \pi a^2 /d_f). \]  

While for uniform, linear heat flow in the substrate:

\[ \Gamma_s = (K_s \pi a^2 /d_s). \]  

As Thornburg and Johnson point out [14] this is correct only for \( a \to \infty \), because non-uniform heat flow, or fringing, can increase the effective thermal conductivity of the substrate. A simple approach to the problem of thermal fringing is to assume that the heat flow near the device is uniform over a hemispherical isotherm of radius \( a \). Placing a mirror image of the heat-sink isotherm above the device produces a thermal flux pattern, in a “homogeneous” medium, which maintains the boundary conditions \( (\partial T/\partial Z) = 0 \). The effective thermal conductance \( \Gamma'_s \) between this isotherm and the heat sink may be found by the method of images, where, in this case, an infinite series of images is produced on either side of the device.
Summing the infinite series obtained, term by term (it is conditionally convergent), the result obtained is

$$\Gamma_s^* = \left[2K_s \pi a/(1 + (a/d_5) \log 2)\right].$$

(13)

The effective conductivity of the substrate is increased over that for linear flow by the factor

$$\left[2(d_5/a)/(1 + (a/d_5) \log 2)\right].$$

(14)

This factor can be very large for typical geometries (e.g. for $a = 50 \mu m$, $d_5 = 0.085 cm$ the factor is $\approx 35$). For typical device geometries ($a \gg d_5$),

$$\Gamma_s^* \approx 2K_s \pi a.$$

(15)

The two limiting cases of a poorly conducting substrate ($\Gamma_t \gg \Gamma_s$) and a highly conducting substrate ($\Gamma_s^* \gg \Gamma_t$) may now be examined for the simple thermal model. In practice, the use of borosilicate glass and silicon substrates correspond, respectively, to these two cases. The discriminating criterion is given by

$$(K_t a/K_s d_5) > 1, \quad \Gamma_t > \Gamma_s.$$

4.1. Poorly conducting substrates

The overall conductance $\Gamma \rightarrow \Gamma'$, and using eq. (15)

$$V_t \approx 0.6 T_a \left(\frac{2K_s d_5}{K_{Ta}}\right)^{1/2} \exp(\Delta E/2K_{Ta}).$$

(16)

Note that this predicts a dependence on device diameter. The electro-thermal case will, of course, be less sensitive to this parameter.

It can now be shown that the assumption of uniform device temperature is well justified in this case. At the film–substrate interface the heat flows from the film and through the substrate may be equated:

$$K_F \left[\frac{\partial T}{\partial z}\right]_{z=d_5} = (\Gamma_s a)^2 \left[T(z) - T_a\right]_{z=d_5}.$$

The maximum temperature drop across the film can be estimated [14]:

$$[T(0) - T(d_5)] < 0.5 d_5 \left[\frac{\partial T}{\partial z}\right]_{z=d_5}.$$

Hence, the ratio of this temperature drop to that across the substrate is

$$(\Delta T_f/\Delta T_s) < (K_s d_5/K_{Ta}).$$

Note that this ratio is much larger than that estimated by Hayes and Thornburg [15] for uniform heat flow, but for a typical device on a glass substrate ($d_5 \sim 1 \mu m$, $a \sim 50 \mu m$, $K_s \sim K_{Ta}$) this ratio is still small ($\sim 0.02$).
Table 1

<table>
<thead>
<tr>
<th>Material</th>
<th>Thermal conductivity $K$ (W cm$^{-1}$K$^{-1}$)</th>
<th>Thermal capacity $C$ (J cm$^{-3}$K$^{-1}$)</th>
<th>Refs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>chalcogenide glass</td>
<td>$2 \times 10^{-3}$</td>
<td>1.0</td>
<td>[16], [17]</td>
</tr>
<tr>
<td>borosilicate glass</td>
<td>$1 \times 10^{-2}$</td>
<td>2</td>
<td>[19]</td>
</tr>
<tr>
<td>Si</td>
<td>1.5</td>
<td>2</td>
<td>[18]</td>
</tr>
</tbody>
</table>

4.2. Conducting substrates $\Gamma_s > \Gamma_f$

The overall conductance $\Gamma \rightarrow \Gamma_f$,

$$V_{f}(\text{max}) = 0.6 \frac{J_f}{T_a(2kK/dT_f)} \exp(\Delta E/2kT_a). \quad (17)$$

The turnover voltage is now independent of device diameter, although the “electrothermal” $V_f$ will still be proportional to $d_f$. This expression is less satisfactory than that for poorly conducting substrates as there will now be a significant temperature drop in the film itself, and $\Gamma_f$ may be larger than the computed value. For example, for a typical device on a silicon substrate, the ratio:

$$(\Delta T_f/\Delta T_s)_{\text{max}} \approx 1.5 \quad (a = 50 \mu m).$$

However, the value derived for $(\Delta T_f/\Delta T_s)$ is a “worst case” result, and a number of factors will tend to reduce this ratio in practice. For example, the temperature dependence of the electrical conductivity will ensure that power dissipation in the cooler region near the substrate interface is greater than that in the upper, hotter region. It should be noted that since, in this case, the thermal resistance of the film itself controls $\Gamma$, then variations in the thermal properties of the surrounding medium would not be expected to affect switching behaviour.

Table 1 lists approximate values of thermal properties of the materials used.

5. The conductivity of Ge$_{15}$Te$_{81}$S$_2$Sb$_2$ glass

The temperature and field dependence of the isothermal conductivity of the Ge$_{15}$Te$_{81}$S$_2$Sb$_2$ glass was measured in the same structures used for the switching experiments. The results are shown in figs. 5 and 6. The power dissipation was kept low in these measurements in order to avoid Joule heating – e.g. short (0.3 $\mu$s) voltage pulses were used at room temperature (note, in fig. 6, the different results for static and pulsed measurements at 296 K). The conductivity has the functional form of eq. (6) for $\epsilon < \epsilon_1 \approx 2.5 \times 10^5$ V/cm, while $\epsilon > \epsilon_1$, it follows eq. (7).

From the experimental data in figs. 5 and 6, the values of the parameters in
Fig. 5. The logarithm of the isothermal conductivity versus reciprocal temperature of the Ge₅Te₈S₂Sb₂ glass. The different points (○, ●, and ○) correspond to devices with glass films of different thicknesses in the range 0.3–5 μm. The slope of the line gives an activation energy of 0.44 eV.

Fig. 6. Field dependence of the isothermal conductivity of the Ge₅Te₈S₂Sb₂ glass, with temperature as a parameter.
eq. (6) and (7) are

\[
\sigma_0 = 600 \text{ (ohm-cm)}^{-1} \quad \text{and} \quad \Delta E = 0.44 \text{ eV},
\]

\[
\sigma(T) = \sigma_0 - dT, \quad \text{with} \quad \sigma_0 = 32 \text{ Å} \quad \text{and} \quad d = 7 \times 10^{-2} \text{ Å K}^{-1}.
\]

\[
\epsilon_1 = 2.5 \times 10^5 \text{ V cm}^{-1}, \quad \hbar = 85 \text{ Å}.
\]

6. Switching and 'turnover' in Ge\textsubscript{15}Te\textsubscript{85}S\textsubscript{2}Sb\textsubscript{2} devices

Static dc \textit{I–V} measurements were made by standard voltmeter–ammeter methods, and also by a slow-ramp method employing the 150 V sweep output from a Tektronix 549 storage oscilloscope through a switched bank of limiting resistors. Scan times of 1–10 s were used.

6.1. Turnover voltage \(V_t\) versus temperature

Figure 7 shows the temperature dependence of \(V_t\), for a device with \(a = 50 \mu\text{m}\) on a glass substrate. The measured thermal conductance for this device is \(\Gamma = \)

![Figure 7. Turnover voltage \((V_t)\) versus temperature (on a log-linear scale) for devices on borosilicate glass substrates. The different points (○, □) are both for pore devices of radius \(a = 50 \mu\text{m}\) and thickness \(d_f = 0.95 \mu\text{m}\). The solid line (1) correspond to \(V_t\) versus \(T\) calculated from the simple thermal model [eq. (5)], line (2) to the single-slope electrothermal calculation [eq. (10)] and line (3) to a calculation which takes into account both slopes in the field dependence [i.e. eq. (7)].](image)
0.33 mW K$^{-1}$, a value obtained by comparing dc and pulsed $I-V$ characteristics. Using this value for $\Gamma$, the turnover voltage was computed using the following:

1. Simple thermal runaway - eq. (5).
2. Single exponential field dependence - eq. (6) and (10).
3. Total isothermal characteristics - eq. (7) and (10).

At low temperatures, when $V_t \geq 7$ V, the simple thermal prediction departs rapidly from the observed $V_t$, while the "single-slope" electrothermal model fits well, up to $\sim 25$ V (corresponding to $\epsilon_1$). Computed curve (3), incorporating the second steep high-field regime, agrees well with the observed data over the whole temperature range.

Calculation of the device thermal conductance, using eqs. (11) and (13) gives

$$\Gamma_f = 6.4 \text{ mW K}^{-1}, \quad \Gamma_s' = 0.3 \text{ mW K}^{-1}.$$ Here, the substrate conductance controls cooling and is in very good quantitative agreement with the measured value. Note that $\Gamma_s$ is $\sim 35$ times greater than its value for linear heat flow.

The device behaviour can be described the three 'regimes' 
1. $V_t < V$. Simple thermal runaway.
2. $7$ V $< V_t < 25$ V. Single-slope electrothermal runaway.

This is not merely an exercise in curve-fitting. It demonstrates that runaway (switching) under these circumstances may be fully described by Joule-heating considerations alone.

### 6.2. Turnover voltage versus temperature and film thickness

Figure 8 shows $V_t$ versus $1/T$ for a range of film thicknesses ($0.3$ $\mu$m$-$4.6 $\mu$m, $a = 50$ $\mu$m, glass substrate). The expected slope predicted by the simple thermal model is also shown ($0.22$ eV). It can be seen that this behaviour, corresponding to regime 1, is approached as $d_t$ increases and as temperature increases, while for thinner films and/or lower temperatures, the temperature dependence is less, as predicted by the electrothermal model.

Figure 9 depicts the film thickness dependence of $V_t$, at fixed temperatures. At room temperature and above, $V_t \propto d_t^{1/2}$ approximately, while at lower temperature, $V_t \propto d_t$. This behaviour is again consistent with the thermal analysis, tending to "simple" thermal turnover (regime 1) for higher temperatures and thicker films, and electrothermal turnover at lower temperatures and for thinner films.

### 6.3. Turnover power versus film thickness

The data given above were for devices on a poorly conducting substrate ($\Gamma_s < \Gamma_f$). For this situation, the turnover power $P_t$ may be written

$$P_t = \frac{V_t I_t}{\Gamma_s (T_1 - T_a)},$$
$$P_t = \frac{(2K_s m a T_0^2 k)}{[\Delta E - (ea(T) V_t/d_t)]}.$$
Provided $V_1 \ll (\Delta E_d/ea(T))$, the turnover power is independent of field parameters, and allows an easy check on $\Gamma$. Note that in this case, $\Gamma'(\sim \Gamma'_d)$ and hence $P_t$, should be independent of film thickness. Experimentally, $P_t$ was found to be between 5 and 6 mW, while the value calculated from eq. (18), is $\sim$5.8 mW. Fig. 10 plots $P_t$ versus $T_a$ for several film thicknesses. Although there is some scatter, the data are practically independent of film thickness; the curve through the data is obtained from eq. (18).

Fig. 8. Turnover voltage $V_t$ as a function of reciprocal temperature for devices of thickness 4.6 $\mu$m (○), 1.7 $\mu$m (●), 0.95 $\mu$m (□) and 0.30 $\mu$m (▲) on borosilicate glass substrates. The dashed line has a slope corresponding to an activation energy of 0.22 eV.

Fig. 9. Device thickness versus turnover voltage for devices on borosilicate glass substrates, with temperature as a parameter. The dashed line shows the threshold voltage measured under pulsed operation $V_{tp}$ at 295 K (see text).

Fig. 10. The turnover power $P_t$ versus ambient temperature $T_a$ for devices on glass substrates. The different points (○, ●, □ and ▲) correspond to devices of different thickness in the range 0.3–5 $\mu$m. The solid line was calculated from eq. (18).
6.4. Turnover voltage versus device radius

6.4.1. Borosilicate glass substrates

Figure 11 plots $V_t$ versus pore radius, varying from 7 V for a 75 μm radius, to ~22 V for a 5 μm radius pore. For a simple thermal model, Eq. (16) predicts, $V_t \propto a^{-1/2}$, and hence fig. 12 plots $V_t^2$ versus $a^{-1}$, including data for a 500 μm crossover device. In the region where a simple thermal model is appropriate, a good linear fit is obtained. The intercept for $(1/a) = 0$ should give $V_t(\infty)$, i.e. $V_t$ for a device of infinite radius and corresponding to linear heat flow, for which $\Gamma_0 = (K_s/d_s)$. From fig. 12, $V_t(\infty) \approx 2.5-3$ V, while using $\Gamma_0$ in eq. (5) gives $V_t(\infty) \approx 2.2$ V, in reasonable agreement.

Once again the thermal fringing analysis gives good agreement with the observed turnover, even though the expression for fringing conductance $\Gamma'_0$ is a simple approximation. It is pertinent to compare the values for $\Gamma'_0$ obtained from the data for turnover power and eq. (18), and the values predicted by the fringing analysis. This is shown in fig. 13. The agreement is reasonable for the larger devices, but for small-area devices the actual thermal conductance is somewhat higher than that predicted. This may be due to the approximation used, and the contribution of the electrodes.

6.4.2. Silicon substrates

An identical range of "variable-radius" pore devices were fabricated on Si substrates incorporating a thin insulating layer of SiO₂ (~0.4 μm). The results, $V_t$ versus pore radius, are shown in fig. 11. Higher turnover voltages are obtained, and

![Fig. 11. Turnover voltage $V_t$ as a function of pore radius $a$ for devices of thickness $d_t = 1$ μm. The points ○ are for borosilicate glass substrates and △ are for silicon substrates.](image-url)
Fig. 12. $V_t^2$ versus reciprocal of pore radius for devices of thickness $d_f \approx 1 \, \mu\text{m}$ on glass substrates.

$V_t$ is much less sensitive to pore radius than for devices on glass substrates. In this case $\Gamma_f$, the device–substrate conductance, controls cooling. Including the thermal resistance of the 0.4 $\mu\text{m} \, \text{SiO}_2$ layer, in $\Gamma_f$,

$$\Gamma_s = 47 \, \text{mW} \, \text{K}^{-1}, \quad \Gamma_f = 4.5 \, \text{mW} \, \text{K}^{-1}$$

(for $d_f = 1 \, \mu\text{m}$, $a = 50 \, \mu\text{m}$, $d_s = 0.02 \, \text{cm}$).

According to eq. (17), $V_t$ is independent of device radius, and the computed value at 295 K is $V_t$ (thermal) = 49 V, while $V_t$ (electrothermal) = 24 V. For small pore diameters the measured $V_t \approx 24 \, \text{V}$, but is slightly less for larger devices.

The power density (mW/cm$^2$) at turnover in these devices is independent of pore radius, which is to be expected if the film itself controls cooling, while for the glass

Fig. 13. The variation of thermal conductance with pore radius for devices of thickness $d_f \approx 1 \, \mu\text{m}$. ○, ● – glass substrates, △ – silicon substrates. The solid lines were computed from eq. (18).
substantially the power density increases with decreasing pore radius. Also plotted in Fig. 13 are the device thermal conductances, obtained from the turnover power and eq. (18), and the theoretical $\Gamma_f$ (min). The agreement is good, giving $\Gamma_o a^2$ as predicted.

6.5. Turnover voltage versus temperature and device thermal parameters

It is now appropriate to consider how the $V_t$ versus $T$ dependence is affected by varying the pore radius. Fig. 14 shows $V_t$ versus $T$ for devices on glass — a 500 $\mu$m crossover device, 50 $\mu$m and 10 $\mu$m radii pore devices — and for a 50 $\mu$m radius pore device on silicon. The features evident in this figure are: (1) at low temperatures, in regime 3, $V_t$ is relatively insensitive to thermal parameters; and (2) the transition to regimes 2 and 1 occurs at lower temperatures for devices with inefficient cooling. It is clear that very different results may be obtained with devices of different radii, even when the film thickness is the same — a fact which has not often been considered.

7. Surface temperature measurement

A surprising aspect of the work reported is that the results imply that the metal electrodes do not necessarily provide the major heat dissipation path. Substrate and film parameters alone can predict turnover behaviour, if "fringing" effects are
included. As the thermal conductivity of the electrode material is typically $10^3$ times higher than the film, or glass substrates (e.g. for Mo, $K \approx 1.4$; for chalcogenides $K \approx 2-5 \times 10^{-3}$ W cm$^{-1}$ K$^{-1}$), it might seem as though the electrodes could be treated as isothermal infinite heat sinks, held at ambient. However, in the absence of field effects this precludes the observed turnover and CCNDR [11], an argument sometimes used against thermal models, and the assumptions mentioned above must be checked.

1. Electrodes held at ambient? As the electrodes in this work are typically deposited on a substrate of poor thermal conductivity this assumption is incorrect — i.e. the electrodes must heat up with the device.

2. Infinite heat sink? The electrodes will probably have a smaller thermal capacity than the substrate.

3. Electrodes are isothermal and good lateral thermal conductors? Initially, this seems to be the case, as the electrode material has a very high thermal conductivity. The point in question is whether the thermal conductance of the electrodes is comparable to that of the substrate (not the film).

It is impossible to treat the two elements — electrode and substrate — separately, as they are strongly coupled along the electrode length. This is the key to the problem. Whichever element dominates the cooling will control the temperature distribution around the device, and quite different patterns are to be expected if the substrate, or electrodes, are dominant. Where electrode conduction dominates, the temperature gradient along the electrodes will be much less than that along the substrate, and the tangent to the isotherms will change direction rapidly at the electrode pattern edges. Where substrate conduction dominates, the isotherm pattern should be almost circularly symmetric.

To check this experimentally, a thin layer of cholesteric temperature-sensitive liquid crystal was deposited on top of the pore devices by spinning. The material used indicated isotherms in the temperature range $36-40^\circ$C, by changing colour through the spectrum from red to blue. The device (top electrode) temperature and lateral temperature distribution were monitored during $I$--$V$ curve measurements. It was found that at turnover, the electrode temperature immediately above the device was uniformly $\sim 36^\circ$C, i.e. $\sim 17^\circ$C above ambient, as predicted by the thermal model. This implies that the electrodes do heat up, and that the device temperature (up to turnover) is uniform. Beyond turnover, in the CCNDR region, the $36-40^\circ$C isotherms shifted outward from the device, and were nearly circular, implying that substrate conduction dominates, as suggested earlier.

8. Slow-ramp switching in Ge$_{15}$Te$_{81}$S$_2$Sb$_2$

Devices could be made to switch in the course of slow-ramp ($\geq 10$ s sweep) $I$--$V$ measurements. On glass substrates this occurred well into the CCNDR region, as shown in fig. 3, while on silicon substrates it occurred at, or close to, turnover. For
devices on glass substrates, switching occurred at almost constant current (large series $R$), to the ON-state curve, and the point at which this occurred ($V_{SW}, I_{SW}$) was reproducible from device to device, and in the same device after "resetting" ($I_{reset} = 100$ mA $\mu$s). It should be pointed out that the whole $I-V$ characteristic could be reproduced, including $V_i$ and $V_{SW}$ (see fig. 3), through many such operations.

It is evident that such a switching event is not a simple device characteristic such as load-line instability, but may depend implicitly on material parameters. There are a number of possibilities, e.g.

(1) Uniform heating to a critical temperature, such as $T_g$ (395 K for Ge$_{15}$Te$_{81}$S$_2$-Sb$_2$) above which structural rearrangement may occur.

(2) Sudden contraction of the current distribution to filamentary form.

(3) A combination of (1) and (2).

The behaviour of devices on silicon is significantly different with switching occurring near turnover. The power density at switching is independent of device area (implied by fig. 13), i.e. $\sim 10^3$ W cm$^{-2}$, and the average current density is roughly the same as in devices on glass substrates at switching. This situation approximates more closely to the "well-heat-sunk" device, in which CCNDR should not occur, with uniform device current. It is possible that the switching at turnover and the apparent inaccessibility of a CCNDR region is due to sudden channelling of the current distribution. An examination of the form of the thermal conductance reveals how this might occur in devices on silicon, but not in devices on glass substrates.

\[ \Gamma_{\text{per unit area}} \text{on Si} = (2K_i/d_f), \]
\[ \Gamma_{\text{per unit area}} \text{on glass} = (2K_g/a_t) \]

If $a$ is the effective radius of a hot area (rather than the whole device), then for glass substrates the cooling efficiency (per unit area) increases as $a$ shrinks, but does not change for silicon substrates. This would make the current distribution on glass substrates less prone to a channelling instability.

9. Pulsed characteristics

The static and slow-ramp measurements give information which allows assessment of the thermal behaviour of devices under pulsed conditions. Fig. 15 shows a simple thermal equivalent circuit of a device from which a number of thermal time constants may be derived, each relevant to a particular regime of operation. Values for a 100 $\mu$m diameter pore on a glass substrate are as follows:

\[ \tau_{ff} = (c_i/\Gamma_i) = (c_i d_f^2/2K_f) = 1 \mu s, \]  
\[ \tau_{sf} = (c_s/\Gamma_s) = (c_s d_f^2 d_f/K_f a) = 60 \text{ ms}, \]  
\[ \tau_{ss} = (c_s/\Gamma_s) = (c_s d_f^2/K_s) = 2 \text{ s} \]
Fig. 15. Thermal equivalent circuit.

where $\tau_{ff}$ is the short-pulse device cooling time constant, i.e. before significant substrate heating, $\tau_{sf}$ the substrate heating time constant and $\tau_{ss}$ the substrate cooling time constant.

9.1. Pulsed threshold voltage $V_{tp}$ on glass substrates

For short pulses, before the glass substrate is significantly heated, the heat generated in the device is dissipated in the substrate via the device thermal conductance $\Gamma_f$. For $\tau_{ff} < t_{\text{pulse}} < \tau_{sf}$, the sample is in “quasi-thermal equilibrium”, and $V_{tp}$ may be obtained from the relevant steady-state turnover equation, but using $\Gamma_f$ in place of $\Gamma_s$. For virgin devices, this will result in a pulse threshold voltage $V_{tp}$ greater than the dc turnover voltage $V_t$, and, in fact, equal to the dc turnover voltage $V_t$ for similar devices on silicon substrates. In addition, devices which display simple thermal turnover (regime 1) at room temperature under “dc” conditions, will now be operating in an electrothermal regime.

In fig. 9, the dashed line shows the pulsed threshold voltage versus film thickness for the same range of thickness as the dc tests, at room temperature (pulse width 10 $\mu$s). The two points to note are (1) $V_t \propto d_f$, while $V_t \propto d^{1/2}$ and (2) for a 1 $\mu$m film, $V_{tp} = 19$ V $\approx V_t$ (Si substrate).

After a number of threshold operations, “forming” occurred, and $V_{tp}$ fell to a lower value, often close to $V_t$. By applying a suitable ‘reset’ pulse (150 mA, 10 $\mu$s), the original $V_{tp}$ could almost be recovered for the next pulse.

9.2. Switching delay time $t_D$ and forming

For very short voltage pulses, such that $t_p < \tau_{ff}$, all of the input energy raises the device temperature. Eq. (1) may be integrated to find, say, the time taken to raise the device temperature to some critical value $T_c$. Ignoring the thermal conductance, then,

$$\frac{c_f d_f^2}{V^2 \sigma_0} \int_{T_a}^{T_c} \exp\left[\frac{(\Delta E - ea(T)\varepsilon)}{kT}\right] dT = t. \tag{22}$$

Assuming that above room temperature, $a(T)$ [in eq. (22)] tends to a constant value


\[ t_D = \frac{c_t d_f}{V^2 a_0} \left\{ T_a^2 \exp\left(\frac{\Delta E - ea'e}{kT}\right) \right. - T_c^2 \exp\left(\frac{\Delta E - ea'e}{kT}\right) \right\}. \]

(23)

The second term may be ignored if \((T_c - T_a) > 50^\circ C\), giving

\[ t_D = \frac{c_t d_f^2 kT_a^2}{\sigma_0 \Delta E} \frac{1}{V^2} \exp\left(\frac{ea'V}{d_f kT_a}\right). \]

(24)

This is plotted as a function of \(V\), in fig. 16, for 0.1 \(\mu s < t_D < 10 \mu s\), and for a 100 \(\mu m\) diameter pore device \((d_f = 1 \mu m)\), although for such a device, it is not strictly valid when \(t_D > 1 \mu s\).

It was found experimentally that during “running-in” \(V_{tp}\) would drop by a factor of 2–3 below its virgin value, while \(t_D\), for a given voltage dropped considerably more (up to two orders of magnitude). Also shown in fig. 16, are the measured values of \(t_d\) versus \(V\), in several “time-regions” of the ‘forming’ sequence.

The curve closest to the computed curve was obtained during the first few operations, where some “forming” must have already occurred, since the computed \(t_D\) is an underestimate for \(t_D > 1 \mu s\). The lower curves were obtained after a number of threshold operations (~100, and ~10³). It can be seen that while \(V_{tp}\) drops from 20 to 8 \(V\), \(t_D\) (at 25 \(V\)) drops from 4 to 0.15 \(\mu s\), a considerable shift.

A简单 explanation for the extreme sensitivity of \(t_D\) to forming is based on the electrothermal model, and Thomas and Bosnell’s [20] suggestions regarding forming in threshold devices of STAG glass. Forming in STAG devices is thought to involve the precipitation of small crystallites of Te from the amorphous matrix. The

![Fig. 16. The variation of delay time \(t_D\) with pulse voltage for a pore device of radius \(a = 50 \mu m\) and thickness \(d_f = 1 \mu m\) on a glass substrate. The solid line on the right was calculated from eq. (24) and the arrow indicates how the experimental curves (○, △ and ●) shift to the left as forming proceeds.](image-url)
data for Ge$_{15}$Te$_{83}$S$_2$Sb$_2$ [21] suggest that Te is also the major crystallization product (Ge—Te crystallization being suppressed), and since the material is near the edge of a glass-forming region this forming mechanism is probably even more appropriate. It is assumed that the result is to reduce the effective “electrical thickness” of the film, $d_{fe}$, in the region affected (which may be a relatively narrow channel). It is also probable that the effective film thickness (say, $d_{nt}$) for thermal considerations, e.g. heat capacity, is not changed. If pulsed switching is in the electrothermal regime, then, from eq. (10),

$$V_{tp} \text{ (formed)} \propto d_{fe}.$$ 

While for the delay time, from equation (24),

$$t_{D} \text{ (formed)} \propto d_{fe} \exp(-ea'V/kT_a d_{fe}). \quad (25)$$

This assumes that the resistivity of the amorphous matrix is not much changed on forming and that the field is uniform in the matrix. It can be seen that the presence of the field exponential in eq. (25) can lead to a large reduction of $t_D$ during forming. It should also be noted that eq. (24) is independent of pore radius and therefore the argument should also apply to “forming” within a small channel. Using the data given above, a reduction of $V_{tp}$ from 20 to 8 V (factor 0.3) should be accompanied by a reduction in $t_D$ (for a 25 V pulse), by a factor $\sim0.03-\sim0.12 \mu$s, in reasonable agreement with the measurements. (Note that since the computed “reduced” $t_D < 1 \mu$s, the value obtained should be valid.)

10. Conclusions

The threshold characteristics of chalcogenide glass memory switches under typical operating conditions (quasi-static or pulsed) can be quantitatively or, at least, semiquantitatively described by an electrothermal model. It is necessary to take into account the influence of thermal fringing on the effective thermal conductance of the substrate, and the full field-dependence of the isothermal conductivity of the chalcogenide glass. The latter is crucial and there is no doubt that the field dependence arises from electronic mechanisms which as yet are not properly understood. However, from the viewpoint of the initiation of switching (i.e. the threshold conditions), the electronic mechanism is immaterial; the field dependence simply serves to deliver energy to the device more rapidly than would otherwise happen.

It must be emphasized that these conclusions do not necessarily apply to the threshold ON-state of these devices. Here, there may be significant electronic factors, although it is obvious from the way in which the permanent ON-state is formed ( precipitation of a crystalline filament) that appreciable heating is involved during the memory setting pulse. Even so, in some of our experiments on the esta-
blishment of the permanent ON-state there are indications that non-thermal factors may also play a role, but this will be reported elsewhere.

Acknowledgements

Most of the material in this paper is based on reports made to the UK Science Research Council (SRC) over the period October 1972 to May 1974 and we are grateful to the SRC for research grants which made the work possible. The authors also thank Messrs. D. Reynolds and G. Borzucki for their technical assistance in preparing the materials and devices.

The first author (A.E. Owen) is very grateful for the hospitality extended by D. Esteve, J. Buxo and their colleagues of the Laboratoire d'Automatique et d'Analyse des Systèmes (Toulouse, France) while he was on sabbatical leave in their laboratory (September 1977—September 1978) and during which period this paper was written.

Finally, but most importantly, we acknowledge with sincere thanks our debt to Prof. Sir Nevill Mott. Our group at Edinburgh was one of the first in the UK to concern itself with amorphous semiconductors; its research has been concerned with chalcogenides, transition-metal oxide glasses and other oxides, and it has covered various aspects of the subject from fundamental optical and transport properties to applications in switching devices, etc. Over many years we have been fortunate to have the benefit of innumerable informal discussions and much correspondence with Sir Nevill; whatever the topic he has never failed to respond with helpful comments and criticisms — always perceptive and thought-provoking, but always impartial to his own point of view. He has, we know, given the same consideration to many other groups in the UK and elsewhere. The debt which the subject owes to Sir Nevill is not simply for his own incomparable contribution, but also for the energy and enthusiasm with which he has helped and encouraged groups like our own.

References

[19] Handbook of Chemistry and Physics (Chemical Rubber Co., Cleveland, Ohio, 1974).
New amorphous-silicon electrically programmable nonvolatile switching device

Prof. A.E. Owen, P.G. Le Comber, G. Sarrebarroyse and W.E. Spear

Indexing terms: Switching, Semiconductor devices and materials

Abstract: The paper reports preliminary data on the characteristics of a new electronic switching device based on amorphous silicon structures. The device is polar, and is switched from OFF to ON (WRITE) or ON to OFF (ERASE) by voltages of opposite sign, the threshold voltage for WRITE being 4-8 V, depending on the device, and for ERASE it is about 1 V. The OFF and ON resistances for WRITE are about 10 MΩ and in the range 50-300 Ω, respectively. Experimental devices have been switched through 10⁴ WRITE and ERASE cycles, and the prospects are that this device can be permanently modified by this forming procedure.

The paper reports first results on a novel amorphous silicon (a-Si) memory device fabricated from the device reported here indicate that the new a-Si memory is potentially for nonvolatile, programmable storage [2].

In this paper we report the first results on a novel electrically programmable nonvolatile semiconductor memory device (MNOS) or floating gate avalanche-metal-oxide-semiconductor (FAMOS) devices currently used in semiconductor memories most showed the possibility of electrically programmable different junction configurations have been investigated, and the total thickness lay between 1 mm in diameter) was deposited onto the surface of the device performance. After completion of the a-Si deposition device as an electrically programmable and nonvolatile memory.

For most of the configurations investigated it was found that the threshold voltages are only slightly dependent on temperature.

All of the experiments were carried out on a-Si junctions deposited by the glow-discharge technique and doped from the gas phase [3]. Stainless-steel substrates were used, and the probe or a thin wire attached to the metal dots with silver paste.

1 Introduction and device fabrication

In this paper we report the first results on a novel electrically programmable nonvolatile semiconductor memory device fabricated from amorphous silicon (a-Si) [1]. The observations reported here indicate that the new a-Si memory is potentially superior, in terms of speed, retention time, operating voltages, stability etc., to the metal-nitride-oxide-semiconductor (MNOS) or floating-gate-avalanche-metal-oxide-semiconductor (FAMOS) devices currently used in semiconductor memories for nonvolatile, programmable storage [2].

All of the experiments were carried out on a-Si junctions deposited by the glow-discharge technique and doped from the gas phase [3]. Stainless-steel substrates were used, and the total thickness lay between 0.5 and 1.0 μm. A number of different junction configurations have been investigated, and most showed the possibility of electrically programmable switching to varying degrees. In its simplest form the device consists of a p- and an n-layer deposited onto conducting stainless-steel substrates. It was generally found, however, that a third undoped quasi-intrinsic (0) layer tended to stabilise device performance. After completion of the a-Si deposition a series of gold (Au) or aluminium (Al) dots (approximately 1 mm in diameter) was deposited onto the surface of the specimens, and the top contact was completed either by a probe or a thin wire attached to the metal dots with silver paste.

2 Forming and static characteristics

For most of the configurations investigated it was found that the first switching operation is unique; all following cycles occur, reproducibly, at a considerably lower threshold voltage. This first operation is therefore called 'forming', and it seems to be an essential precursor for subsequent operation of the device as an electrically programmable and nonvolatile memory. Fig. 1 illustrates the forming process for a typical p-n-i structure. Curve a in Fig. 1A is the initial static I/V characteristic in the conventional forward direction (i.e. the electrode adjacent to the p-layer is made positive with respect to the other electrode).

Curve b represents the initial reverse characteristic. When the applied forward potential is increased to values between 20 and 25 V, a rapid rise in current takes place and the device is brought into a highly conducting formed state, represented by curve c in Fig. 1B. The resistance is now of the order of 50 Ω and the electrical properties of the specimen have been permanently modified by this forming procedure.

The device now functions as a nonvolatile memory device, and Fig. 2 shows a typical set of static (DC) characteristics measured point-by-point. Immediately after forming, the device is in its ON state so that small positive and negative voltages trace out curve ab; ON-state currents of 10 mA or more are generally observed. On increasing the reverse potential (i.e. a negative voltage applied to the p-doped region) a reverse threshold voltage V_{thR} is reached, beyond which the device switches to an OFF state with a resistance of the order of 1 MΩ, represented in Fig. 2 by the characteristic cd. The reverse threshold voltage V_{thR} is about 1 V for the particular batch of specimens described. The OFF state is stable for voltage swings of ± 4 V. If now the forward potential is increased beyond a value V_{thF}, the forward threshold voltage, the device switches back into its high-conductivity state ab, in some cases through an intermediate state such as e or f.

The above cycle has been repeated up to 10⁵ times without observable changes in characteristics or threshold voltages. Devices set in ON and OFF states have been monitored for at least several weeks, and there has been no detectable change in any characteristic. With the present design the threshold voltages for stable switching to the ON or OFF state are of opposite polarity, although in a few devices some form of switching could be observed with both positive and negative voltages. The actual values of V_{thF} and V_{thR} depend on a number of variables, including the thicknesses of the various layers, which need further investigation.

As noted above, the OFF-state resistance is of the order of 1 MΩ in the present devices, which have effective areas of about 0.01 cm² (defined by the top evaporated contacts of 1 or 2 mm diameter). The ON-state resistance is determined by the current allowed to flow through the device during the switching transient (and so by any series resistance). This can become as low as 10 Ω, but it is preferable to prevent the ON-state resistance falling below about 200 Ω to achieve reproducible cycling of the device.

Preliminary experiments on the effect of temperature on V_{thF} and V_{thR} indicate only small changes in either threshold voltage with temperature (e.g. in a particular experiment, V_{thR} varied from ~ 4 V at room temperature to ~ 3 V at 450 K). The threshold voltages are also unaffected by illumination from filament lamps.

3 Dynamic characteristics

The dynamic responses of a typical p-n-i a-Si switch through the OFF→ON and ON→OFF transitions are illustrated in Figs. 3a and b, respectively. In both cases a 10 V pulse of 100 ns duration superimposed on a 1 V pulse of approximately 0.5 μs

---

Paper 18331, first received 18th January and in revised form 8th February 1982

Prof. Owen is with the Department of Electrical Engineering, University of Edinburgh, King's Buildings, Edinburgh EH9 3JL, Scotland. Dr. Le Comber and Dr. Spear are with the Carnegie Laboratory of Physics, University of Dundee, Dundee DD1 4HN, Scotland. Dr. Sarrebarroyse is with the Laboratoire d'Automatique et d'Analyse des Systemes du CNRS, 7 Avenue du Colonel Roche, 31400 Toulouse, France.
duration is applied to the device and, as before, the positive polarity (Fig. 3a) means that the p-layer of the p-n-i structure is positively biased. The purpose of the 1 V pulse is to 'read' the state of the switch; it is not a holding pulse to maintain the ON state, and separate experiments clearly established that the 10 V, 100 ns pulse switches the device into a nonvolatile conducting state. This latter point is also evident from Fig. 2 which shows that the ON-state characteristic passes through the origin. The characteristics shown in Fig. 3a and b were taken from switching transients recorded on a fast storage oscilloscope, and they demonstrate that either transition is completed within 100 ns.

A particularly notable feature is that, for both switching transients, the device current responds essentially instantaneously to the voltage signal, i.e. on the time-scale of these experiments, at least, there is no delay time involved in the response of the switch. During the switching pulse the current through the device may reach 100 mA or more, and it rises and falls too rapidly to be clearly recorded in the present experiments. The relatively slow current decay on the trailing edges of the current transients is determined by the time constants of the measuring circuit. It is estimated that the energy absorbed during either transition is extremely low, typically in the range $10^{-6}$ to $10^{-8}$ J.

**Fig. 2** Complete static current/voltage characteristics for formed p-n-i-a-Si device, showing forward and reverse threshold voltage, $V_{ThF}$ and $V_{ThR}$, respectively.

Preliminary experiments have been carried out on the effect of varying the height and duration of the switching pulse. As far as the pulse height is concerned the important factor is probably the overvoltage, i.e. the amount by which the pulse voltage exceeds the static threshold voltage, defined in Fig. 2. For the devices reported here $V_{Th}$ (see Fig. 2) is typically 4 to 6 V, and it has been observed that if the height of the switching pulse is reduced to 8 V, the pulse duration must be increased to about 400 ns to ensure that the device is switched into a permanent ON state. It is not clear at the moment, however, whether this effect is the direct result of the lower pulse height, or a consequential reduction in the current during the switching transient.

### 4 Discussion

The mechanisms underlying the switching phenomena described above are unknown at present, and it does not seem profitable to speculate about their possible nature on the basis of the above preliminary observations. More detailed experimental results are required to establish the critical features and parameters of the device, e.g. the relationship between switching performance and the thickness of the various layers. It is worthwhile, however, to draw some comparisons with other related switching devices.

First, it is relevant to note that, although threshold switching has been observed in a-Si [4], this is the first report of memory switching in an a-Si device. Also, threshold switching is well known in related structures fabricated on monocrystalline silicon, i.e. in the so-called MISS device which has the structure $p^+n^+i$ or $n^+p^+i$ [5-8]. The $i$-layer in these devices is usually an SiO$_2$ film thin enough to pass significant tunnelling currents (i.e. $< 40 \text{ A}$) but it may be significant, in the present context, that MISS devices can also be fabricated in 'all-Si' form, using polycrystalline Si as the $i$-layer [8b]. MISS devices switch to a nonpermanent ON state when the $p^+n$ (or $n^+p$) junction is forward biased, i.e. the same polarity that causes memory switching in the a-Si structures reported here. In the MISS
device the switching action is associated with minority-carrier injection and accumulation at the interface of the i-layer, normally leading to punch-through to the injecting contact which causes the device to switch ON. It is conceivable that in the a-Si structure the initiation of nonvolatile memory switching involves similar processes, but it must be noted that the OFF-ON transition time, required to set the a-Si device into a permanent ON state, is at least an order of magnitude faster than the delay time observed in MISS switches for threshold switching. Moreover, as mentioned earlier, on the time-scale used in the present experiments there is no observable delay time in the switching action of the a-Si device, whereas for the MISS device well-defined voltage-dependent delay times of at least 100 ns are commonly observed.

In the field of amorphous semiconductors much attention has been given over the past 10–15 years to memory switching devices fabricated from multicomponent chalcogenide glasses in which the reversible memory action is associated with the growth and destruction of a crystalline filament [9–12]. The switching phenomenon in the a-Si structures reported here is clearly very different, at least operationally. The most obvious difference is the completely nonpolar character of switching in chalcogenide glass devices, in contrast to the marked polarity-dependence of the a-Si memory switches. More importantly, perhaps, the switching times for the a-Si device are much faster (100 ns or less for either the OFF-ON or ON-OFF transition, compared with at least several milliseconds in chalcogenide devices) and the energy involved in the switching process is considerably lower (1 µJ or less compared with 1 mJ or more). Also, chalcogenide glass devices require voltage pulses of magnitude 25–30 V (for a device ~ 1 µm thick) to establish the ON state, and very often they need 100 or more ‘forming’ cycles before reasonably stable operation is achieved. This again contrasts with the operation of the a-Si memory switch, which (for a total device thickness of ~1 µm) there is a single forming step with a threshold voltage of about 30 V, and for all subsequent operations the forward threshold (V_{TPF}) is 4 to 6 V.

The closest parallel to the a-Si devices described in this paper seems to be the observation of memory switching in heterojunctions of n-type ZnSe grown epitaxially on p-type (single-crystal) Ge substrates, reported by Hovel and Urgell [13]. The ZnSe-Ge heterojunction devices are polar and the transition times for the OFF-ON and ON-OFF operation are both in the region of 100 ns or less. Similar but not so well-substantiated memory switching characteristics have also been briefly reported in devices fabricated by forming Schottky contacts on n-type GaAs and Si (single crystal) [14]. Hovel and Urgell have tentatively and qualitatively explained switching in the heterojunction by a model involving the filling and emptying of traps in the ZnSe, with the formation of a current filament in the ON state. Even in this case of superficially similar characteristics however, there are notable differences. Most significantly, the polarity required for switching in the ZnSe-Ge heterojunctions is the opposite to that found in the a-Si devices and the OFF-ON threshold voltage for the ZnSe-Ge switch decreases substantially with temperature (from about 1 V at 200 K to less than 0.1 V at 400 K), whereas V_{TPF} for the a-Si devices investigated so far is at the most only weakly temperature dependent.

The MNOS and FAMOS devices which, as mentioned in Section 1, are currently used in the semiconductor industry for nonvolatile memory applications, are both basically field-effect transistors which rely on charge injection and storage in the gate insulator(s) for their nonvolatility [15, 16]. They are therefore inherently slow, the WRITE and ERASE voltages are relatively large (e.g. 20–30 V) and they degrade significantly after cycling through 10⁴ to 10⁶ WRITE/ERASE (ON/OFF) operations.

Finally, it should be noted that in all switching devices, except those of the field-effect type mentioned latterly, it is almost certain that the ON state, whether permanent (nonvolatile) or temporary (as in a threshold switch), involves the formation of a current filament. The present experimental evidence, although indirect, suggests that a conducting filament is also formed in setting the a-Si switch into its ON state. The extent of the filament, the conductivity implied in the formation of a filament, and the consequences of filament formation, are all important factors in determining the performance of a-Si devices. Further study of these factors is needed in order to determine the role of filament formation in the switching processes.

![Fig. 3 Dynamic switching characteristics of p-n-i a-Si device, using 10 V, 100 ns pulse superimposed on 1 V, ~ 0.5 µs pulse (drawn from oscilloscope traces)](image)

IEE PROC., Vol. 129, Pt. 1, No. 2, APRIL 1982
of semimetals. It could perhaps be relevant that conductivities approaching $10^4$ S m$^{-1}$ have recently been observed in doped microcrystalline films of Si [17].

5 Acknowledgments

The authors thank J.J. Delima (of the Department of Electrical Engineering, University of Edinburgh) and S. Kinmond (of the Carnegie Laboratory of Physics, University of Dundee) for valuable assistance with the experimental work.

6 References

MEMORY SWITCHING IN AMORPHOUS SILICON DEVICES

A.E. OWEH*, P.G. LE COMBER, W.E. SPEAR and J. HAJTO

* Department of Electrical Engineering, University of Edinburgh, King's Buildings, Edinburgh EH9 3JL, Scotland, U.K.
+ Carnegie Laboratory of Physics, University of Dundee, Dundee DD1 4HN, Scotland, U.K.

Recent experimental observations on high-speed memory switching in p-n-i structures of amorphous silicon are described. Particular emphasis is given to the first switching operation of a virgin device which is effectively a unique forcing process for all subsequent operations. There is a characteristic delay for forming, varying over ten orders of magnitude from \(-10^4\) to \(-10^8\) s. Evidence is presented to show that forming is a charge-controlled process which occurs at a constant field across the n-layer, but details of the switching mechanisms in the a-Si p-n-i devices remain obscure.

1. INTRODUCTION

Widespread interest in the phenomena of threshold (monostable or volatile) and memory ( bistable or non-volatile) electrical switching in chalcogenide glasses provided the impetus for much of the pioneering research in the field of amorphous semiconductors in the late 1960s and early 1970s.1,2 There are still some controversial features but by and large generally acceptable models for both threshold and memory switching in chalcogenide glasses are now reasonably well established.3,4

By contrast, and despite the almost unprecedented growth of research and development on amorphous silicon (a-Si) since the mid-1970s, practically nothing has been reported on electrical switching in that material. Some cursory observations of threshold switching in homogeneous films of evaporated a-Si, with very tentative evidence for memory behaviour, were described by Feldman and Moorjani contemporaneously with some of the early literature on switching in chalcogenide glasses.5 There seems to have been no further activity until the work of Dey and Fong who observed threshold switching in homogeneous films of evaporated a-Si with titanium contacts.6,7 Much more recently, den Boer has also reported threshold switching in a-Si, but specifically in hydrogenated n-i-n' structures (i stands for “intrinsic”).8 To the authors’ knowledge, the only other relevant paper, again on hydrogenated a-Si, is the slightly prior publication of Gabriel and Adler9 who searched unsuccessfully for switching in homogeneous films with molybdenum contacts. They concluded that unlike the chalcogenide glasses, hydrogenated
a-Si does not have the fundamentally requisite properties for reversible switching behaviour.

The present paper is concerned with studies of the electrical switching behaviour of heterogeneous structures of a-Si which evolved from a collaborative project between the amorphous semiconductor groups at the Universities of Dundee and Edinburgh in the U.K. A preliminary description of this work has already been published\(^{10}\) and in the following sections a more detailed account of experimental observations will be presented with emphasis on the initial switching operation of freshly prepared (and previously unswitched) devices. It is believed, on the basis of the present evidence, that the first operation is a charge controlled process. The research is still in its early stages however and the switching mechanisms remain obscure.

2. DEVICE STRUCTURE AND FABRICATION

A variety of configurations has been investigated but the majority of the observations to date, and all of the results reported in this paper, have been obtained on devices having a-Si doped and intrinsic (i) layers in the sequence p\(^+\)-n-i. The a-Si was deposited by the glow discharge technique, with gas phase doping, as described by Spear.\(^{11}\) Stainless steel substrates were generally used and the total thickness of the deposited a-Si layers was between 0.5 and 1.0 \(\mu\)m. After completion of the a-Si deposition a series of gold (Au), aluminium (Al) or nichrome (NiCr) dots (up to approximately 1 mm in diameter) was evaporated onto the surface of the samples and the top contact was completed either by a probe or by a thin-wire attached to the metal dots with conducting silver paste.

3. STATIC CURRENT-VOLTAGE CHARACTERISTICS OF VIRGIN DEVICES

Typical current-voltage (I-V) characteristics for a freshly prepared (unswitched) device are illustrated in figure 1(a) in both the forward and reverse directions (the forward direction is defined such that the substrate, and hence the p\(^+\)-region, is positively biased). It must be emphasised here that these measurements were taken "by hand", point-by-point, in a manner which required a few seconds for each point to be measured. The significance of this remark will become apparent in the next section. In the forward direction there is a region of ohmic behaviour over an appreciable voltage followed by an abrupt change to a markedly non-ohmic region until, at the point indicated by the arrow, the device was unstable and it was impossible to continue with point-by-point measurements. The change from ohmic to non-ohmic behaviour is more clearly apparent in the conductivity vs. field plot of
A. L O. c., at / Mernun- sw:wI, ,i,,g in ,jmmpIs,ius nhcun devices

FIGURE 1(a)
Static current-voltage characteristics of a virgin (unformed) device. The positive quadrant corresponds to the $p^+$-layer positively biased.

FIGURE 1(b)
The data for 30°C in figure 1(a) plotted in the form log.conductivity vs. field, and illustrating the change from ohmic to non-ohmic behaviour.

of figure 1(b). As ambient temperature increases the onset of non-ohmic behaviour moves to lower voltages. In the reverse direction there is also an initial ohmic region (and this part of the I-V curve is symmetrical) but the change to non-ohmic behaviour is much more gradual and eventually the device breaks down.

4. FORMING - STATIC CHARACTERISTICS

As noted in the previous section, during point-by-point measurements under forward bias, at room temperature, the a-Si $p^+$-n-i device tends to become unstable when the applied bias is about 25 V. At higher temperatures the instability occurs at lower voltages, as indicated by the arrows in figure 1. On attempting to increase the voltage still further the device switches into a low resistance ON-state. Typical I-V characteristics for both polarities in the ON-state are shown in figure 2; the I-V curve is ohmic, it extrapolates through the origin (i.e. the ON-state is permanent) and it is slightly asymmetrical. Note that the current is now measured in milliamps and the voltages across the device are small. On increasing the voltage in the forward direction the ON-state current continues to increase apparently indefinitely, subject only to any current limiting resistor, and the device is eventually destroyed, presumably by Joule heating. In the reverse
FIGURE 2  
Current-voltage characteristics of a device in the ON-state, at different temperatures, in the forward (FWD) and reverse (REV) directions. The forward direction corresponds to the p+-layer positively biased.

FIGURE 3  
Complete static current-voltage characteristics of a formed a-Si p'-n-i device, showing the forward and reverse threshold voltages, $V_{\text{ThF}}$ and $V_{\text{ThR}}$ respectively.

direction however another instability is observed and at about -1 V (typically) the device switches back into a high resistance OFF-state. The OFF-ON transition may now be repeated by biasing in the forward direction but on the second and all subsequent switching operations the forward threshold voltage $V_{\text{ThF}}$ occurs at a much lower voltage than the first operation e.g. at -5 V compared with the 25 V observed under the conditions obtaining for the measurements shown in figure 1. The first OFF-ON transition, occurring at a relatively high voltage, seems therefore to be unique and by analogy with the usage of switching in chalcogenide glasses, it is referred to as "forming".

The formed a-Si p'-n-i device may be cycled through ON and OFF states by a sequence of biasing in forward and reverse directions with critical points at $V_{\text{ThF}}$ and, in the reverse direction, $V_{\text{ThR}}$. A complete and typical characteristic obtained on a curve tracer is illustrated in figure 3. On occasions the device appears to go through a number of intermediate ON-states during the OFF-ON transition and this is indicated in figure 3. In addition, there is often an observable and appreciable region of negative resistance in the reverse biased OFF-state characteristic of a formed device (also indicated in figure 3).
The forming process, i.e. the first OFF-ON transition, does not occur instantaneously when a voltage step or pulse is applied to the a-Si p'-n-i device. Initially there is a delay time, \( t_0 \), during which the device current remains essentially constant at the OFF-state value appropriate to the voltage across the device. Only after this delay does the current begin to increase and it then rises essentially instantaneously to its ON-state value. The forming delay time is an extremely sensitive function of the applied voltage and typical data, obtained at three temperatures, are illustrated in figure 4.

The forming delay time varies over nearly ten orders of magnitude from a few hundred seconds at low forming voltages to about 10 ns at high voltages. In particular, at a temperature dependent critical voltage \( V_{CR} \) there is virtually a discontinuous change in \( t_0 \) as a function of bias. The voltages \( V_{CR} \) indicated in figure 4 are roughly coincident with the voltages at the points of instability marked by the arrows in figure 1; \( V_{CR} \) also corresponds to the forming voltage which would be obtained in an experiment on a virgin device with a curve tracer. Note in figure 4 that above and below \( V_{CR} \) the delay time tends to a value which seems to be approximately independent of both voltage and temperature; for the particular results illustrated \( t_0 \) is in the range \( 10^2 \)-\( 10^3 \) s for \( V < V_{CR} \) and in the range 10-100 ns for \( V > V_{CR} \).
The results plotted in figure 4 for $V > V_{CR}$ correspond of course to voltages less than the point of instability indicated in figure 1. There does appear to be a lower limit to the forming voltage however and present results indicate that the limiting voltage coincides with the bias at which the I-V characteristics change from their ohmic to non-ohmic behaviour (see figure 1 and section 3). Several experiments have shown that virgin devices fail to switch (form) even if held for many hours at a forward bias only just below the non-ohmic region. In other words, forming occurs at any forward bias within the non-ohmic region of the I-V characteristics but at voltages below the point of instability $t_0$ is comparatively long. It must also be emphasised again that the device current remains constant at its OFF-state magnitude during the delay time, even when $t_0$ is 100 s or more.

Preliminary experiments have been carried out to determine the effect of device geometry on the forming voltage $V_F$. It was found that $V_F$ increases linearly with the thickness of the n-layer and data for typical p'-n-i devices are plotted in figure 5; $V_F$ in this case was measured by a curve tracer and as already mentioned therefore it coincides with the voltage $V_{CR}$ indicated in figure 4. Note that the results do not extrapolate to zero voltage for zero n-layer thickness.
6. DYNAMIC SWITCHING OF FORMED DEVICES

The principal experimental features of the pulsed operation of formed a-Si p'-n-i devices have already been described. A representative diagram drawn from an oscilloscope trace of the OFF-ON transition on applying a fast voltage ramp is shown in figure 6. The main points to note are as follows:

(i) When biased with a pulse in the forward direction the device switches ON provided the pulse height exceeds the static threshold voltage \( V_{ThF} \) as defined in figure 3.

(ii) On the time scale of \( 1 \text{ ns} \) or less, there is no observable delay time in the response of a formed device. The device current follows the applied voltage instantaneously on this timescale.

(iii) Provided the pulse is long enough the ON-state is permanent and the pulse duration required for switching to a memory state increases as the pulse height decreases towards \( V_{ThF} \). In typical cases a permanent ON-state is obtained with pulse durations of a few tens of nanoseconds and magnitude \( 5 \text{ V} \) in excess of \( V_{ThF} \).

(iv) Similarly, on biasing in the reverse direction with a pulse of height \( V_{ThR} \) the device switches from OFF-ON and again there is no observable delay in response.

(v) The ON-state appears to be truly permanent. No detectable changes have been observed in devices stored in their ON-state for a year or more.

7. DISCUSSION

The authors' earlier paper on the new a-Si p'-n-i device focusses mainly on the properties of the formed switch, particularly its dynamic response, and the operational characteristics were compared with those of other switching devices. The switching mechanisms in the new device remain obscure and much more detailed research needs to be done. At this stage therefore there is little to add to the problem of memory switching in formed devices but recent results do provide some insight into the forming process.

It should be noted first that the charge which flows through or into the device during the delay time (i.e. during forming) is an approximately constant function of pulse height for voltages \( V_{CR} \). In fact, as illustrated in figure 7, the ratio of charge to n-layer thickness \( (Q/d_n) \) seems to be roughly independent of both pulse height and the thickness of the n-layer (sample 403 in figure 7 has an n-layer thickness of 215 nm, sample 349 of 445 nm and 223 of 780 nm). By contrast both the current and the power vary markedly with the pulsing conditions.

These observations suggest very strongly that forming is an electronic process associated with a critical charge density in the n-layer. In terms
of electrons the critical charge is \( 1.2 \times 10^{19} \) electrons cm\(^{-3} \). Moreover, the results in figure 5 imply that forming occurs at a constant field across the n-layer. For the particular devices shown the field is \( 2 \times 10^5 \) V cm\(^{-1} \), corresponding to a geometric charge which is rather less than the critical charge. Typically, the ratio of the critical charge to geometric charge varies from about 30 in sample 403 \((d_n = 215 \) nm\) to about 100 in sample 223 \((d_n = 780 \) nm\). Finally, assuming a carrier mobility of \( 10^4 \) cm\(^2\) V\(^{-1}\) s\(^{-1}\), the transit times across the n-layers are in the range of 1-4 ns for the samples shown. This is not inconsistent with the delay times for \( V \cdot V_{CR} \) (see figure 4).

ACKNOWLEDGEMENTS

The authors are grateful to the Venture Research Unit of British Petroleum International Ltd. for a grant supporting the research described in this paper. We also thank Dr. F.F. Carasco, Mr. M.C. Holland (both of Dundee University) and Mr. W.K. Choi for assistance with the experimental work and for enthusiastic help in preparing this paper.

REFERENCES

Electronic Switching in Amorphous Silicon Junction Devices

P. G. LeComber  
Carnegie Laboratory of Physics  
University of Dundee  
Dundee, Scotland

A. E. Owen  
Department of Electrical Engineering  
University of Edinburgh  
Edinburgh, Scotland

W. E. Spear  
Carnegie Laboratory of Physics  
University of Dundee  
Dundee, Scotland

J. Hajto  
Department of Electrical Engineering  
University of Edinburgh  
Edinburgh, Scotland

W. K. Choi  
Department of Electrical Engineering  
University of Edinburgh  
Edinburgh, Scotland

I. Introduction

Electronic switches are solid-state devices that can be changed from a nonconducting OFF state to a conducting ON state by an appropriate electrical signal. The importance of such devices in the development of solid-state digital electronics has been enormous and is likely to remain so in the foreseeable future, especially with the increasing demand for memory elements.

275
The present article is concerned with two-terminal switching devices. Generally, these have one or the other of the two types of current-voltage (I-V) characteristics shown schematically in Fig. 1a,b. In Fig. 1a the device switches from its OFF to its ON state at a critical threshold voltage $V_{th}$, but if the ON state conditions fall below a critical holding point ($I_h$, $V_h$), the device reverts spontaneously to its OFF state. Devices of this kind are called threshold switches; they are nonpermanent, or “volatile,” and they always revert to the OFF state in the absence of an appropriate bias. In Fig. 1b there is again a critical switching voltage for the OFF to ON transition, but both ON- and OFF-state characteristics extrapolate through the $I-V$ origin. Devices of this kind are therefore permanent, or nonvolatile, and they are called memory switches. Memory devices can remain in either the ON or OFF state more or less indefinitely, whether or not a bias is applied and the ON to OFF transition is usually triggered by a current pulse.

By the early 1970s many examples of threshold and memory switching had been reported in homogeneous thin films of a variety of amorphous materials, including simple oxides, transition-metal oxides, elemental selenium, and boron. By far the most important materials, however, were the chalcogenide glasses in which, depending on composition, reproducible characteristics of the kind illustrated in Fig. 1a (threshold switching) or Fig. 1b (memory switching) may be obtained (Ovshinsky, 1968; Owen and Robertson, 1973). A substantial specialist literature on electrical switching

![Fig. 1. Current-voltage characteristics for (a) threshold switching and (b) memory switching.](image-url)
in the chalcogenide glasses has developed, and although there are still some controversial features, generally accepted models of at least a semiquantitative kind are now reasonably well established for both types of switching (Adler et al., 1978; Owen et al., 1979).

This article is concerned with recent studies of a rather different switching behavior in heterogeneous structures of amorphous hydrogenated silicon, which evolved from a collaborative project between the authors’ groups at the Universities of Dundee and Edinburgh in the United Kingdom. A preliminary account of the work has already been published (Owen et al., 1982), and in the following sections a more detailed description of the experimental observations will be presented.

Before proceeding, however, it is worth noting that, by contrast with the chalcogenide glasses and despite the almost unprecedented growth in research and development on amorphous silicon (a-Si) since the mid-1970s, very little has been reported on electrical switching in the latter material. To put the recent observations in context it is pertinent therefore to review briefly the relatively few previous reports of switching in a-Si.

II. Previous Work on Electrical Switching in Amorphous Silicon

Some cursory observations of threshold switching in homogeneous films of a-Si, with very tentative evidence for memory behavior, were described by Feldman and Moorjani (1970) and Moorjani and Feldman (1970) contemporaneously with some of the early literature switching in chalcogenide glasses. More detailed experiments on the same structures were reported later (Feldman and Charles, 1974; Charles and Feldman, 1975). The authors studied vacuum-evaporated films of a-Si in the range 0.3–2.0 μm thick, fitted with titanium electrodes. They also, incidentally, made similar observations on evaporated films of germanium, boron, and boron plus carbon. As threshold switches these a-Si structures had threshold voltages $V_{th}$ of 5–10 V, OFF resistances in the range 1–30 kΩ, and an ON resistance of about 100 Ω. In common with the chalcogenide glasses, there was a delay time before switching of 20–50 μsec or more (at room temperature) and the actual switching time was at least several microseconds. Feldman et al. did not, however, report any initial “forming” process, unlike the situation in chalcogenide glass switches (see also Part V). As already noted, there was some tentative indication of memory switching, but that feature was apparently not substantiated. Feldman and Charles (1974) interpreted the results in terms of a simple and qualitative electrothermal model involving the formation of a conducting filament; a similar, more quantitative model has been developed for switching in chalcogenide glasses (Owen et al., 1979).

The work of Feldman and his colleagues, which originated in the early
1970s, seems to be the only investigation of switching in a-Si until the later studies of Dey and Fong (1977, 1979) and Dey (1980). These authors report results very similar to those of Feldman et al. They studied thin films of a-Si in the range 0.3 - 1.5 \( \mu m \) thick, deposited by electron-beam heating in a vacuum evaporator. Titanium contacts were again used, either in the form of evaporated films or as probes. Dey and Fong reported only threshold switching, with \( I-V \) characteristics similar to those in Fig. 1a; they did not mention any evidence for memory behavior. In contrast to Feldman et al., however, Dey and Fong did observe forming effects; that is, the initial threshold voltage was relatively large but it decreased to a more or less constant value after a number of switching cycles. In Dey and Fong's devices the threshold voltage varied systematically from about 6 \( V \) for the thinner films (=0.3 \( \mu m \)) to about 9 \( V \) for the thicker films (=1.2 \( \mu m \)). The delay time before switching was in the range 2-60 \( \mu sec \), varying in a systematic way with pulse height, pulse duration, and repetition rate, again in a manner very similar to switching in calcogenide glasses (Adler et al., 1978; Owen et al., 1979). Dey and Fong also interpreted their results in terms of a simple one-dimensional electrothermal model, but one developed a little more quantitatively than that by Feldman et al.

It should be noted that both Feldman et al. and Dey and Fong used a-Si films deposited by vacuum evaporation. This probably accounts for the relatively low OFF-state resistances which they both found (100 - 100 k\( \Omega \)). It is now well established that vacuum-evaporated a-Si is a very different material from the hydrogenated form of a-Si obtained, for example, by the carefully controlled glow-discharge decomposition of silane (e.g., Spear, 1977).

Three papers concerned specifically with switching in hydrogenated amorphous silicon by Gabriel and Adler (1982), den Boer (1982), and Owen et al. (1982) appeared almost concurrently early in 1982, each reporting very different effects observed in different a-Si structures. Our own work (Owen et al., 1982), including recent results, is described in detail in the following sections.

Den Boer studied \( n^+ - i - n^+ \) structures of a-Si prepared by the glow-discharge decomposition of \( \text{SiH}_4 \) ("i" stands for "intrinsic" or undoped material). The \( n^+ \) layers were 50 nm thick and prepared by adding 1\% PH\(_3\) to the \( \text{SiH}_4 \) gas flow; the \( i \) layer in different devices ranged in thickness from 2.5 to 5 \( \mu m \). Den Boer found that the \( n^+ - i - n^+ \) devices functioned as threshold switching devices with nonpolar characteristics similar to those sketched in Fig. 1a. For the first switching cycle the threshold voltage was in the range 40 - 100 \( V \), but for all subsequent operations it was only 10 - 35 \( V \), depending on the \( i \) layer thickness (as the \( i \) layer thickness increased the threshold voltage also increased). The OFF-state resistance of the \( n^+ - i - n^+ \) switches
was about 1 MΩ, and the ON-state resistance about 1 kΩ. There was again an observable delay time before switching, ranging from a few microseconds when the applied voltage was about 8 V greater than \( V_{th} \) to about a millisecond for voltages within 1 V of \( V_{th} \). The \( n^+ - i - n^+ \) threshold switches could be cycled through at least \( 10^5 \) stable switching operations. Den Boer also compared structures with chromium or a combination of chromium and \( n^+ \) contacts (i.e., \( Cr-n^+-i-Cr \) and \( Cr-i-Cr \)). The former had rectifying characteristics while the latter switched but were very unstable.

Gabriel and Adler (1982) prepared their a-Si films by sputtering from a polycrystalline silicon target in an argon—hydrogen plasma. In all cases their devices were notionally homogeneous thin films of intrinsic a-Si:H with molybdenum contacts. The samples were fabricated under a wide range of deposition conditions in two sputtering systems, and although results from some of the devices were rendered rather doubtful because of contamination problems, in no case did Gabriel and Adler observe any evidence of reversible switching. They concluded that, in contrast to the chalcogenide glasses, a-Si does not have the electronic and structural properties required for reversible switching.

III. Device Structure and Fabrication

We turn now to the work on electronic switching carried out in the authors' laboratories. Although a number of different a-Si multilayer structures have been investigated, all the results discussed in the following refer to \( p^+ - n - i \) devices deposited in this sequence by the glow-discharge technique, with gas-phase doping. Stainless steel substrates were generally used and the total thickness of the deposited a-Si layers was between 0.5 and 1.0 μm. After completion of the a-Si deposition a series of gold (Au), aluminum (Al), or nichrome (NiCr) dots (up to approximately 1 mm in diameter) was evaporated onto the surface of the samples, and the top contact was completed either by a probe or by a thin wire attached to the metal dots with conducting silver paste.

IV. Static Current—Voltage Characteristics of Virgin Devices

Typical current—voltage (\( I-V \)) characteristics for a freshly prepared (unswitched) device are illustrated in Fig. 2a in both the forward and reverse directions. (The forward direction is defined such that the substrate, and hence the \( p^+ \) region, is positively biased.) It must be emphasized here that these measurements were taken “by hand,” point by point, in a manner that required a few seconds for each point to be measured. (The significance of
this remark will become apparent in the next sections.) In the forward direction there is a region of ohmic behavior over a limited voltage range followed by an abrupt change to a markedly nonohmic region until, at the point indicated by the arrow, the device is unstable and it is impossible to continue with point-by-point measurements. The change from ohmic to nonohmic behavior is more clearly apparent in the effective conductivity versus applied voltage plot of Fig. 2b. As the ambient temperature increases, the onset of nonohmic behavior moves to lower voltages. In the reverse direction, corresponding to a negative potential applied to the $p^+$ side, there is an initial ohmic region which is symmetrical for positive and negative voltages. However, the change to nonohmic behavior is much more gradual in the reverse direction, leading to the eventual breakdown of the device.
V. Forming: Static Characteristics

As noted in the previous section, during point-by-point measurements under forward bias, the a-Si $p^+ - n - i$ device tends to become unstable when the applied bias is about 24 V at room temperature. At higher temperatures the instability occurs at lower voltages, as indicated by the arrows in Fig. 2a. On attempting to increase the voltage still further, the device switches into a low-resistance ON state. Typical $I - V$ characteristics for both polarities in the ON state are shown in Fig. 3; the $I - V$ curve is ohmic, it extrapolates through the origin (i.e., the ON state is permanent), and it is slightly asymmetrical. Note that the current is now measured in milliamperes and that the voltages across the device are small. On increasing the voltage in the forward direction the ON state current continues to increase apparently indefinitely, subject only to any current-limiting resistor, and the device is eventually destroyed, presumably by Joule heating. In the reverse direction, however, another instability is observed, and at about $-1$ V (typically) the

![Graph showing current-voltage characteristics](image)

**Fig. 3.** Current-voltage characteristics of a device in the ON state, at different temperatures, in the forward and reverse directions. The forward direction corresponds to $p^+$ layer positively biased. (●) forward bias, 80°C; (△) forward bias, 30°C; (○) reverse bias, 80°C; (△) reverse bias, 30°C.
device switches back into a high-resistance OFF state. The OFF-ON transition may now be repeated by biasing in the forward direction, but on the second and all subsequent switching operations the forward threshold voltage $V_{TF}$ occurs at a much lower voltage than the first operation, e.g., at $\approx 5 \text{ V}$ compared with the $25 \text{ V}$ observed under the conditions obtaining for the measurements shown in Fig. 2. The first OFF-ON transition, occurring at a relatively high voltage, seems therefore to be unique and by analogy with the usage of switching in chalcogenide glasses, it is referred to as "forming."

The formed a-Si $p^+-n-i$ device may be cycled through ON and OFF states by a sequence of biasing in forward and reverse directions with critical points at $V_{TF}$ and, in the reverse direction, $V_{TR}$. A complete and typical characteristic obtained on a curve tracer is illustrated in Fig. 4. On occasions the device appears to go through a number of intermediate ON states during the OFF-ON transition, and this is indicated in the figure. In addition, there is often an observable and appreciable region of negative resistance in the reverse-biased OFF state characteristic of a formed device, denoted by $N$ in Fig. 4.

VI. Forming: Dynamic Characteristics

The forming process (i.e., the first OFF-ON transition) does not occur instantaneously when a voltage step or pulse is applied to the a-Si $p^+-n-i$ device. Initially there is a delay time $t_D$ during which the device current remains essentially constant at the OFF-state value appropriate to the voltage across the device. Only after this delay does the current begin to increase, and it then rises almost instantaneously to its ON-state value. The forming delay time is an extremely sensitive function of the applied forming

![Figure 4](image_url)
voltages $V_F$, and typical data, obtained at three temperatures, are given in Fig. 5. The forming delay time varies over nearly 10 orders of magnitude, from a few hundred seconds at low forming voltages to about 10 nsec at high $V_F$. In particular, at a temperature-dependent critical forming voltage $V_{CR}$, there occurs virtually a discontinuous change in $t_D$. The forming voltages $V_{CR}$ indicated in Fig. 5 are approximately the same as the voltages at the points of instability marked by the arrows in Fig. 2a; $V_{CR}$ also corresponds to the forming voltage obtained in experiment with a curve tracer, operated in ac mode at a frequency of 50 Hz. It can also be seen in Fig. 5 that above and below $V_{CR}$ the delay time tends to a value that seems to be approximately independent of both voltage and temperature; for the particular results illustrated, $t_D$ is in the range $10^2$–$10^3$ sec for $V < V_{CR}$ and lies between 10 and 100 nsec for $V > V_{CR}$.

The results plotted in Fig. 5 for $V < V_{CR}$ correspond of course to voltages less than the point of instability indicated in Fig. 2a. There does appear to be a lower limit to the forming voltage, however, and present results indicate that the limiting voltage coincides with the bias at which the $I$–$V$ characteristics change from their ohmic to nonohmic behavior (see Fig. 2b). Several
experiments have shown that virgin devices fail to switch (form) even if held for many hours at a forward bias only just below the nonohmic region. In other words, forming occurs at any forward bias within the nonohmic region of the $I-V$ characteristics, but at voltages below the point of instability $t_D$ is comparatively long. It must also be emphasized again that the device current remains constant at its OFF-state magnitude during the delay time, even when $t_D$ is 100 sec or more.

Preliminary experiments have been carried out to determine the effects of device geometry on the forming voltage $V_F$. It was found that $V_F$ increases linearly with the thickness $d_n$ of the $n$ layer. Data for typical $p^+-n-i$ devices are plotted in Fig. 6. In this case $V_F$ was measured by applying a voltage ramp and its value coincides with the voltage $V_{CR}$ indicated in Fig. 5. It can be seen that the results do not extrapolate to zero voltage for zero $n$ layer thickness.

The charge $Q = \int_{t_D}^{t} i dt$, which flows through or into the device during the forming delay time, has been determined for $V_F > V_{CR}$. Figure 7 shows that in this range of $V_F$ the ratio $Q/d_n$ is approximately independent of $V_F$ and $d_n$ for $n$ layer thicknesses between 215 and 780 nm. This could mean that forming occurs when a critical volume charge has accumulated in the $n$ region.

VII. Dynamic Switching of Formed Devices

The principal experimental features of the pulsed operation of formed a-Si $p^+-n-i$ devices have already been described by Owen et al. (1982). A representative diagram drawn from an oscilloscope trace of the OFF $\rightarrow$ ON

![Diagram](image-url)

Fig. 6. The dependence of the forming voltage $V_F$ on the thickness of the $n$ layer.
transition on applying a fast voltage ramp is shown in Fig. 8. The main points to note are as follows:

(i) When biased with a pulse in the forward direction, the device switches ON, provided the pulse height exceeds the static threshold voltage $V_{TH}$ as defined in Fig. 4.

(ii) On the time scale of $\sim 1$ nsec or less, there is no observable delay time in the response of a formed device. The device current follows the applied voltage instantaneously on this time scale.

(iii) Provided the pulse is long enough, the ON state is permanent and the pulse duration required for switching to a memory state increases as the
P. G. LECOMBER et al.

pulse height decreases toward $V_{\text{TWF}}$. In typical cases a permanent ON state is obtained with pulse durations of a few tens of nanoseconds and magnitude $\sim 5 \text{ V in excess of } V_{\text{TWF}}$.

(iv) Similarly, on biasing in the reverse direction with a pulse of height $> V_{\text{TWF}}$ the device switches from ON $\rightarrow$ OFF and again there is no observable delay in response.

(v) The ON-state appears to be truly permanent. No detectable changes have been observed in devices stored in their ON state for a year or more.

VIII. Discussion of Possible Switching Mechanisms

The mechanisms underlying the switching phenomena in the a-Si devices are not understood at present; clearly, more data will be required to explain these exceptional properties. In the following we therefore draw some comparisons with other related switching devices and only briefly speculate on possible mechanisms.

Although there is no observation of memory switching in analogous crystalline Si (c-Si) devices, threshold switching is well known in c-Si $p^+ - n - i$, $n^+ - p - i$, and related structures (Yamamoto and Morimoto 1972, Yamamoto et al., 1976; Buxo et al., 1978; Sarrabayrouse et al., 1980; Simmons and El-Badry, 1978; Kroger and Wegener, 1973, 1975). The $i$-layer in these devices is usually a SiO$_2$ film ($\leq 40 \text{ Å}$), thin enough to pass appreciable tunneling currents, but it may be significant that these metal–insulator–semiconductor–semiconductor (MISS) devices can also be fabricated in an "all-Si" form using polycrystalline Si as the $i$-layer (Kroger and Wegener, 1975). These devices switch to a nonpermanent ON state when the $p^+ - n$ (or $n^+ - p$) junction is forward biased, which is the same polarity producing the memory ON state in the a-Si structures. In the MISS device the switching action is associated with minority-carrier injection from the $p - n$ junction and accumulation at the interface of the $i$ layer, normally leading to punchthrough to the injecting contact, which causes the device to switch ON. The values of the threshold voltage are similar to those observed in the a-Si forming process, and it is conceivable that for this operation the processes are similar. The experimental results described in Part VI, which suggest that the forming process is likely to be charge controlled, would not be inconsistent with this model.

For the c-Si MISS $p^+ - n - i$ threshold device the values of threshold voltage are predicted (Simmons and El-Badry, 1978) to depend on $(d_n - W)^2$, where $d_n$ is the thickness of the $n$ layer and $W$ is the width of the depletion region, and for low donor concentrations this dependence is supported experimentally. In contrast, the results in Fig. 6 show that in the a-Si devices the forming voltage varies linearly with $d_n$. It is difficult to decide at present
whether this disagreement suggests a different mechanism or whether it arises from our attempt to extrapolate from a model developed for a crystalline threshold device to an amorphous memory junction.

An alternative model, based on a regenerative process, has also been suggested for the c-Si MISS devices (Sarrabayrouse et al., 1980), taking into account a carrier multiplication mechanism at the "inverted" Si–SiO₂ interface. The model correctly accounts for a number of MISS properties; for example, it predicts a threshold independent of $d_n$, which agrees with the observations on $p^+–n–i$ MISS structures when the $n$ layer is more heavily doped. But this model is also in disagreement with the data for the a-Si devices shown in Fig. 6, and just as for the "punchthrough" model, it is therefore difficult at the present stage to decide whether the regenerative model could be relevant to the understanding of a-Si memory switching.

In the field of amorphous semiconductors much attention has been given over the past 10–15 years to memory switching devices fabricated from multicomponent chalcogenide glasses in which the reversible memory action is associated with the growth and destruction of a crystalline filament (Ovshinsky, 1968; Cohen et al., 1972; Steventon, 1974; Owen and Robertson, 1973; Owen et al., 1979). Although it is very likely that in the a-Si devices some form of filament formation (not necessarily crystalline) is taking place in the OFF–ON transition, the switching phenomena are clearly very different from those in the chalcogenides, at least operationally. The most obvious difference is the completely nonpolar character of switching in chalcogenide glass devices, in contrast to the marked polarity dependence of the a-Si memory switches. More important, perhaps, the switching and setting times for the a-Si device are much faster (~10 nsec for either the OFF–ON or ON–OFF transition, compared with at least several milliseconds in chalcogenide devices) and the energy involved in the switching process is considerably lower (1 μJ or less, compared with 1 mJ or more). Also chalcogenide glass devices require voltage pulses of magnitude 25–30 V (for a device ~1 μm thick) to establish the ON state, and very often they need 100 or more "forming" cycles before reasonably stable operation is achieved. This again contrasts with the operation of the a-Si memory switch, in which (for a total device thickness of ~1 μm) there is a single forming step with a threshold voltage of about 30 V, and for all subsequent operations the forward threshold ($V_{thF}$) is 4–6 V.

The closest parallel to the a-Si devices described in this paper seems to be the observation of memory switching in heterojunctions of $n$-type ZnSe grown epitaxially on $p$-type (single-crystal) Ge substrates, reported by Hovel (1970) and by Hovel and Urgell (1971). The ZnSe–Ge heterojunction devices are polar and the transition times for the OFF–ON and ON–OFF operation are both in the region of 100 nsec or less. Similar, but not so
well-substantiated, memory switching characteristics have also been briefly reported in devices fabricated by forming Schottky contacts on n-type GaAs and Si (single crystal) (Moser, 1972). Hovel and Urgell (1971), have tentatively and qualitatively explained switching in the heterojunction by a model involving the filling and emptying of traps in the ZnSe, with the formation of a current filament in the ON state. However, even in this case of superficially similar characteristics, there are notable differences. Most significantly, the polarity required for switching in the ZnSe–Ge heterojunction is the opposite to that found in the a-Si devices, and the OFF–ON threshold voltage for the ZnSe–Ge switch decreases substantially with temperature (from about 1 V at 200°K to less than 0.1 V at 400°K), whereas for the a-Si devices investigated so far is at the most only weakly temperature dependent. In addition, the memory state of the ZnSe–Ge heterojunctions is generally lost within a few weeks, whereas no change in the characteristics of the a-Si memory devices has been observed after storage over a period of 18 months.

To conclude this chapter it is perhaps worthwhile to note that the switching phenomena observed in the a-Si memories are not the only nanosecond processes known for this material. Drift mobility studies, which show that electron transit times across about 1-μm-thick films are of the order of 10 nsec or so, have been known for over a decade (LeComber and Spear, 1970; Spear, 1983). More recently, it has been demonstrated that hydrogenated amorphous silicon can be used to modulate light at subnanosecond speeds (Phelan et al., 1981; see also Chapter 13 by Phelan of this volume). In pulsed laser annealing of a-Si it has been suggested that the electron–hole plasma generated by the laser could produce rapid second-order phase transitions (van Vechten et al., 1979). The challenge and excitement in understanding the a-Si memories lies in discovering whether the origins of the fast switching processes are electronic, structural, or both.

ACKNOWLEDGMENT

The authors are grateful to the Venture Research Unit of British Petroleum International PLC for a grant supporting the research described in this paper.

REFERENCES

15. ELECTRONIC SWITCHING IN a-Si JUNCTION DEVICES

Inorganic resists based on photo-doped As-S films

A.P. Firth, P.J.S. Ewen, A.E. Owen and C.M. Huntley

Department of Electrical Engineering, University of Edinburgh, Edinburgh, EH9 3JU, U.K.

Abstract

In recent years there has been considerable interest in inorganic resist systems based on the photo-doping of amorphous chalcogenide films, the majority of the research being devoted to Ge-Se films. This paper presents a detailed investigation of inorganic resists based on the photo-doping of Ag into As-S films. It is shown that high resolution patterns can be produced in such resists using holography or optical lithography and that they are compatible with wet-chemical or plasma etching. Structural studies using Raman spectroscopy indicate that for best resolution the composition of the As-S film should be close to \( \text{As}_{33}\text{S}_{67} \) since on photo-doping it will yield a single-phase homogeneous material. A possible mechanism for the photo-doping process is described based on a tarnishing-type photo-chemical reaction. It is shown that the actinic radiation initiating the photo-dissolution effect is absorbed primarily in the photo-doped layer, close to the interface with the undoped As-S region.

Introduction

Interest in inorganic resist systems based on amorphous chalcogenide semiconductors stems from their potential to extend photolithography into the submicron region. In terms of resolution capability and linewidth control the performance of these systems matches that of organic resists and they are compatible with industry-standard processing techniques, such as plasma-etching and spin-coating. They can also be used as resists for X-ray, e-beam and deep-UV exposure, with an expected resolution capability of 0.3 \( \mu \text{m} \) in the latter case. Because chalcogenide glasses can be deposited by vacuum evaporation very thin, uniform films are obtainable, which makes these materials particularly suitable for multilevel resist techniques for producing chromium photomasks and also for ion-beam lithography.

The most extensively studied chalcogenide resist system is Ag/As-Se/Ge-Se in which the basis of image formation is the metal photo-dissolution effect. Resists based on the many other photo-induced phenomena which occur in chalcogenide glasses have yet to be fully investigated. In the present paper we report a study of resists based on the photo-dissolution of Ag into As-S films. The first section is concerned with the resolution capabilities and etching properties of these resists and the following section discusses the influence of As-S film composition on resist performance. The final section describes in qualitative terms a possible mechanism for the photo-dissolution process.

Ag/As-S resists

The resolution capability of the Ag/As-S resist system is demonstrated in Figures 1 and 2, which show patterns produced by the photo-dissolution of Ag into As-S films evaporated onto Si wafers. Figure 1 is an electron-micrograph of a grating produced in a film of composition \( \text{As}_{30}\text{S}_{70} \) by conventional holographic techniques using 488 nm (optical) illumination from an Ar-ion laser. The grating spacing is \( \approx 0.3 \mu \text{m} \). Figure 2 is a micrograph of a test pattern containing a range of line-widths and line-spacings and was also generated in an \( \text{As}_{30}\text{S}_{70} \) film but using a UV wafer stepper (Eaton Optimetrux 8010); line-widths and spacings down to \( \approx 0.6 \mu \text{m} \) were readily resolved. The thicknesses of the \( \text{As}_{30}\text{S}_{70} \) and Ag films were 3000 and 1000 \( \AA \) respectively. Typical exposure times on the wafer stepper for an incident intensity of 150 mW/cm\(^2\) were 3-4\( s\) in the case of the film thicknesses quoted above. These results are comparable with what can be achieved with the more widely studied Ag/As-Se resist.

The Ag/As\(_{30}\text{S}_{70}\) system is compatible with both plasma and wet-chemical etching. The most satisfactory wet etchants were found to be iron (III) nitrate solution (2.5 g/100 ml of methanol) for the Ag film and a saturated ammonia solution in methanol for the As-S film. These etchants do not attack photo-doped material and give etch rates of 100-300 \( \AA/s \). In most instances it was found that aqueous solutions caused the films to flake, although this effect can be reduced by annealing the films. Plasma etching was done in an IPC 2000 Series plasma system using CF\(_2\) gas. Photo-doping using an Ag layer above the As-S film was found to be preferable for plasma-etching since surplus Ag remaining on the surface after photo-doping can be easily removed before etching is commenced. When the Ag layer is below the As-S film it is exposed to the plasma when
the As-S is etched off, which causes it to flake.

**Optimization of the As-S film composition**

Most studies of Ag/As-S resists have used amorphous A52S3 as the chalcogenide film. The compound A52S3 is commercially available in bulk glassy form and it has the additional advantage that as a prototype amorphous material it has been extensively studied, so that a considerable amount of information on its properties exists. However, it is not obvious that A52S3 is necessarily the optimum resist composition and there is a need to investigate other As-S compositions. The glass-forming region in the As-S system extends from 5 to 43 atomic percent and any composition in this range can be deposited as an amorphous film. Bulk samples of these glasses for evaporation sources are easily made.

One of the basic questions regarding the photo-dissolution process is how the photo-dissolved Ag is incorporated in the As-S film? Is it accommodated interstitially, does it form a new homogeneous compound or does it produce a phase-separated mixture of new compounds? The compounds which might form depend on the composition of the As-S film and of the Ag source. Thus film composition influences the photo-dissolution process and hence, through it, resist performance.

The principal compounds of the As-S-Ag system together with the glass-forming regions are shown in the phase diagram of Figure 3. Also drawn are five tie-lines: four linking the compositions A52S3, A53S17, A55S19 and A58S28 with Ag and one between A52S3 and Ag25, which can also be used as a source of photo-dissolved Ag. When elemental Ag is used as the source to photo-dope one of these four As-S films the resulting overall composition during photo-dissolution must move along the appropriate tie-line terminating at the Ag vertex. None of these four tie-lines passes through any of the known ternary compounds and only the tie-line for A53S17 passes through the central glass-forming region. Thus of these four compositions only A55S19 can yield a homogeneous photo-doped film. Raman results obtained in the present study and shown in Figure 4 suggest that photo-doped as-evaporated A53S17 films are indeed homogeneous, having a structure similar to that of a bulk glass within the central region of glass formation and close to the Ag-As30 S70 tie-line. Spectra A and B of Figure 4 are for an As30 S70 film before and after photo-doping respectively; spectrum C is for a bulk glass of composition Ag30 As22 S48 and is clearly similar to the photo-doped spectrum B. Photo-doped films of A55S19, A53S17 and A58S28 on the other hand yield spectra indicative of a phase-separated structure: photo-doped A55S19 and A53S17 consist of a mixture of amorphous Ag and an amorphous Ag-As-S phase while photo-doped A58S28 is a mixture of Ag2S (not necessarily stoichiometric) and an amorphous Ag-As-S phase.

The existence of phase-separated regions in the photo-doped film may have important technical consequences since it could ultimately affect edge definition and resolution. In the case of binary (undoped) As-S compounds, for example, phase-separation occurs for As contents greater than 43 atomic percent and results in the formation of ~1 μm diameter pockets of crystalline AsS3 embedded in an amorphous As-S network. Such inhomogeneities in photodoped As-S resist films would be undesirable. The optical and chemical sensitivity of the resist may also be influenced by the phase separation which appears to be occurring in some of the photodoped As-S compositions. It has recently been shown that formation of Ag which can be incorporated into an evaporated As-S film on dipping in AgNO3 solution is at its lowest for the composition A93AsS3, which lies at the centre of the range of compositions whose tie-lines to the Ag vertex pass through the central glass-forming region. This result might be due to the fact that additional Ag can be incor-

The tie-line between A52S3 and Ag2S passes through the ternary compounds AgA5S2 and Ag25/As2S3, which suggests that a homogeneous photo-doped phase will be produced if Ag2S is used as the Ag source. This has been confirmed by Raman experiments carried out in the authors' laboratories, the photo-dissolution product being Ag2S/AgS. An e-beam resist based on Ag25/As2S3 has been developed by Singh et al. and has been shown to be capable of producing 0.1 μm lines and edge-to-edge spacings down to 0.03 μm. The sensitivity of this resist, however, is very low (4 × 10⁻⁹ C/cm²).

**The photo-dissolution mechanism in the Ag/As-S system**

An important objective in the study of the photo-dissolution effect is to identify where the actinic radiation which initiates the process is absorbed. The spectral dependence of the photo-dissolution rate suggests that the basic optical excitation occurs in the amorphous film but as the effect can be stimulated by light of energy less than the band-gap energy of the As-S glass some investigators conclude that the actinic photons are...
absorbed in the Ag film. A third possibility is that the actinic radiation is absorbed in the photo-doped region between the Ag and the undoped As-S; our results support this and show, in addition, that absorption takes place primarily near the interface between the doped and undoped regions.

Figure 5 is a diagram of the sample geometry used in experiments designed to investigate the lateral migration which can occur when photo-dissolution is carried out on a conducting substrate. A film of gold (which does not significantly photo-dissolve into As-S films) was first deposited over a section of the surface of a glass slide and a band of Ag 3 mm wide and 600 nm thick was evaporated on top. Finally a 60 nm thick As-S film of composition As;S; was evaporated over the whole surface. The thickness of the Ag band is considerably larger than that of the As-S film in order to provide a reservoir of Ag for lateral photo-dissolution over the substrate surface.

By focussing the light source through a microscope objective it was possible to illuminate small areas of the sample. Normal photo-dissolution occurred when the region above the Ag band was illuminated whereas illumination of the film above the gold had no effect. However, when a photo-doped region and a neighbouring undoped region above the gold were exposed simultaneously the Ag migrated into the undoped region. Figure 6 is a micrograph of a photo-doped strip that was drawn out by tracking the illuminated spot across the film over the gold layer. To produce the strip only a segment of the interface between the photo-doped and undoped regions had to be in the illuminated area - it was not necessary to illuminate the photo-doped region between this interface and the Ag band or the Ag band itself. Hence the actinic radiation is being absorbed in the photo-doped region close to the interface with the undoped region.

The photo-doped strip could be drawn out indefinitely, provided the Ag source was not consumed. Figure 7 shows the time dependence of the distance advanced by the tip of the photo-doped strip as the illuminated spot was tracked stepwise across the As-S film over the gold. The plot is linear, indicating a constant speed of dissolution, and its slope gives the rate of advance as ~ 1 μm per minute.

The lateral photo-dissolution occurs only over the conducting gold-coated part of the substrate and this observation suggests that a supply of electrons is necessary for photo-dissolution to proceed. Ordinarily, the electrons would be supplied by the Ag source and would have to diffuse through the reaction products to reach the interface but in the present experiment electrons can be supplied by the gold and they do not have to diffuse further than the thickness of the As-S layer.

Recent measurements in the authors' laboratory of the change in transmissivity of Ag/As-S systems during photo-dissolution indicate that a single, highly-absorbing photo-doped layer builds up in the As-S film, its thickness increasing as the square root of exposure time. This, together with the structural information and lateral dissolution results, suggests a mechanism involving a solid state photo-chemical reaction analogous to tarnishing. When Ag is exposed to S a tarnishing reaction occurs which results in the formation of an Ag;S layer on the Ag surface. This reaction is well understood and it involves Ag+ ions and electrons from the Ag diffusing through the Ag;S reaction product to the Ag;S/S interface and there reacting with S, so that the Ag;S layer thickness continues to increase although the Ag and S are separate. Similarly, the photo-dissolution mechanism may involve Ag+ ions and electrons from the Ag source passing through the photo-doped layer to reach the undoped As-S region, where they react to produce a homogeneous compound or mixture of compounds. This mechanism is illustrated schematically in Figure 5. The Ag/As-S reaction product is likely to be a mixed conductor (i.e. ionic and electronic) so that electrons as well as Ag+ ions will be able to diffuse through it and indeed the speed at which they diffuse may govern the photo-dissolution rate since the photo-doped region advances much more quickly for the case of lateral photo-dissolution, where electrons are supplied by the conducting substrate.

The role of illumination in the process is probably to stimulate the diffusion of Ag+ ions through the Ag/As-S region to the interface with the undoped As-S. Why this occurs is not clear but a similar effect is observed in bulk Ag/As-S glasses and crystalline Ag;As-S (proustite) and results in the formation of a silvery deposit on the illuminated surface. Illumination with light of energy greater than the band-gap energy of the As-S film will also create electron-hole pairs in the undoped region. Electrons in As-S glasses are trapped immediately at defects, possibly 3-fold co-ordinated S atoms. Thus at the interface between the photo-doped and undoped regions the electrons available for neutralising the Ag+ ions may be associated with S atoms, thus facilitating the formation of Ag-S bonds. The existence of two illumination-dependent processes, one in the undoped As-S region and one in the photo-doped Ag/As-S region, would explain why the process has a spectral dependence which resembles the absorbance profile of the As-S film but can still be stimulated by below-band-gap radiation.
If the proposed mechanism is correct then the speed at which the photo-doped layer builds up will be governed by the rate at which Ag\(^+\) ions from the Ag source diffuse through the photo-doped region to reach the interface with the undoped As-S. This implies that the thickness of the photo-doped layer will increase as the square root of the exposure time so that unreasonably long exposure times will be required to photo-dope thick films.

### Conclusions

Inorganic resist systems based on the photodoping of Ag into amorphous As-S films have been investigated, particularly with regard to the influence of film composition on resist performance. It is shown that submicron resolution is achievable with As\(_{30}\)S\(_{70}\) films and that such resists are compatible with plasma and wet-chemical etching. Structural studies using Raman spectroscopy suggest that As-S compositions such as As\(_{33}\)S\(_{67}\), which are close to As\(_{33}\)S\(_{67}\), will yield the best resolution, although they may not be as optically or chemically sensitive as As\(_{33}\)S\(_{67}\).

Experiments on the photo-dissolution of Ag into As-S films indicate that the actinic radiation initiating the effect is absorbed primarily in the photo-doped layer, close to the interface with the undoped region of the As-S film. A possible mechanism was described based on a tarnishing-type photo-chemical reaction.

### Acknowledgements

The authors are grateful to the Edinburgh Microfabrication Facility for the use of their UV wafer stepper and ancillary equipment. This work was supported by the Science and Engineering Research Council.

### References

Figure 1. Electron-micrograph of a grating produced holographically by Ag photo-doping of an As$_{30}$S$_{70}$ film.

Figure 2. Micrograph of a test pattern produced in an As$_{30}$S$_{70}$ film by Ag photo-doping using a UV wafer stepper.

Figure 3. Phase diagram of the Ag-As-S system showing the principal ternary compounds and the glass-forming regions (cross-hatched areas). Tie-lines are discussed in the text.

Figure 4. Raman spectra of films of composition As$_{28}$S$_{70}$. Spectrum A is prior to photo-doping with Ag and Spectrum B that of the photo-doped film. Spectrum C is for a bulk glass of composition Ag$_{3}$As$_{2}$S$_{48}$.
Figure Micrograph of a photo-doped strip drawn out in an As-S film above a gold layer. The Ag source is at the left.

Dimensions: 1 cm = 100 μm.

Figure 6. Micrograph of a photo-doped strip drawn out in an As-S film above a gold layer. The Ag source is at the left. Dimensions: 1 cm = 100 μm.

Figure 5. The sample geometry used in the lateral photo-dissolution experiments: (a) plan view, (b) cross-sections at indicated regions.

Figure 7. Distance travelled by the photo-dissolution front (dissolution length) against time. The straight line is a linear fit to the data points and has a gradient of ~1 μm per minute.

Figure 8. Diagram showing the formation of a photo-doped layer of Ag-As-S material as a result of a tarnishing-type reaction involving the diffusion of Ag⁺ ions and electrons. a-b is the interface between the Ag and the photo-doped layer and c-d the interface between the photo-doped layer and the undoped region. Interface c-d advances through the As-S region as the square root of the exposure time.
THE SWITCHING MECHANISM IN AMORPHOUS SILICON JUNCTIONS


*Carnegie Laboratory of Physics, University of Dundee, Scotland.
*Department of Electrical Engineering, University of Edinburgh, Scotland.

Extensive new results have been obtained on memory switching in a-Si p+ni junctions. It is shown that the ON-state has its origins in a highly conducting filament less than \( \mu \)m in diameter. The physical mechanisms that could play a role in the switching operations are discussed.

1. INTRODUCTION

Threshold and memory switching in the chalcogenide glasses greatly stimulated research on these materials in the late 1960s and early 1970s\(^1,2\). In spite of the wide-ranging work on amorphous Si (a-Si) in recent years, memory switching phenomena were not observed and this led to a general belief that homogeneous films of this and other tetrahedrally bonded amorphous materials do not possess the physical properties required for switching behaviour\(^3\). However, some three years ago we demonstrated that heterogeneous junction layers of a-Si could be made to exhibit reliable, fast, polarity dependent memory switching\(^4,6\). The present paper is concerned with recent results on these specimens.

As prepared, the a-Si layers, generally with a p+ni structure, require one forming operation after which they are in a non-volatile low resistance ON-state. They can be switched back (ERASED) to a non-volatile high resistance OFF-state by the application of a negative potential to the p layer and switched ON again (WRITE operation) by a positive voltage. This cycle can be repeated many times\(^5,6\). The particular exciting aspect of these devices is their remarkable switching speed. Pulses of a few volts in height and a few tens of nsecs duration\(^5,6\) are sufficient for both the WRITE and ERASE operations.

In the first part of this paper our latest results, mainly for formed devices, will be presented. This will be followed by a discussion of possible physical mechanisms that could be playing a role in the switching operation.

2. SPECIMEN PREPARATION

The samples were prepared by the RF glow discharge decomposition of silane or appropriate silane gas mixtures. The majority of the results have been obtained on specimens having a p+ni structure where i denotes an undoped layer. We have also observed similar results for n+pi and related structures and recently
Gangopadhyay et al.\(^7\) have reported memory switching in p\(^{+}\)n\(^{-}\) samples. The samples were deposited onto Corning 7059 glass substrates at 300°C and generally prepared with Cr bottom electrodes and Al or Cr top contacts, although a number of other metals have been investigated. The active area of the samples, defined by photolithographic techniques, ranged from about 10\(^{-7}\) cm\(^2\) to 10\(^{-3}\) cm\(^2\). The work has also been extended to specimens in which the a-Si \(i\)-layer was replaced by an insulator. The results for these samples showed the same general features as those reported for the p\(^{+}\)ni layers and will be described in subsequent publications.

3. STATIC EXPERIMENTAL RESULTS

A number of dc experiments have been carried out on formed devices in an attempt to provide additional information primarily about the nature of the ON-state. The current-voltage characteristics of recent p\(^{+}\)ni specimens had the same general features as those reported previously\(^4\)\(^-\)\(^6\), although by carefully controlling the forming process their ON-state resistances were generally kept at a somewhat higher level (\(\sim 1\,\text{k}\Omega\)).

3.1. Area Dependence of ON- and OFF-state Resistances

Fig. 1 shows the area dependence of the OFF-state resistance, \(R_{\text{OFF}}\), for samples of different area fabricated from a single p\(^{+}\)ni deposition run. These results were obtained on specimens that had been previously switched many times. Within the experimental scatter, \(R_{\text{OFF}}\) clearly scales with the reciprocal of the area \(A\) demonstrating that in the OFF-state the current flows throughout the whole area of the specimen.

![FIGURE 1](image1)

Area dependence of \(R_{\text{OFF}}\).

![FIGURE 2](image2)

Area dependence of \(R_{\text{ON}}\).
In complete contrast, the values of the ON-state resistances, $R_{ON}$, for the same specimens show no area dependence at all as can be seen from fig. 2. This result can be understood only if the ON-state has its origin in a highly conducting filament, less than a few $\mu$m in diameter, that extends at least part of the way through the specimen thickness. This result, although not entirely unexpected, will obviously be of central importance in developing a model for the switching process. A direct proof of the existence of a filamentary ON-state is provided by the results in section 4.

3.2. Temperature Dependence of the ON- and OFF-states

The temperature dependence of both the ON- and OFF-state conductance has been measured over the temperature range from about 230K to 400K. Results for a typical $p'n'i$ specimen are shown in fig. 3. The OFF-state conductance $G_{OFF}$ varied slowly with temperature, increasing by less than a factor of three between 230K and 360K. The ON-state conductance was even less temperature dependent, increasing by only 10% over this temperature range. The insensitivity of these device parameters to temperature is also observed in other properties. For example, the magnitudes of the WRITE and the ERASE voltage, measured under pulsed conditions, both increase by only a factor of 2.5 as the temperature is decreased from 400K to 200K. Clearly the general insensitivity of the switching parameters must also be a feature of any model proposed to explain the switching behaviour.

3.3. Magnetoresistance Measurements of the ON-state

Transverse magnetoresistance measurements of the ON-state of a number of specimens have been carried out at room temperature up to fields of 0.5T. Within the experimental error, $\Delta G/G$ was proportional to $B^2$ and found to be positive. The values of $\Delta G/(G B^2)$ ranged from $(0.5$ to $2.0) \times 10^{-4}T^{-2}$. Experiments on phosphorus doped a-Si after thermal crystallisation also gave a positive magnetoresistance with the same $\Delta G/(G B^2)$ dependence. Unfortunately, no results on homogeneous undoped glow discharge a-Si films have been reported.

It is therefore difficult to draw any definite conclusions about the amorphous or crystalline
nature of the ON-state filament from the present results. However, it is probably correct to associate the $B^2$ dependence of the memory ON-state with a longer mean free path than is normal in amorphous solids. Using conventional crystalline theory the magnitude of the measured $\Delta V/(PB^2)$ would lead to an effective mobility $\sim 100 \text{cm}^2\text{V}^{-1}\text{s}^{-1}$ in support of this suggestion.

4. OBSERVATION OF ON-STATE FILAMENT

In order to learn more about the ON-state a series of experiments have been carried out in which the surface of the a-Si device was covered with a thin layer of thermochromic liquid crystal. By passing current through the device in the ON-state it is possible to observe the current path from the changes produced by Joule heating in the reflected colour of the liquid crystal (LC). The results for two 20µm diameter structures, viewed through a high power optical microscope, are given in fig. 4. Fig. 4(a) shows the LC surface in the absence of any current flow through the specimen and fig. 4(b) shows a sample with current flowing in the ON-state. The change in the LC appearance produced by Joule heating within the filament can be clearly seen. Fig. 4(c) represents another device in which the filament has been made more visible by passing a larger current through it. We estimate from a number of measurements that the ON-state is associated with a filament of maximum diameter 0.5µm. The result therefore has interesting implications for the understanding of the switching process. The visual observation of the current filament has also enabled us to establish that switching a device OFF and ON again produces the current filament in the same place and this implies that the switching processes are not destructive.

In another series of experiments the specimens were covered by a thin layer of a liquid crystal, 4-cyano-4 altylbiphenyl, which undergoes a nematic-liquid phase transition at 35.3°C. The phase boundary may be observed quite easily in cross-polarised light and the transition was found to be fast and without

![FIGURE 4](image-url)

Photographs of LC surface covering 20µm diameter memories. The arrows in (b) and (c) denote the positions of the current filaments.
hysteresis. If the sample temperature is fixed using a thermostatically controlled stage, the difference in temperature between a region of local heating (at a temperature $T_h$) and the surrounding film (at $T_f$) may be determined. The results indicate that no observable temperature rise occurred as a result of applying electrical power to the pore in either the unformed stage just prior to forming, or in the formed OFF-state just prior to switching. However, as described above, in the ON-state local heating which results on applying continuous power could be clearly seen. The effect of changing the RMS power applied to a 50µm diameter pore was studied using a continuous train of 300ns pulses. The stage was maintained at 21°C, thus the phase boundary represented the locus of points $(35.3-21) = 14.3°C$ above the film temperature. These isotherms were circular and in the particular case studied were symmetric about the centre of the pore.

A plot of the phase boundary diameter $d_1$ vs. RMS power $P$ is shown in fig. 5. Although there is considerable scatter it can be seen that the relationship between $d_1$ and $P$ is substantially linear for $d_1 \geq 2\mu m$. Below this the accuracy of the measurements is limited by the resolution of the microscope used. A linear dependence of $d_1$ on $P$ is obtained as a solution of the heat conductivity equation for an idealised system of this kind, in which the source of local heating is assumed to be much smaller than $d_1$. Thus this data indicates that the diameter of the ON-state filament must be much less than 2µm, in agreement with the above experiments.

5. DYNAMIC BEHAVIOUR

In an earlier publication the existence of a voltage dependent delay time had been established for the forming operation. In the course of our recent work we have observed a number of other time dependent effects and these will be described in the following.

5.1. Current Instabilities in Unformed Structures

Current instability phenomena have been observed at voltages below those required for forming, in unformed a-Si p+ni structures with thin i-layers. Typical results obtained under pulsed operation are shown in fig. 6 for three pulse heights decreasing in amplitude from (i) to (iii).
It can be seen that the current increases after a delay time such as $\tau_{dd}$. At first sight the results in fig. 6(a) for the unformed a-Si memories appear to be similar to those observed in crystalline Si MISS devices. However, the a-Si p'nn device does not remain in a high conductivity state whilst the voltage is maintained, as is the case for the MISS. The current increases fairly rapidly and then decays over a somewhat longer time-scale, resulting in an asymmetric current pulse as shown in fig. 6(b). However, it is worth emphasising that these current instabilities were observed at voltages just below those required for forming. If the voltage pulse was left on for many tens of microseconds then, some time after the first, a second current pulse would propagate through the sample, and then a third, etc. As the applied voltage $V$ approached the forming voltage $V_F$ these current pulses appeared to merge until at $V = V_F$ the current level remained high as the device formed. These current instabilities therefore appear to be an important precursor to the forming process.

5.2. WRITE and ERASE Delay Times

We have recently observed a delay in the WRITE operation which is a strong function of the WRITE pulse magnitude as shown in fig. 7. Note that these delay times $t_d$ are significantly faster than the forming delay times reported previously. If the results in fig. 7 are expressed in the form $t_d = t_o \exp (-V/V_o)$ then $t_o = 335$ns and $V_o = 4.5$V. Similar results have been obtained for all the specimens investigated although the $V_o$ values ranged from about 0.5V to 13V.
Experiments to measure any ERASE delay time were unsuccessful implying that any delay was less than the rise time of the measuring circuit, i.e. typically less than 5ns.

6. DISCUSSION OF POSSIBLE SWITCHING MECHANISMS

The work described above contained a number of important new results. Probably the most crucial of these to an understanding of the physical processes underlying the switching mechanism, was the proof of the existence of a filament in the memory ON-state. Filamentary conduction has been observed previously in a wide range of materials. These include single crystal silicon, gallium arsenide, zinc telluride, gallium arsenide phosphide and polycrystalline silicon, all of which can show current-voltage characteristics associated with threshold switching. This applies as well to the amorphous chalcogenide glasses in which both threshold and memory behaviour can be observed. In the following we shall begin by discussing the filament formation and then describe some of the models that have been used for these materials and discuss their relevance to the switching in a-Si.

6.1. Filament formation

During the forming process which, as has been demonstrated, leads to a current filament, the metal-insulator barrier of the device will be under reverse bias. This suggests that the forming is likely to be associated with extremely high local fields across the thin a-Si i-layer which may approach $10^6 \text{Vcm}^{-1}$. Under these conditions tunnelling of field emitted electrons from the top of the electron distribution in the metal electrode becomes the dominant transport mechanism, injecting appreciable electron densities into the semiconductor. The current instabilities described in section 5.1 would certainly be consistent with such a model.

How is the reproducible filament produced during the forming process? A possible answer is suggested by the extensive work on thin film composite
materials in which small isolated metallic particles are dispersed in a dielectric medium\textsuperscript{11}. This work has given a great deal of insight into the tunneling process between isolated metal particles and its dependence on average particle size and separation, applied fields, etc. It is feasible that the high fields and current densities developing locally during forming could lead to enhanced diffusion of metallic particles from the electrode into the thin insulator. Such a region would become the preferred current path carrying the electron current in the ON-state. Further experimental work is required to confirm such a model; the observed field and temperature dependence in the ON-state is certainly consistent with results established in the previous work on these systems.

6.2. Possible Mechanisms for Memory Switching

The above considerations refer to the initial formation of the filament but cannot as such explain the subsequent memory switching. We should now like to critically discuss a number of possible mechanisms to explain this behaviour.

6.2.1. Thermal models

At first sight this model\textsuperscript{12}, used to explain the behaviour of the amorphous chalcogenide memories, might appear to offer a basis for explaining the a-Si switching processes. In the chalcogenides the ON-state is associated with a filament of crystalline material which is formed after sufficient power has been applied to the layer to melt a small area of the material. Switching OFF is achieved by burning out this filament using a number of relatively short high-power pulses and allowing rapid quenching to reform the highly resistive amorphous phase. However, there are a number of important differences between the a-Si and the chalcogenide memories: (a) it has been established that the a-Si memories do not form or WRITE at constant power; in general forming occurs at much lower powers (\(<10^{-8}\) J) than in the chalcogenides\textsuperscript{10^{-4} - 10^{-3}} J); (b) the forming, WRITE and ERASE operations for the a-Si memories are generally polarity dependent; (c) no rise in the temperature of the a-Si specimens can be observed immediately prior to switching; and (d) the a-Si WRITE and ERASE times are many orders of magnitude faster than those for the chalcogenides e.g. for the WRITE operation\textsuperscript{10^{-8}} s compared with\textsuperscript{10^{-3}} s.\textsuperscript{2} For all these reasons we do not believe that the crystalline/amorphous thermal model is applicable to a-Si.

6.2.2. Models based on Trapped Space Charge

In many respects the behaviour of the amorphous Si layers appears to be closely related to that of crystalline Si MISS structures in that both show fast polarity dependent switching, both show current instabilities during some stage of the forming processes and both have high conductance states associated with current filaments. The major difference of course is that the MISS structures are threshold switches which always revert to the OFF-state when the power is...
removed, whereas the a-Si layers have the additional advantage (and complexity) of non-volatile memory behaviour.

It is nonetheless possible that the memory switching in the a-Si devices is closely related to the mechanisms proposed for the MISS layers. Essentially two models have been used to explain the MISS behaviour\textsuperscript{13}. These are generally referred to as "punch-through" and the "avalanche-mode" and both rely for switching on establishing high fields across space charge barriers in the films. In addition, in both models the low impedance state is produced by injected charge causing inversion of the Si at the Si/insulator interface. It is tempting to suggest that the "permanent" memory of the a-Si layers may be produced by a similar mechanism in which the charge is trapped in deep gap states at the insulator-semiconductor interface for which the probability of release is very small. However, the a-Si layers retain their memory without any observable change in their properties for as long as we have measured them, namely, for over one year at room temperature and 24 hours at 95°C.

Using the thermal release time from deep mid-gap states as a measure of the persistence of the trapped space charge, then from the well-known expression for the average thermal release time we estimate that the capture cross-section of these centres should be less than $10^{-18}$ cm$^2$. Although extremely small, such values would be consistent with Coulomb repulsive centres identified in crystalline materials\textsuperscript{14}. However, the problem is that recombination of the trapped distribution through tunnelling or diffusion may well invalidate the above estimate by leading to a much faster decay of any trapped space charge distribution. All one can say at present is that a model in which the observed memory is associated with a trapped space charge cannot be excluded but, in view of the remarkable non-volatility of the memory states, may not be the correct explanation.

6.2.3. Other suggestions

It should be remembered that in the random network of the a-Si layers significant amounts of hydrogen are incorporated. The possibility therefore exists that memory switching may be associated with some atomic motion of hydrogen in the material. For instance, it has been reported that in ambient sensors containing Pd/a-Si Schottky barriers, hydrogen plays an important role in lowering the barrier\textsuperscript{15}. Also the polarity dependence of the threshold voltages for the a-Si memories could be understood on the basis of field assisted hydrogen diffusion. Like all the possibilities mentioned above, this remains at present somewhat speculative and further work is required to identify the most likely memory switching mechanism.

7. CONCLUSIONS

A number of new results, including the observation of a filamentary ON-state,
are reported in this paper. These provide important information relating to
the physical processes underlying the operation of the a-Si memories.

At present we believe that the initial stages of memory formation are likely
to be associated with high fields and/or trapped charges in a manner analogous
to that responsible for the threshold switching in crystalline Si MISS devices.
However, neither the nature of the ON-state current filament nor the physical
mechanisms responsible for the permanent memory of the a-Si layers, have been
established with any certainty at the present time.

ACKNOWLEDGEMENTS

We should like to thank S. Kinmond and A. Carrie for their help with the
specimen preparation. The financial support of the B.P. Venture Research Unit
for some of the work described in this paper is gratefully acknowledged.

REFERENCES

   59 and 60 (1983) 1273.
6) P.G. LeComber, A.E. Owen, W.E. Spear and J. Hajto, in: Semiconductors and
   pp 275 - 289.
7) S. Gangopadhyay, J. Geiger, B. Schröder, H. Rübel and S. Iselborn, in print.
9) J. Buxo, A.E. Owen, G. Sarrabayrouse and J.P. Sebaa, Revue de Physique
   Appl. 13 (1978) 767.
14) A. Rose, Concepts in Photoconductivity and Allied Problems (Krieger, N.Y., 1978)
15) A. D'Amico and G. Fortunato, in: Semiconductors and Semimetals, Vol. 21D,
A theoretical analysis of transmissive gratings of the type likely to be produced by the metal photodissolution effect in chalcogenide glasses has been carried out and used to predict diffraction efficiencies for a typical chalcogenide system, namely As$_{30}$S$_{70}$ glass photodoped with Ag. This was done to assess the potential for using these devices as diffractive elements for the 8-14 $\mu$m region. For bulk diffractive devices with a rectangular modulation profile, predicted efficiencies in the multiwave regime were over 90% and in the volume regime up to 100%, irrespective of profile. Surface relief gratings with a rectangular profile yielded a maximum efficiency of at least 73%.

The refractive index of Ag-doped and undoped As$_{30}$S$_{70}$ glass as a function of wavelength was measured for use in the grating analysis and was found to be approximately constant above 1 $\mu$m, the value depending on the amount of Ag photodoped into the glass. For As$_{30}$S$_{70}$ containing approximately 27 at% Ag the refractive index at 2$\mu$m is 2.8 compared with 2.2 for undoped As$_{30}$S$_{70}$, and this large difference is expected to be retained beyond 10 $\mu$m.

1. Introduction

Chalcogenide glasses exhibit a wide variety of photoinduced phenomena that enable them to be used as optical imaging or storage media (Owen et al 1985). The use of these glasses in optical imaging has been extensively studied over the past decade because of their potential applications as photoresists for VLSI lithography (Tai et al 1982). Their use as optical storage media, particularly for holography, has also received considerable attention (Bordogna and Keneman 1977), though little work has been done on modelling the performance of gratings produced in these materials.

Chalcogenide glasses are also well known as $\mu$m-transmitting materials, having pass bands (depending on composition) from the visible to beyond 15 $\mu$m (Savage 1985), so that, by making use of the photoinduced phenomena that occur in these glasses, it should be possible to produce diffractive elements for use at IR wavelengths. Such elements have several potential uses, for example in beam combining, filtering and spectral analysis, and also have advantages over conventional IR refractive elements (typically Ge lenses) as regards weight, cost and ease of manufacture (Close 1975).

This paper is chiefly concerned with a theoretical analysis of the performance of chalcogenide gratings, particularly with a view to their use in the IR. Since refractive index is one of the main parameters governing grating performance, we also present some refractive index measurements for a typical chalcogenide system that might be used to fabricate these gratings in practice, namely As–S glasses containing photodissolved Ag. The metal photodissolution effect is probably the most useful of the various photoinduced phenomena as far as these applications are concerned because it produces the largest change in the properties of the chalcogenide, particularly the refractive index and etch resistance.

2. Refractive Index measurements

Samples were prepared by evaporating onto Corning 7059 glass substrates first a layer of Ag and then a layer of As$_{30}$S$_{70}$ glass, there being no break in the vacuum between evaporations. The composition As$_{30}$S$_{70}$ was chosen for these measurements because earlier studies have shown that it is around the optimum composition with regard to resolution and sensitivity and because it exhibits negligible photodarkening. The As$_{30}$S$_{70}$ film thickness was approximately 5000 Å for all samples but the thickness of the underlying Ag layer was different for each sample, varying from 500 to 1500 Å; by varying the thickness of the Ag layer in this way, it was possible to control the amount of Ag that photodissolved into the As$_{30}$S$_{70}$ and hence the Ag concentration in the final photodoped material. The maximum possible Ag concentration is around 33 at% for As$_{30}$S$_{70}$ (Janai 1981) but may vary with As–S composition. Photodissolution
The thickness of the initial Ag layer and these thicknesses are indicated in the figure. The Ag concentration in the films was controlled by varying the thickness of the initial Ag layer and these thicknesses are indicated in the figure.

was achieved by illuminating samples with a 200 W mercury lamp.

The transmission of the samples was measured as a function of wavelength using a Perkin-Elmer Lambda 9 spectrophotometer and the refractive index as a function of wavelength using a Perkin-Elmer Lambda 9 spectrophotometer. The refractive index obtained over the wavelength range 500–2000 nm for undoped As₂S₃ and for four photodoped samples containing various amounts of Ag.

The curves are similar in shape but are displaced to higher values of refractive index at the Ag content is increased. For phase gratings an important quantity is the difference between the refractive indices of the doped and undoped material, and from the figure it is seen that for the most heavily doped sample, which contained approximately 27 at% Ag, Δn > 0.5 over this wavelength range. Single-oscillator fits were obtained for each of the curves and extrapolation of these to longer wavelengths showed that for the most heavily doped sample Δn > 0.5 at 10.6 μm.

Although the optical constants of As–S–Ag glasses have not been reported previously, those for two of the crystals in the system have been measured: smithite (AgAsS₂) and proustite (Ag₃AsS₄) were both found to be weakly absorbing from 0.65 to 12 μm and had an approximately constant refractive index in the IR, the values being 2.33 for smithite (Wehmeier et al 1968) and 2.52 (n₀) and 2.73 (n₁) for proustite (Hulme et al 1967). As the structure of the As–S–Ag glasses is believed to contain the same structural units as smithite, it is likely that for them also n remains constant above 2 μm so that extrapolation of the oscillator fits to longer wavelengths is valid.

3. Theoretical analysis

In this section we analyse the performance, in particular the efficiency, of both bulk and surface relief gratings.

The grating is assumed to be phase-modulated, lossless, unslanted and uniform with depth, with a rectangular modulation profile. This situation approximates mask exposure in chalcogenides (Owen et al 1985). A coupled wave approach is used (Moharam and Gaylord 1981, 1982).

The grating is taken to be in the form of an infinite parallel-sided slab of thickness d in the y–z plane, with a normal in the x direction. The input and output surfaces of the slab are bounded by homogeneous regions. An infinite, monochromatic plane wave, polarised perpendicular to the grating vector and of amplitude A₀, is incident upon the slab.

In the slab (0 < x < d) the grating parameters may generally vary as a function of y and z, but the slab is assumed to be locally plane (Sym and Solymar 1981). The rectangular modulation profile can be written as a Fourier series:

$$e(r) = e_0 + \sum_{i=1}^{N} e'_i \cos(k \cdot r)$$

with

$$e_0 = e_{\text{min}} + \Delta \varepsilon \mu$$

$$e'_i = (2/\pi)\Delta e \sin(i\mu x)$$

where $e_0$ is the bulk relative dielectric constant of the slab and $e'_i$ is the phase modulation of the ith harmonic of the grating profile. K is the grating vector (with $|K| = 2\pi/\Lambda$, where $\Lambda$ is the grating period), $\Delta \varepsilon = e_{\text{max}} - e_{\text{min}}$ and $\mu$ is a fill (mark/space) parameter.

The key coupled wave equations describing propagation inside the grating can be written as the set (for $m = -N, \ldots, -1, 0, 1, \ldots, N$)

$$-k_1 \frac{1}{d^2} A_m + \frac{d A_m}{d z} + m \omega (m + P) A_m$$

$$+ \sum_{i=1}^{N} k_i (A_{m+i} + A_{m-i}) = 0$$

where $A_m$ is the amplitude of the mth diffraction order, $k_i = \beta e_i / \varepsilon_i \beta$, $\beta = 2\pi N e_0 / \Lambda / \Lambda$ is the propagation constant in the grating, $\lambda$ is the free-space wavelength of the radiation, $\varepsilon = \varepsilon_0 / \sqrt{e_0} \theta_0$ is a modulation parameter, $\Omega = K^2 / 28 \varepsilon_0$ is the volume parameter (Solymar and Cooke 1981), $P = (\sin \theta_0)2B/\Delta$ is a Bragg parameter and $\theta_0$ is the propagation angle of wave zero. Thus each diffraction order $A_m$ is coupled to the m ± i orders $A_{m+i}$ and $A_{m-i}$ by the ith coupling coefficient $k_i$. The coefficients of $A_m$ are a measure of the mismatch in phase velocities of the diffracted orders.

The diffracted amplitudes $A_m$ are obtained by solution of these equations, subject to grating boundary conditions (matching tangential electric and magnetic fields) at $x = 0, d$. Both the cases of bulk and surface relief gratings can be analysed in this way. However, it is possible to simplify the calculations by neglecting second derivatives in (3). This is valid if $e'_i / \varepsilon'_0 < 1$, generally a figure of 0.2 being the limit. For bulk and surface relief gratings in photodoped chalcogenides, $e'_i / \varepsilon'_0 = 0.1$ and 1.0 respectively.
So for the bulk case (3) becomes (for $m = -N, \ldots, 1, 0, 1, \ldots, N$)

$$\frac{dA_m}{dz} - jm\Omega(m + P)A_m$$

$$+ j \sum_{r=1}^{N} N \lambda_r (A_{m+r} + A_{m-r}) = 0.$$ (4)

No backward diffracted waves are now possible. The transmission grating boundary conditions are

$$A_m(x = 0) = 0 \quad (m \neq 0), \quad A_0(x = 0) = 1.$$ (5)

Analytic solutions of (4) are possible for the thin ($\Omega < 0.1$) and volume ($\Omega > 10$) diffraction regimes. Defining efficiency as the ratio of power in the first diffracted order to that in the incident wave, a maximum of 40.5% is found for a square wave grating in the thin regime (e.g. Magnusson and Gaylord 1978) while 100% is possible in the volume regime, irrespective of profile (Slinger 1988). For the multiwave regime, numerical solution of (4) is necessary.

4. Numerical results

4.1. Bulk gratings

Equations (4) were solved numerically to determine the efficiency of first-order diffraction, $\eta_1$, as a function of $\Omega$ and $\zeta$. The results for a square wave grating are shown in figure 2. Fresnel reflection losses have been neglected. The thin, multiwave and volume diffraction regimes can all be seen. Perhaps surprisingly, high diffraction efficiencies of over 90% are possible in the multiwave region. For the majority of grating applications, operation in the volume regime is usually desired. Note, however, that this may not always be possible—the high refractive index of the chalcogenide may require operation at lower values of $\Omega$.

Using a bulk square wave grating with refractive indices of 2.22 (undoped) and 2.42 (photodoped region) and at ($\zeta$, $\Omega$) = ($\pi/2$, 5) requires a grating thickness of 3.6 $\mu$m—around 38 $\mu$m at $\lambda = 10.6 \mu$m.

4.2. Surface relief gratings

Here the rigorous theory (equations (3)) must be used, due to the large values of the ratio $e_i/e_0$. The value of the $\Omega$ parameter still governs the diffraction regime that the grating operates in, but there are no simple expressions predicting the performance in the limiting cases of thin or volume behaviour. Figure 3 shows the theoretical variation in efficiency for a typical photodoped grating.

For the grating analysed above, a maximum efficiency of 73% at $d = 1.36 \mu$m is seen. Here ($\zeta$, $\Omega$) = (3.7, 0.35). The grating is in the multiwave regime. This efficiency is lower than the bulk grating case but should not be regarded as a maximum. It may be increased by the use of suitable coatings or higher values of $\Omega$. Perhaps more importantly, the required grating depth is significantly less than for the bulk grating, being about 14 $\mu$m for operation at 10.6 $\mu$m. This could be of value in minimising grating exposure times.

5. Summary

In conclusion, for rectangular profiles, high efficiencies are achievable in the IR for both bulk and surface relief
gratings, provided the thicknesses are sufficiently large, the required thickness for the surface relief device being approximately a third of that for the bulk device.

Acknowledgments

This work was carried out with the support of the Procurement Executive, Ministry of Defence. CWS would like to thank F S M Clube at GEC Marconi Research Centre for assistance with some of the surface relief software.

© 1988 Controller HMSO.

References


Savage J A 1985 in Optical Materials and their Antireflection Coatings (Bristol: Adam Hilger) p 79
DRY ETCHED HIGH RESOLUTION POSITIVE AND NEGATIVE INORGANIC PHOTORESIST

E. HAJTO, R.E. BELFORD, P.J.S. EWEN and A.E. OWEN

Department of Electrical Engineering, University of Edinburgh, Edinburgh EH9 3JL, Scotland

The etching properties of spin coated As-S and Ag:As-S films are reported. Novel results obtained on both the light induced effects and the dry etching of these films opens real opportunities for their use as high resolution photoresists.

1. INTRODUCTION

Light induced changes in the properties of evaporated chalcogenide thin films have been studied for their potential use as inorganic photoresists\(^1\)\(^-\)\(^2\). The advantages of their use in fabricating integrated circuits (chemical and physical durability, broad spectral sensitivity, high resolution, edge sharpening and step-like function capabilities) have been described in recent publications\(^3\)\(^-\)\(^6\).

The technique of spin-coating offers the additional advantage that the layers can be deposited in the same way as conventional organic photoresists. The films deposited by this method have basically the same properties as those obtained by vacuum evaporation or sputtering techniques. Amorphous As-S films can be deposited in a wide range of thicknesses by spin coating, such films films are uniform and free of microstructure\(^7\). Their composition closely follows that of the bulk material. The annealed spin coated As-S film structure and optical properties have been investigated\(^8\)\(^-\)\(^9\).

Patterns can be generated by two different methods: photo-darkening of the As-S film and photo-dissolution of Ag into the As-S film, the former being a single step process and the latter a bi-layer resist process involving the deposition of a Ag layer, as well as the chalcogenide film. Both negative and positive type resists are available using the above methods. The properties of these resists were investigated using standard IC processing techniques.

2. EXPERIMENTAL METHODS

2.1. Sample preparation

Solutions of various strengths were made by dissolving the powdered As\(_2\)S\(_3\) glass in n-propylamine, the resulting film thickness (200Å - 2μm) was dependent on the concentration of the solution. Thin films were obtained by spinning these solutions onto substrates of either Si wafers or microscope slides. The spun films were annealed at 120°C for 1hr in order to eliminate the organic solvent from the film. For the silver photodissolution experiments, evaporated Ag films were used beneath the spin coated chalcogenide film as a silver source. A commercial wafer stepper (Optimetrix 8010 λ=430-600nm, power density=150mW.cm\(^{-2}\)) was used to illuminate the samples.

2.2. Plasma etching

Plasma etching experiments were carried out in an International Plasma Corporation 2000 series system, which is a basic barrel-type plasma etcher. The chamber pressure (atmospheric - 10\(^{-1}\) Torr) was the sole means of controlling the rate of flow of gas within the chamber. Etching was carried out with different plasma strengths (input powers) and barrel pressure combinations for each gas. The gases used were CF\(_4\)(g) for As-S material and S(g) for Ag:As-S material. Sulphur vapour was generated in vacuo from a solid source and sublimed before entry into the chamber.

An "in situ" optical technique was used to evaluate the etching rates using a He-Ne laser. The light is reflected from both the front (receding) and the back
(stationary) interfaces of the film, and the reflected light intensity is detected by a photodiode and plotted on a chart recorder.

3. RESULTS AND DISCUSSION

3.1. The single layer system: Photodarkening

It was observed that illumination caused photodarkening of the annealed spin coated films (associated with structural changes). This results in differences in the chemical etch rate, therefore the effect can be used for pattern generation in the film. A single As$_2$S$_3$ layer (thickness ~0.5μm) was deposited by spin coating and annealed at 120°C for 1 hour. Half of the layer was then illuminated for 10 min. The exposure time used here was for comparative work and is not a measure of sensitivity. A plasma of CF$_4$ (g) was used to etch differentially the illuminated and unilluminated areas of the film. The optimum etching conditions were: CF$_4$ gas pressure 0.7 Torr, r.f. power 100W. The progress of the etching was monitored continuously by the in-situ technique described above. The recorded interference patterns are shown in Fig. 1.

![FIGURE 1](image1.png)

Reflected laser intensity curves as a function of etching time. (a) unilluminated (b) illuminated As$_2$S$_3$ film

The etching rates were calculated by the positions relative to time, of successive maxima and minima. The amount of film removed at any given time is; 
\[ d = M\lambda/2n \]
where M is an integer, $\lambda$ is the wavelength of laser light used (6328Å) and n is the refractive index of the film. The refractive indices were found to be 2.24 for the unilluminated and 2.36 for the illuminated spin coated As$_2$S$_3$ (as determined by optical measurements). The etching time required to remove the illuminated film is significantly shorter than that required for the unilluminated areas. The calculated etch rate for illuminated spin coated As$_2$S$_3$ film is ~230 Å/sec, and for the unilluminated film is ~92 Å/sec. The results obtained for the etching kinetics are shown in Fig. 2.

![FIGURE 2](image2.png)

Etching kinetics for (a) unilluminated (b) illuminated As$_2$S$_3$ film

The etch rate is constant i.e. the plots are linear for both exposed and unexposed areas. The etch rate of the illuminated film is 2.5 times larger than that of the unilluminated areas. A differential etch rate of this magnitude makes spin coated As$_2$S$_3$ films suitable for pattern delineation, and thus use as a positive photoresist with all the advantages this material offers.

3.2 The bilayer system: Photo-dissolution of Ag into As-S

As demonstrated in the previous section As$_2$S$_3$ film can be etched in CF$_4$ plasma but no suitable etchant has been reported for the removal of the Ag doped material, this has restricted the use of Ag-As-S system as negative photoresist. A plasma of S(g) was found to be very effective and selective to the doped material, yielding an amphoteric resist system i.e either negative
or positive depending on the etchant used. The etching was monitored in the same way as described above. The optimum etching conditions were: Si(g) pressure = 0.65 Torr, r.f. power = 10W, and an etch rate of ~160 Å/sec was observed. Fig. 3 shows the experimentally obtained interference pattern and Fig. 4 shows the etching kinetics for the Ag-As-S bilayer system.

![Fig. 3](image)

**FIGURE 3**
Reflected laser intensity curve as a function of etching time for Ag-As-S film.

![Fig. 4](image)

**FIGURE 4**
The information obtained from Fig. 3 in the form of thickness etched (%) vs. time.

In this case the samples were used to obtain optical resolutions of 0.5μm, which is the limit of the wafer stepper used. The samples used were Si wafers with a vacuum evaporated Ag layer of 1000Å and a spun As-S film of 300Å. These very thin films have the same etching characteristics as those shown in Figs. 3-4, but they are too thin to give information as regards etch rates and thus vacuum evaporated samples were used to obtain this data.

4. CONCLUSIONS

Both photodarkening and silver-photodiffusion effects in spin-coated films of As₂S₃ can be used for pattern generation. The photodarkening effect offers a factor of 2.5 in the etching rate difference for illuminated and unilluminated films. This affords a positive type photoresist involving only a single layer deposition. The Ag-As-S double layer system has the drawback of having two deposition steps, but can act as either a negative or positive photoresist with all the advantages described.

REFERENCES

INTERFERENCE GRATING FABRICATION IN SPIN-COATED As$_2$S$_3$ FILMS

E. HAJTO AND P. J. S. EWEN
Department of Electrical Engineering, University of Edinburgh, Edinburgh EH9 3JL (U.K.)
R. E. BELFORD
Department of Physics, Napier College, Colinton Road, Edinburgh EH10 5DT (U.K.)
A. E. OWEN
Department of Electrical Engineering, University of Edinburgh, Edinburgh EH9 3JL (U.K.)
(Received August 29, 1990; revised November 16, 1990; accepted December 5, 1990)

The optical properties and solubility of spin-coated As$_2$S$_3$ films significantly change after silver is introduced by photodiffusion. The different types of holographic grating that can be fabricated in spin-coated Ag/As$_2$S$_3$ double film structures by photodiffusion and subsequent selective etching are described in this paper. The silver diffusion occurs under the influence of visible monochromatic light and the grating pattern is produced by a holographic arrangement.

1. INTRODUCTION

Light-induced changes in the physical and chemical properties of evaporated As—S and other chalcogenide thin films have been widely studied.$^{1-3}$ As—S films can also be deposited by the spin coating technique and such films have basically the same properties as those obtained by vacuum evaporation.$^{4-5}$ Amorphous As—S films can be deposited in a wide range of thicknesses by this technique and the resulting films are uniform.$^6$ Their composition closely follows that of the bulk material dissolved for spin coating and the structure and optical properties of the annealed, spin-coated As—S films are found to be very similar to those of the evaporated material.$^7$ Patterns can be generated in As—S films either by photo-darkening, which produces a difference in etch rate between the illuminated and unilluminated part of the film,$^8$ or by light-enhanced silver diffusion into the As—S film$^9$, which results in an even larger difference in the etch rate between the illuminated (i.e. silver-doped) and unilluminated film.

Since the chalcogenide glasses are also well-known IR transmitting materials, patterning based on these photo-induced effects can be used to produce diffractive optical elements for use in the IR region.$^{10}$ Such elements have potential applications in beam combining, filtering and spectral analysis and have advantages over conventional refractive components as regards weight, cost and ease of manufacture. Spin coating is a common technique for depositing materials, such as dichromated gelatin, which are used to produce diffractive optical elements for visible operation, and hence it would be useful to determine whether gratings can be created in chalcogenide films deposited by this technique.
In this paper the light-induced diffusion of silver into spin-coated As$_2$S$_3$ films is studied. It is shown that, using a holographic illumination method and subsequent selective etching, a surface relief grating of submicron resolution can be obtained. It is also demonstrated that the optical constants (absorption coefficient $\alpha$ and refractive index $n$) of the silver-doped As$_2$S$_3$ films differ considerably from those of the undoped films, so that the changes can be used to record either amplitude or phase holograms in the Ag/As$_2$S$_3$ double film structures.

2. EXPERIMENTAL DETAILS

The Ag/As$_2$S$_3$ double film structures are prepared in two steps. First a thin (about 0.3 $\mu$m) silver layer is deposited by vacuum evaporation onto a glass substrate; then an amorphous As$_2$S$_3$ layer (about 1 $\mu$m) is deposited using the spin coating technique$^{11}$. For spin coating, solutions of As$_2$S$_3$ were made by dissolving the powdered bulk As$_2$S$_3$ glass in propylamine. Thin films were obtained by spinning these solutions onto glass substrates previously covered with a silver film. Typically 2 g of As$_2$S$_3$ dissolved in 10 ml $n$-propylamine and spin deposited at 3000 rev min$^{-1}$ for 20 s, produced an As$_2$S$_3$ film of about 1 $\mu$m thickness. The samples were annealed (in the dark) at 120°C for 30 min in order to eliminate the remains of the organic solvent from the As$_2$S$_3$ film. The optical properties of the films were measured using a UV–visible–near-IR spectrophotometer (Perkin Elmer Lambda 9). The optical constants $n$ and $\alpha$ were calculated from the transmittance and the reflectance data using the method suggested by Swanepoel$^{12}$ and Abeles$^{13}$. For these measurements the samples were deposited onto Corning glass 7059 substrates with refractive index $n_{\text{substrate}} = 1.5049$. The effect of silver doping was also studied. For this purpose the Ag/As$_2$S$_3$ structures were illuminated from the As$_2$S$_3$ side (i.e. from the top of the double structure) with UV light using a Carl Suss mask aligner system. As a result of illumination, the silver layer completely diffused into the As$_2$S$_3$ film. The optical constants of the silver-doped As$_2$S$_3$ films were then measured.

The structure of the thin films and the principal steps in the formation of the surface relief grating are shown in Fig. 1. The Ag/As$_2$S$_3$ double structure is illuminated with the required pattern and the silver layer diffuses into the As$_2$S$_3$ layer in the illuminated areas. After illumination, methanol saturated with ammonia is used for developing the surface relief grating.

The optical arrangement for pattern generation is shown in Fig. 2. In this experiment we used the simplest form of interference grating in which the

Fig. 1. Sample configuration used for holographic imaging.
photosensitive films are exposed to a sinusoidal fringe pattern generated by two interfering beams of light. A continuous-wave argon ion laser is used for illumination, operating at a single wavelength of 5145 Å. The laser beam is focused first through a spatial filter, and then expanded and collimated. The beam is divided into two parts of the same intensity using a prism, which is a more stable arrangement than the usual holographic method using a beam splitter and mirrors, because the optical path length is shorter and the number of optical elements is fewer. A stable optical arrangement is important in this case because of the relatively long exposure time required for the silver diffusion. When the two coherent beams of light intersect at an angle $2\theta$, they will generate interference fringes within the volume common to both beams with a spacing (also termed grating period) $A$ given by

$$A = \frac{\lambda}{2} \sin \theta$$

where $\lambda$ is the wavelength of the light and $\theta$ is the angle of incidence. Therefore the grating period in the Ag/As$_2$S$_3$ structure is determined by the wavelength of the illuminating light and the angle of incidence $\theta$ which is in turn determined by the angle of the prism. The finest spacing that can be obtained is $\lambda/2$ which corresponds to $\theta = 90^\circ$. This cannot be achieved in practice since it requires both beams to be incident along the surface of the sample, but a value of $\theta = 60^\circ$ yields a spacing of $0.6\lambda$. In our experiments two different prisms were used: prism (1) provided an angle of incidence $\theta = 20^\circ$ and prism (2) provided an angle of incidence $\theta = 0.7^\circ$. The performance of the gratings produced (resolution, spectral purity, efficiency) was assessed using a low intensity He–Ne laser.

3. RESULTS AND DISCUSSION

Curve A of Fig. 3 shows the absorption coefficient of the undoped, spin-coated As$_2$S$_3$ films (i.e. without silver) as a function of a wavelength in the visible range of
the optical spectrum. Curve B of Fig. 3 shows the absorption coefficient for the silver-doped As$_2$S$_3$ films. The effect of silver doping on the refractive index is shown in Fig. 4 (curve A for the undoped As$_2$S$_3$ film, and curve B for the silver doped film). It can be seen that the absorption coefficient and the refractive index increase as a result of the silver incorporation into the As$_2$S$_3$ network. These results correspond

Fig. 3. Wavelength dependence of the absorption coefficient of spin-coated As$_2$S$_3$ films (curve A) and of silver-photodoped As$_2$S$_3$ films (curve B).

Fig. 4. Wavelength dependence of the refractive index of spin-coated As$_2$S$_3$ films (curve A) and of silver-photodoped As$_2$S$_3$ films (curve B).
to a silver concentration of about 22 at.\% in the As$_2$S$_3$ films as calculated from the initial thickness of the silver and As$_2$S$_3$ films.

The large changes observed in the optical constants of the silver-rich (i.e. the illuminated) parts of the films are accompanied by a drastic change in other physicochemical properties, for example their solubility in a suitable solvent such as methanol saturated with ammonia. This agent dissolves the non-illuminated part of the As$_2$S$_3$ film but the silver-doped part becomes practically insoluble. Figure 5 is a scanning electron micrograph of the surface relief grating produced in a spin-coated As$_2$S$_3$ film by light-enhanced silver diffusion and subsequent selective etching using the above solvent. In this experiment the sample was illuminated with the holographic arrangement of Fig. 2 using prism (1), i.e. the intersecting beams reach the surface of the sample with an angle of incidence $\theta = 20^\circ$. According to eqn. (1) this will produce a standing wave with the distance $\Lambda$ between the neighbouring interference maxima (the grating period) of about 0.75 \mu m. The electron micrograph in Fig. 5 shows that the standing wave pattern is reproduced as a surface relief grating after selective etching. The applied laser power was 13.5 mW and the illuminated area was 1.84 cm$^2$, resulting in a laser power density of 7.34 x 10$^{-3}$ W cm$^{-2}$. The duration of the illumination was 10 min, i.e. an illumination energy of 4.4 J cm$^{-2}$ was obtained. This illumination energy caused the silver to diffuse to a depth of about 0.6 \mu m, as can be estimated from the surface profile shown in Fig. 6 which is an electron micrograph of the grating taken in a tilted position with an angle close to 90°. Since the groove profile is basically a sine wave, it can be concluded that the silver diffusion depth (which determines the profile of the surface relief grating) is governed primarily by the light intensity profile.

The period of the recorded fringes can be increased significantly if a small angle of incidence is used for illumination with the same wavelength (see eqn. (1)). This is achieved by using prism (2) which produces an angle of incidence $\theta = 0.7^\circ$ for the two intersecting beams. This resulted in a grating period $\Lambda \approx 22 \mu m$ as seen in Fig. 7.

![Fig. 5. Scanning electron micrograph of the interference grating produced in spin-coated As$_2$S$_3$ by silver photodiffusion.](image)
The surface relief gratings prepared in this way are considered as thin holographic gratings. The distinction between thick and thin holographic gratings is usually made with the aid of the $Q$ parameter defined as

$$Q = \frac{2\pi \lambda d}{n\Lambda^2}$$

where $\lambda$ is the illuminating wavelength, $n$ is the refractive index, $d$ is the thickness and $\Lambda$ is the spacing of the recorded fringes. The holographic grating is considered thick when $Q \geq 10$ and thin otherwise. In our case a value $Q \approx 1.4$ can be obtained using the experimentally determined parameters of $\lambda = 0.6 \mu m$, $n = 3.0$ (see Fig. 4), $d = 0.56 \mu m$ (see Fig. 6) and $\Lambda = 0.7 \mu m$ (see Fig. 6).

Holographic gratings are also classified by the mechanism by which the illuminating light is diffracted. In the amplitude grating, the interference pattern is
As$_2$S$_3$ interference gratings

recorded as a density variation of the recording medium and the amplitude of the illuminating wave is modulated. In the phase grating a phase modulation occurs when the illuminating wave passes through the film. In our case it is primarily the phase of the illuminating wave which is modulated and diffracted from the surface relief grating. The change $\Delta \phi$ in the phase is given by

$$\Delta \phi = \frac{2\pi}{\lambda} \{ d \Delta n + (n - 1) \Delta d \}$$

where $\Delta n$ is the difference in the refractive index, $\Delta d$ is the difference in the thickness and $\lambda$ is the illuminating wavelength. If the hologram is physically thin, $d$ is very small and the contribution to $\Delta \phi$ from the term $d \Delta n$ is negligible so that

$$\Delta \phi = \frac{2\pi}{\lambda} (n - 1) \Delta d$$

suggesting that the surface relief modulates the light by the differences in the thickness of the film.$^{15}$

For an ideal sinusoidal grating one would expect the light to be diffracted only into the angles determined by the grating equation$^{14}$

$$\Lambda (\sin \alpha + \sin \beta) = m \lambda$$

where $\Lambda$ is the grating period, $\alpha$ and $\beta$ are the angles of incidence and diffraction, $m$ is the order number and $\lambda$ is the wavelength of the light used. Therefore the grating period $\Lambda$ can be determined from the measured angle of the diffracted beam. It was found that if the measuring light ($\lambda = 0.6328 \mu m$) is used at normal incidence relative to the surface relief grating, the diffracted beams exit the surface at angles $\beta = \pm 56.4^\circ$. This gives $\Lambda = 0.76 \mu m$ for the period of the grating, in good agreement with the micrographs of the surface relief grating (seen in Figs. 5 and 6).

The efficiency of a grating is defined as the fraction of the incident radiation that is diffracted into the required order. Figure 8 shows the measured efficiency of the surface relief grating using an He–Ne laser ($\lambda = 0.6328 \mu m$) at normal incidence. These measurements were obtained on the same surface relief grating but used in reflection mode. (For these measurements the sample was covered with a thin aluminium layer in the order to increase the reflectivity.) The efficiency is determined

![Fig. 8. Diffraction efficiency of the grating with 0.7 $\mu$m period as measured with 6328 Å light at normal incidence.](image)

primarily by the groove profile and if the grating is illuminated at normal incidence it cannot be higher than 50\% in any order because of the symmetry. In practice the efficiency seldom exceeds 33\% in this configuration\textsuperscript{14}. Figure 8 shows that approximately 18\% efficiency is measured at the diffracted orders of ±1 (at angles $\beta = \pm 56.4^\circ$), suggesting that the profile of the surface relief grating is close to the "ideal" sinusoidal profile. Similar diffraction efficiencies were obtained in a vacuum-evaporated chalcogenide–silver double-layer system\textsuperscript{16}. The roughness of the surface (see Fig. 5) might introduce a random variation in phase and amplitude of the diffracted light which might generate diffuse scattering. However, according to the theoretical analysis\textsuperscript{17} the spectral image of the diffracted light does not change significantly; the efficiency may decrease slightly but the distribution of light among the diffracted orders is much the same. It should also be emphasized that we did not observe spurious diffracted orders or diffuse scatter which might arise from periodic errors or imperfections across the surface of the grating.

We suggest that the large variation observed in the optical constants of the As$_2$S$_3$ films as a result of light-enhanced silver diffusion could also be used for recording thin amplitude or phase holograms. In this process no selective etching is necessary because the hologram can be recorded as changes in the absorption coefficient or the refractive index, which are in turn determined by the amplitude and the phase of the illuminating beam. It can be seen from Fig. 3 that silver doping increases the absorption coefficient (measured at 0.6 $\mu$m) from $\alpha = 1.35 \times 10^3$ cm$^{-1}$ to $\alpha = 4.67 \times 10^3$ cm$^{-1}$, which can be used for recording an amplitude hologram. In addition, the refractive index (measured at 0.6 $\mu$m) increases from $n = 2.22$ to $n = 3$ (see Fig. 4) as a result of light-induced silver diffusion. This effect can be used to produce a phase shift of the illuminating wave, i.e. to record a phase hologram. In our case these two effects occur simultaneously, indicating that the amplitude transmittance of such a hologram will be a complex function that describes the change in the amplitude and the phase of an illuminating wave on transmission through the material.

4. CONCLUSIONS

The optical constants and solubility of spin-coated As$_2$S$_3$ films significantly change after silver is incorporated by photodiffusion and these changes can be used to record grating patterns in the films. The silver diffusion depth, which determines the profile of the surface relief grating, is governed primarily by the light intensity profile. The recorded pattern can be developed using selective etching or, alternatively, the observed changes in the optical constants as a result of silver photodiffusion can enable the material to modulate the amplitude and the phase of the transmitted light, thereby facilitating the recording of amplitude or phase holograms.

ACKNOWLEDGMENT

E. H. is grateful to the U.K. Science and Engineering Research Council and Pilkington Brothers plc. for the provision of a CASE studentship.
AS$_3$S$_3$ INTERFERENCE GRATINGS

REFERENCES

Analogue memory and ballistic electron effects in metal–amorphous silicon structures

By J. HAJTO, A. E. OWEN and A. J. SNELL
Department of Electrical Engineering, University of Edinburgh, Edinburgh EH9 3JL, Scotland

and P. G. LE COMBER and M. J. ROSE
Department of Applied Physics and Electronic and Manufacturing Engineering, University of Dundee, Dundee, DD1 4HN, Scotland

[Received 11 June 1990 and accepted 15 June 1990]

ABSTRACT

We present experimental results showing that p⁺ amorphous silicon memory structures exhibit polarity-dependent analogue memory switching. The effect is non-volatile and we propose that it is associated with changes in a tunnelling barrier within the structure. It is also observed that conduction in the memory ON state is restricted to a narrow conducting channel through which the electrons can, under certain conditions, travel ballistically. As a consequence, quantized resistance levels associated with ballistic electron transport are observed under certain circumstances. In the presence of a magnetic field, additional steps in the quantized resistance levels occur. A particular feature of this quantized resistance is that the effect can be observed at relatively high temperatures (up to about 190 K).

§1. INTRODUCTION

Results on metal–p⁺–n–i–metal amorphous silicon devices have provided experimental evidence that such structures exhibit extremely fast non-volatile polarity-dependent digital memory switching phenomena (Le Comber et al. 1985) after initial conditioning by means of a moderately high applied potential (‘forming’). The most important result to emerge on formed devices is that in the ON state the current is carried by a highly conducting filament which is less than 1 μm in diameter. Filamentation has been demonstrated by experiments on the ON state resistance as a function of area (R(on) is independent of area), by thermal imaging techniques with liquid crystals and by direct observation with a scanning electron microscope combined with microanalysis. The latter indicated that the formation of the current filament may be associated with diffusion of the top metal contact into the amorphous silicon resulting in a region of intermingled metal and silicon. More recent experimental results (Rose et al. 1989) have demonstrated a new metal–p⁺–metal amorphous silicon device which, rather than exhibiting a two-state digital operation, has a continuum of stable states which are non-volatile and fully programmable by single 10 ns voltage pulses. It has also been suggested that the new analogue memory devices can be used as non-volatile and reprogrammable memory elements in analogue neural networks (Hajto, Rose, Snell, Le Comber and Owen 1990).

In this paper we present a summary of the new results obtained on non-volatile analogue switching effects in amorphous silicon metal–p⁺–metal devices and discuss the possible physical mechanisms responsible for the phenomena.
§ 2. EXPERIMENTAL

The samples used for this work were amorphous silicon Cr–p+–V thin-film structures. The p+ layer was prepared by r.f. glow-discharge decomposition of SiH4 containing 104 vol. p.p.m. of B2H6. Films of 1000 Å thickness were deposited on Corning glass substrates previously patterned with chromium bottom contacts. The p+ amorphous silicon was then patterned and an insulating layer was used to define an active device area of 10−6 cm². The metal used for the top contact was normally vanadium. However, a number of different metals were also used and their influence on the memory operation will also be described.

In accordance with our previous results (Le Comber et al. 1985), memory devices prepared in this way require an initial forming process. This means that the resistance of the as-deposited (unformed) device has to be lowered from \( R \approx 10^9 \) Ω to \( R \approx 10^3–10^6 \) Ω (i.e. the typical value of an ON state). The forming can be achieved by biasing the sample with a single voltage pulse (duration 300 us; magnitude about 12 V) with positive polarity applied to the top vanadium contact.

The formed samples were mounted in a 24-pin chip carrier and wire bonded. The carrier was then inserted into a standard socket in an Oxford Instruments cryostat in order to investigate the low-temperature characteristics. The current–voltage characteristics were measured using an HP 4145B semiconductor parameter analyser.

§ 3. RESULTS

The metal–p+–metal memory structures exhibit a forming step which is different from the previously investigated metal–p+-n–i–metal structures (Le Comber et al. 1985). The differences are demonstrated in figs. 1 and 2. In the case of metal–p+-n–i–metal structures, the resistance suddenly drops from about \( 10^{12} \) Ω (virgin state) to about \( 10^6 \) Ω after the critical voltage (forming voltage \( V_F \)) has been applied. No change in the virgin resistance occurs when the sample is biased with voltages less than \( V_F \). This type of forming is termed hard forming. In contrast with this, the resistance of the unformed metal–p+-metal structures can be lowered gradually by applying voltage levels with progressively increasing magnitudes. In this case no sudden change in the current or voltage signal can be detected when the sample is biased with a voltage pulse, as seen in fig. 2. This process is called soft forming. Figure 3 shows the device resistance as a function of the soft-forming voltage (pulse duration, 300 ns). On reaching a critical voltage (about 14 V in fig. 3) the device resistance suddenly drops from about \( 10^6 \) to about \( 10^3–10^4 \) Ω. This is the memory ON state of the device. Once the memory device had reached its first (non-volatile) ON state all subsequent switching operations were performed with 10–100 ns pulses 1–5 V in magnitude.

An example of the analogue switching effect is shown in fig. 4, where the sample resistance is plotted as a function of applied alternating WRITE and ERASE pulses (100 ns pulse duration each). It is important to emphasize the polarity dependence of the analogue memory behaviour. In the case of the WRITE pulses, positive polarity is applied to the chromium (bottom contact) while ERASE pulses have opposite polarity. The sample was first switched to an ON state (\( R_{ON} = 2 \times 10^5 \) Ω) and then a series of alternating WRITE and ERASE pulses were applied. The WRITE pulses were kept at a constant magnitude of 3–4 V but the ERASE pulses were incremented by 0.05 V steps from 1.2 to 3.4 V after each WRITE pulse. It can be seen from fig. 4 that the sample resistance changes in an analogue manner as the magnitude of the ERASE pulse increases, that is the difference between \( R_{ON} \) and \( R_{OFF} \) is a function of the magnitude of the ERASE pulses. A voltage range \( \Delta V \) (ERASE) of 1.6 V resulted in a change in...
Waveforms of a single forming pulse applied to an amorphous silicon Cr–p⁺–n–i–Al structure showing 'hard forming'. $V$ and $I$ represent the voltage across the device and the corresponding device current respectively.

Waveforms of single pulses of increasing magnitude applied to an amorphous silicon Cr–p⁺–V structure showing 'soft forming'. $V$ and $I$ represent the voltage across the device and the corresponding device current respectively.
Fig. 3

Resistance of an amorphous silicon Cr-p-V memory structure as a function of forming voltage.

Fig. 4

Memory resistance as a function of ERASE voltage in a Cr-p-V structure.

Resistance from \( R \approx 2 \times 10^3 \Omega \) to \( R \approx 6 \times 10^5 \Omega \). The shaded area in fig. 4 indicates the reproducibility of the analogue memory switching by showing the scattering in the resistance during repeated experiments (data from 100 cycles are included). Figure 5 shows another case where the ERASE pulses were maintained at a constant value of \( V = 3.4 \text{ V} \) but the WRITE pulses were incremented from 1.2 V to 3.4 V in 0.05 V steps. The value of the OFF state resistance remained constant at about \( 6 \times 10^3 \Omega \) (i.e. it changed back to this constant level from every ON state) whilst the ON state resistance decreased through a continuum of intermediate states over a similar \( \Delta V \) to the ERASE operation. The shaded area in fig. 5 again represents the reproducibility of the analogue switching for 100 complete cycles. It is emphasized that the device will switch between any two resistance states within the range from about 1 k\( \Omega \) to 1 M\( \Omega \) by selecting the correct polarity and the magnitude of the WRITE and ERASE pulses. For all devices with a vanadium top contact, the values of \( \Delta V \) range from 1.5 to 2.0 V for both the WRITE and the ERASE operations.

We have repeated the above experiments on metal-p-metal devices with chromium as top metal and observed similar polarity-dependent changes in the memory state resistance. However, in the WRITE and ERASE experiments, intermediate states were found to exist only over a narrow \( \Delta V \) of about 0.2 V as shown in fig. 6. Therefore these devices are considered as 'digital' devices. It is important to
emphasize that both the analogue (seen in figs. 4 and 5) and the 'digital' (seen in fig. 6) memory switching effects are non-volatile. Devices set to ON or OFF states have been monitored over 2 years without any significant change in their resistance. Also, operation at temperatures up to 160°C shows little change in the threshold voltages or in the device stability. However, it is also found that devices with certain top metal contacts such as molybdenum and palladium show a volatile memory switching effect. This is illustrated in fig. 7 where the signal through the device is continuously monitored at a low voltage level (at 0.5 V) that is below the voltage level of the programming pulses. The current decays rapidly after the end of the programming pulse, that is the memory state is volatile. The role of the top metal contact has been investigated by fabricating devices with a range of different top metals but with otherwise identical physical parameters (i.e. about 1000 Å thickness of the p⁺ layer, and a chromium bottom electrode). Using the analogue switching voltage window $\Delta V$ as a guide, it is found that its value is significantly dependent on the top metallization contact. This is illustrated in the table. It can also be seen that the definition of 'analogue' ($\Delta V \geq 1 V$) or 'digital' (0.5 V or less) memory switching is somewhat arbitrary, because there is no sharp boundary between the two types of operation, and they are almost certainly associated with the same underlying physical phenomena. However, in the cases of molybdenum and palladium top contacts, a new type of volatile
Volatile memory effect in a Cr–p⁺–Mo structure.

Effect of top metallization on the switching behaviour.

<table>
<thead>
<tr>
<th>Metal</th>
<th>$\Delta V$ (V)</th>
<th>Switching characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ag</td>
<td>&gt;0.1</td>
<td>Digital, non-volatile</td>
</tr>
<tr>
<td>Al</td>
<td>&gt;0.1</td>
<td>Digital, non-volatile</td>
</tr>
<tr>
<td>Cr</td>
<td>0.2</td>
<td>Digital, non-volatile</td>
</tr>
<tr>
<td>Mn</td>
<td>≈0.5</td>
<td>Digital, non-volatile</td>
</tr>
<tr>
<td>Fe</td>
<td>≈0.5</td>
<td>Digital, non-volatile</td>
</tr>
<tr>
<td>Ti</td>
<td></td>
<td>Unstable switching</td>
</tr>
<tr>
<td>Au</td>
<td></td>
<td>No switching</td>
</tr>
<tr>
<td>Cu</td>
<td></td>
<td>No switching</td>
</tr>
<tr>
<td>W</td>
<td>≈1.0</td>
<td>Analogue, non-volatile</td>
</tr>
<tr>
<td>V</td>
<td>1.8</td>
<td>Analogue, non-volatile</td>
</tr>
<tr>
<td>Ni</td>
<td>2.0</td>
<td>Analogue, non-volatile</td>
</tr>
<tr>
<td>Co</td>
<td>2.0</td>
<td>Analogue, non-volatile</td>
</tr>
<tr>
<td>Mo</td>
<td>2.0</td>
<td>Analogue, volatile</td>
</tr>
<tr>
<td>Pd</td>
<td>2.0</td>
<td>Analogue, volatile</td>
</tr>
</tbody>
</table>

switching effect is observed and in the cases of titanium, gold and copper no reproducible switching effects can be observed. These results suggest that the top metal contact plays a crucial role in determining the type of memory switching phenomena observed in these devices. In this paper we shall concentrate on the p⁺ memory devices with vanadium top contact because these show typical non-volatile analogue memory switching.

The current–voltage characteristics of analogue memory resistance states have been systematically investigated at both room temperature and lower temperatures and the following results obtained. It is found that all the room-temperature current–voltage characteristics show a 'linear-plus-power-law' behaviour in the various analogue memory resistance states (seen in fig. 8). The $I$–$V$ curves can be described by a simple nonlinear relationship

$$ I = C_1 V + C_2 V^n, $$

(1)
where $C_1$ and $C_2$ are constants and the exponent $n$ increases with increasing low-bias (linear region) resistance according to the relationship $n = A + B \log R$. The observed power-law behaviour could indicate the possibility of space-charge-limited conduction at higher biases, although investigation of the thickness dependence has shown this to be unlikely. In accordance with the pulsed analogue memory switching results, a continuous transition of states can be found between the low-bias ON (about $10^3 \Omega$) and OFF (about $10^6 \Omega$) states. The terms ON and OFF seem to be somewhat arbitrary therefore, and are only used in this paper for practical reasons. If the curves in fig. 8 are extrapolated above 1 V, they meet in the region of about 3–4 V, that is at typical programming levels (see figs. 4 and 5). It is also found that the exponent $n$ does not depend on the thickness or the active device diameter but it is primarily determined by the low-bias (linear region) resistance of the analogue memory state. This indicates that the nature of the electrical conduction is quite similar in all memory states. However, repeatedly switching the device into the same resistance state need not always result in an identical value of the exponent. This appears to indicate that the 'same' resistance state can be achieved through different conduction paths within the same device. This is in accordance with the suggestion that the conduction path (filament) might have a structure similar to that of granular metals embedded in an insulating matrix and this structure may provide a variety of conduction paths (Giaever and Zeller 1968).

It is possible to extend the range of the $I$–$V$ characteristics (up to the critical field where switching occurs) using very short single voltage pulses of varying polarities and pulse heights. Figure 9 shows the room-temperature 'pulsed' current–voltage characteristics of the analogue Cr–p+–V memory device. These characteristics are obtained using 400 ns pulses of progressively increasing magnitude and of both polarities. The 400 ns pulse length is long enough to observe a plateau in the pulse signal, that is $RC$ effects are avoided. Positive polarity refers to (positive) voltages applied to the chromium bottom contact. Starting from a $9.5 \times 10^5 \Omega$ OFF state (measured at 0.5 V (fig. 9(a))), the onset of the strong nonlinear rise in the current occurs at about +1.1 V. Figure 9(a) is reproducible (i.e. it can be repeated many times without a change in the device resistance) up to a voltage level of about +1.7 V. Further increase in the voltage height will result in a permanent decrease in the device resistance and consequently a change in fig. 9(a). The decrease in resistance is determined by the magnitude of the maximum voltage applied.
Figure 9 (b) represents a memory ON state ($R = 5.6 \times 10^3 \Omega$ at 0.5 V) of the device and is reproducible up to a voltage level of about 5 V if positive polarity is applied. Further increase in voltage might destroy the device. However, if a negative voltage is applied to the same ON state (fig. 9(c)), an ERASE process is observed at voltage levels of magnitude greater than about 1.7 V. On the other hand, if an OFF state ($R = 9 \times 10^3 \Omega$) is negatively biased fig. 9(d), no change in the OFF state resistance can be observed up to a voltage of about 5 V. The apparent polarity dependence suggests that the analogue memory switching is not determined simply by the magnitude of the applied power or energy. This is further supported by the comparison of individual switching transients with opposite polarity. Figure 10(a) shows the waveform of a single WRITE pulse from OFF ($R = 2 \times 10^2 \Omega$) to ON ($R = 3.8 \times 10^3 \Omega$) and of an ERASE pulse from OFF ($R = 2 \times 10^3 \Omega$) to a slightly higher OFF state ($R = 2.4 \times 10^3 \Omega$). In the second case a large change in the memory state resistance does not occur although similar voltage and current levels are measured. The calculated total charge (flowing through the sample) is also rather similar: $Q_{w} = 8.5 \times 10^{-11}$ C for the WRITE and $Q_{E} = 7.2 \times 10^{-11}$ C for the ERASE pulse. The rather high level of 'ringing' oscillations on these transients was caused by poorly matched cable impedances.

These measurements suggest that the memory state resistance is determined by a combination of applied voltage level and the appropriate polarity. On the other hand, the memory resistance not being determined uniquely by the applied power or energy suggests that the memory switching does not depend significantly on the internal temperature of the device. Figure 11(a) shows a switching transient of an ERASE pulse at 300 K where the device resistance is changed from $R_{ON} = 3 \times 10^3 \Omega$ to $R_{OFF} = 6 \times 10^3 \Omega$. Figure 11(b) shows a similar ERASE transient at much lower temperature (4.2 K) where the device resistance has also changed from $R_{ON} = 3 \times 10^3 \Omega$ to $R_{OFF} = 6 \times 10^3 \Omega$. It can be seen that, despite the large temperature difference (more than two orders of magnitude), the current level and the threshold voltage (for achieving the same ERASE process as at room temperature) increase by a factor of less than two. It should also be emphasized that the device continues to operate even at liquid-helium temperature. The analogue switching effect is still observed at 4.2 K without large changes in either the threshold voltage or the current level. This suggests that the electrical conduction and the memory phenomena are possibly connected to a temperature-independent physical process which we propose may be associated with tunnelling.
Analogue memory in metal--a-Si structures

Fig. 10

Waveforms of single-pulse memory switching: (a) OFF→ON transient; (b) OFF→OFF transient.

Fig. 11

Waveforms of single pulse memory switching: (a) ERASE transient at 300 K; (b) ERASE transient at 4-2 K.

The above suggestion is supported by recent results obtained during investigation of the low-temperature conductivity of the analogue memory states (Hajto et al. 1990). Typical current--voltage characteristics of a formed memory ON state at 4-2 K are shown in fig. 12. In the voltage region from 0 to 0.36 V, the current around zero bias is of the order of $10^{-9}$ A but increases to about $10^{-6}$ A at voltages approaching 0.36 V, that is a strong nonlinear behaviour is found. It is important to emphasize that the room-temperature current--voltage characteristics of the memory ON state are linear. The observed large increase in the resistance around zero bias (fig. 13) is consistent with tunnelling conduction between metallic particles embedded in an insulating matrix (Giaver and Zeller 1968). Further experimental evidence for the tunnelling conduction comes from the temperature dependence of the device current. Figure 14 shows the current at a fixed voltage (0.2 V) plotted against $T^2$. The linear plot indicates a $T^2$ dependence of the current at a constant $V$ which is in accordance with tunnelling conduction.

At 0.36 V, a current jump occurs and the resistance of the sample is lowered to the order of a few kilohms. After the first current jump at 0.36 V, further current steps can be observed at 0.47, 0.53 and 0.70 V (fig. 12 (a)). At 0.36 V (henceforth called the critical voltage $V_c$) there is a dramatic change in the behaviour of the sample. At voltages lower than $V_c$, no discontinuities are observed but, at voltages higher than $V_c$, the resistance
Fig. 12

Current–voltage characteristics at 4·2 K (B) with and (A) without magnetic field. (B) is displaced vertically for clarity.

Fig. 13

Resistance as a function of applied bias at 4·2 K.

Fig. 14

Current (at fixed voltage of 0·2 V) as a function of $T^2$. 
Analogue memory in metal–a-Si structures

is lowered and the current increases in discrete steps. The current–voltage characteristics are symmetrical, that is the same behaviour is observed for the opposite polarity. Figure 12(B) depicts the current–voltage characteristics of the same sample under the influence of a 0.2 T magnetic field. The curve has been displaced by 50 \mu A in the current scale for clarity. The direction of the magnetic field is 30° with respect to the conducting channel (i.e. the filament). Additional steps can be observed at 0.34, 0.42, 0.5, 0.59 and 0.79 V together with the steps observed in the zero magnetic field case (fig. 12(B)). The effect of the magnetic field is reversible, that is, if the magnetic field is removed, the current–voltage characteristics revert to the zero-magnetic-field case. A number of \( I-V \) more characteristics have been obtained in which sharp and well defined steps can be observed at 4.2 K.

The critical voltage \( V_c \) at which the first current jump is observed and the magnitude of the first current jump are dependent on the resistance of the memory ON state investigated. Figure 15 shows the effect of changing the memory ON state resistance on the observed current steps. The curves have been displaced by 100 \mu A on the current scale for clarity. The critical voltage and the magnitude of the first current jump increase with increasing resistance (i.e. \( R, > R, > R, \), but after the first jump the characteristics are rather similar, suggesting a similar conduction mechanism at higher voltages. The first current jump appears to be associated with the formation of a highly conducting path within the structure whose characteristics are independent of the low-bias behaviour (i.e. the different memory states). On the other hand, the position of the current steps is dependent on the direction of the voltage sweep, that is some hysteresis is observed as illustrated in fig. 16.

Further information can be obtained if the resistance of a memory state (corresponding to fig. 12(A)) is plotted as a function of applied voltage as illustrated in fig. 17. We propose that the first large current step (at 0.36 V) is associated with the formation of a narrow, highly conducting channel which significantly lowers the resistance of the sample. With further increase in the applied voltage, the resistance is lowered in steps, corresponding to quantized resistance values \( R = h/2ie^2 \) where \( i \) is an integer. In the voltage range from 0.36 to 0.8 V, there are four steps (corresponding to the observed current rises in fig. 12(A)) with \( i \) being 2, 3, 4 and 5. Higher voltages have not been applied to the sample because this could change the resistance of the particular memory state. If a magnetic field is now applied to the sample, further quantization of resistance is observed at values \( R = h/2i(e^2) \). This is illustrated in fig. 18 (the data correspond to fig. 12(B)). It can be seen that extra steps in the resistance occur at \( i = 2, 5, 3, 3, 4, 5 \) and 5.5. The effect of temperature is illustrated in fig. 19. The curves have been shifted along the current axis for clarity. The observed current steps gradually decrease with increasing temperature until the effect is no longer observable at about 190 K. The results summarized in figs. 15–19 are indicative of ballistic electron transport (Hajto et al. 1990).

§4. DISCUSSION

In analysing the main results of this work, two important facts should be emphasized. Firstly the observation of ballistic electron transport provides a vital clue to the structure of the analogue memory element. Secondly, the programmability of the analogue memory provides information about the possible mechanism of the switching process itself.

The starting point of our discussion is that the analogue memory effect in amorphous silicon metal–p+–metal structures can only be observed if the sample is
Current–voltage characteristics of different memory states at 4·2 K. $R_c > R_p > R_A$ (curves displaced for clarity).

Current–voltage characteristics at 4·2 K showing the hysteresis observed on first increasing and then decreasing the voltage.

Resistance against voltage at 4·2 K with no magnetic field: (---) indicates the idealized contribution from the ballistic transport channel.
subjected to an initial forming process. The forming process is characterized not only by the breakdown of the high-resistance state of the structure but, more importantly, also by presence of a positive feedback mechanism which provides a low-resistance ON state so that the breakdown is non-destructive and repetitive switching is possible. Furthermore, the experimental results suggest a strong influence of the choice of the top metal contacts (summarized in the table) on the type of memory switching observed (i.e. digital or analogue switching) and on the success of obtaining stable and reproducible switching. These results are in good accordance with the previous observation (Gage et al. 1989) that the first switching event (i.e. the forming) causes a local structural modification of the $p^+$ amorphous silicon layer, producing a highly conducting filament which does not revert to the original amorphous material when the device is switched OFF. After the forming process the $p^+$ devices usually exhibit a lower OFF resistance than the unformed device in contrast with our original data for $p^+-n-i$ devices (Le Comber et al. 1985). The temperature dependence of the conductivity is also greatly reduced by the forming process. The area independence of $R_{ON}$ (Le Comber
suggests localized electrical conduction after forming. These results, together with the newly observed ballistic electron transport phenomena (see figs. 12–18), provide strong experimental evidence that the forming process creates a filamentary region consisting of a new material whose properties have changed significantly compared with those of the unformed original material. It is feasible that the high fields and current densities present during forming result in the development of high temperatures locally, which could lead to enhanced diffusion of metallic particles from the electrode into the thin amorphous film. Such a region would become the preferred current path carrying the electron current in the ON state. The current–voltage characteristics in the ON state at room temperature (see fig. 8) suggest that any material rearrangement within the filament occurs so as to destroy the rectifying properties of the original metal–p+ Schottky contact, and it should also be noted that this is the case even in a typical OFF state.

Lowering the temperature of the memory device reveals further information about the possible structure. The low-temperature current–voltage characteristics (see figs. 12 and 13) show that we have observed a zero-bias high-resistance anomaly and quantized resistance steps (associated with ballistic electron transport) in the ON state of amorphous silicon Cr–p+–V structures. The phenomenon of ballistic transport is observed in the case when the mean free path $\lambda$ of the electrons is larger than the length of the conducting channel. The usual approach to the fabrication of devices based on ballistic conduction is to use a very high mobility material (usually high-purity GaAs), where the mobility can reach values of the order of $10^{12}$ cm$^2$ V$^{-1}$ s$^{-1}$, leading to values of the electron mean free path of the order of 1 µm. The value of $\lambda \approx 1$ µm is certainly longer than the device dimensions which can be achieved by modern submicrometre technology. The structure in which ballistic transport is most widely investigated is a GaAs–Al$_x$Ga$_{1-x}$As heterojunction with a split-gate field-effect transistor configuration, usually less than 0.5 µm in length, with a gap of about 0.7 µm (Wharam et al. 1988, Van Wees et al. 1988). As the voltage on the gate is made increasingly negative, the depletion region increases, narrowing the effective gap and hence the width of the conducting channel decreases. As a consequence, the (one-dimensional) channel width becomes comparable with the electron wavelength, that is becomes sufficiently small that quantization occurs. On the other hand, the channel length is sufficiently short that electrons pass through ballistically (i.e. without appreciable scattering). Our structure is quite different and a possible mechanism which explains the ballistic transport in amorphous silicon structures is outlined below.

The current–voltage characteristics (see fig. 13) suggest that the filament has relatively large resistance around zero bias, that is a barrier for the current flow exists. The observed zero-bias anomaly is consistent with tunnelling conduction between metallic particles embedded in an insulating matrix (Giaver and Zeller 1968). The current–voltage characteristics from 0 to 0.36 V show a continuous, but strongly non-linear behaviour owing to the high electric field across the small tunnelling distance. In such a system the injection of free carriers is an activated process related to the increase in the electrostatic potential of a particle when a free electron is added to it. The activation energy can be provided entirely by thermal energy (hence the effect diminishes at higher temperatures) or, in the presence of an applied field, part or all of it can be provided by the field itself. The experimental observation of an approximate $T^2$ dependence of the current at a constant voltage observation of less than $V_m$ (see fig. 14) is also in agreement with the assumption that the electron transport is dominated by field-activated tunnelling processes in this region.
Analogue memory in metal-a-Si structures

These results can be partially explained by assuming that the conducting filament is composed of two parts: small-scale inclusions of permanently changed material connected by conducting channels which are formed, broken, dimensionally changed, re-formed, etc, during switching. Figure 20 illustrates an idealized model, having a single permanent inclusion extending from the top contact, with a narrow channel connecting it to the bottom contact. The evidence is (Le Comber et al. 1985) that the overall diameter of the filament at the top contact is less than 0.5 μm. The length of the channel must be consistent with tunnelling.

With increasing applied voltage, the tunnelling current increases exponentially. This implies that at the critical voltage $V_c$, the tunnelling barrier effectively 'breaks down' and a very large current flow occurs. This is not a destructive effect as the process is completely reversible and no material changes ensue. Because of the curved shape of the metallic inclusion illustrated in fig. 20, the current flow is restricted to a very localized area. Thus the channel can be considered to be an electron waveguide with confinement being brought about by the combination of the applied field and the geometry of the metallic inclusion. The observation of ballistic transport shown in figs. 15-19 can only be explained by assuming that the carrier transit time across the channel must be less than a scattering time, that is $\tau_\text{c} < \tau$. This leads to the condition

$$\mu^2 > \frac{ed^2}{m^*V},$$

where $m^*$ is the effective mass, $\mu$ is the mobility and $d$ is the channel length. With $d \approx 50 \, \text{Å}$ and $V \approx V_c = 0.4 \, \text{V}$ this leads to $\mu > 100 \, \text{cm}^2 \, \text{V}^{-1} \, \text{s}^{-1}$. The question remains, however, of whether it is feasible to have such a mobility in the material of the conducting channel. Assuming that the channel length is about $50 \, \text{Å}$ (typical tunnelling length), the potential gradient along the channel is about $0.36 \, \text{V}/50 \, \text{Å}$, that is $E \approx 7 \times 10^3 \, \text{V/cm}$. The low-field value of mobility in amorphous silicon is about $10 \, \text{cm}^2 \, \text{V}^{-1} \, \text{s}^{-1}$, but this might be increased by the presence of high local fields, when the more energetic carriers might have a larger mean free path because of the reduced scattering cross-section. There is also a possibility that, because of the presence of the metal, alloys could be formed, yielding still higher mobility values. In this connection it is worth noting that we have previously estimated a carrier mobility of $100 \, \text{cm}^2 \, \text{V}^{-1} \, \text{s}^{-1}$ in similar devices from magneto-resistance measurements (Le Comber et al. 1985). Therefore, although there is considerable uncertainty regarding the structure of the channel and the mobility value there, it looks quite certain that it is

Fig. 20

Schematic description of the filament showing the proposed metallic inclusion and the one-dimensional conducting channel.
the small channel length which makes ballistic conduction possible at much lower values of mobilities than in the previous 'conventional' cases (Wharam et al. 1988). It is important to emphasize, however, that in our case the electron mobility does not need to have the same very high value as in the case of GaAs—Al$_x$Ga$_{1-x}$As heterostructures (Wharam et al. 1988). In our case the one-dimensional conducting channel is established by the dramatic increase in the tunnelling current at voltage $V_a$. A further increase in voltage in the ballistic regime results in an increase in the cross-section of the conducting channel. This effect can be explained by assuming that the edge of the permanent inclusion has a curved shape as illustrated in fig. 20, that is the tunnelling barrier will break down at larger areas under the influence of higher voltages. The larger voltage applied will produce the same critical electric field at larger distances between the inclusion and the metal electrode; therefore the area of the one-dimensional channel increases. It should be noted that the voltage has an opposite effect on the cross-section of the conduction channel in the case of split-gate structures where the increased voltage level 'squeezes' the channel width. In our case the electrons are confined by the applied potential in the conducting channel whose cross-section increases with increased voltage levels.

This cross-section can be approximated by a rectangle (in a manner analogous to the 'electron-in-a-box' problem), whose dimensions $L_y$ and $L_z$ are comparable with an electron wavelength (fig. 20). The wavefunction $\psi$ and the energy of the electrons in the lowest quantum level of the one-dimensional channel are given by

$$\psi = L_y^{-1/2} \exp(ikx)g_y(y)g_z(z),$$

$$E = \frac{h^2}{2m^*} k^2 + \frac{h}{2m^*} \left[ \left( \frac{x}{L_y} \right)^2 + \left( \frac{z}{L_z} \right)^2 \right].$$

where $g_y$ and $g_z$ are the normalized standing-wave functions of the forms $2^{1/2} \cos(\pi y/L_y)$ and $2^{1/2} \cos(\pi z/L_z)$ respectively, $L_y$ is the length of the conducting channel; $L_y$ and $L_z$ are the cross-section dimensions of the channel (fig. 20). The first term in eqn. (4) describes the kinetic energy of a ballistic electron in motion along the channel. The second term represents the quantum energy level which is uniquely determined by the dimensions $L_y$ and $L_z$. It is important to emphasize that the electron can accelerate freely along the channel, owing to the collision-free nature of the motion, that is its momentum can increase continuously as the applied voltage increases. On the other hand the electron momentum is quantized in the direction across the channel (defined by $L_y$ and $L_z$). Therefore, as the voltage increases, two opposite effects can be observed; one is characterized by the increase in the electron kinetic energy along the channel and the other is associated with the increase in the cross-section of the conducting channel (i.e. with increase in $L_y$ and $L_z$). According to eqn. (4) this latter effect will decrease the quantized energy levels until the nearest higher energy level (termed sub-band (Landauer 1985)) becomes lower in energy than the Fermi level and will be occupied by the 'more energetic' electrons. Consequently a new sub-band is created at higher voltage (i.e. at larger values of $L_y$ and $L_z$) which in turn will lower the resistance in a quantized manner. This physical scenario was first suggested by Sharvin (1965) as a possible method for studying the Fermi surfaces of metals. It is thus a necessary consequence of the observation of quantized resistance values that the quasi-Fermi level must be within the sub-bands. This we believe to be a direct result of the extremely high current density in the channel, estimated to be of the order of $10^9$ A cm$^{-2}$. 
Analogue memory in metal-a-Si structures

According to the theoretical predictions (Landauer 1985, Imry 1986, Wharam et al. 1988), the resistance of a one-dimensional channel (in which the electrons punch through ballistically) should have a quantized value. This is because, in the absence of scattering, the change of electron kinetic energy is determined directly by the applied field across the one-dimensional channel. The current flow is given by

\[ I = ne\delta v, \]  

where \( n \) is half the number of carriers per unit length within one particular electron energy level (sub-band) and \( \delta v \) is the increase in the electron velocity acquired on transit through the one-dimensional channel. The number of carriers can be obtained by integrating the density of states in one dimension:

\[ n = \int_0^{E_F} N(E)\,dE = \frac{g_s}{2\pi\hbar} \left( \frac{m^*E_F}{2} \right)^{1/2} = \frac{g_s m^*v_F}{2\pi\hbar}, \]  

where \( E_F \) is the Fermi energy and \( v_F \) is the Fermi velocity. It is assumed that the applied voltage is sufficiently small that the velocity acquired by the electrons during the transit in the channel is small compared with the Fermi velocity. In this case the increase in velocity is given by

\[ \delta v = eV/m^*v_F. \]  

Substituting eqns. (6) and (7) into eqn. (5) yields

\[ I_s = \frac{g_s e^2V}{2\pi\hbar}, \]  

for the current and

\[ R_s = \frac{h}{g_s e^2} = \frac{h}{2e^2}, \]  

where \( R_s \) is the sub-band resistance and \( g_s \) is the spin degeneracy \((g_s = 2)\) in the non-magnetic field case. However, the value of \( R_s = h/2e^2 \) applies to only one sub-band. If a new sub-band is created at a higher voltage (i.e. by changing the cross-section of the channel), the total resistance becomes \( h/2ie^2 \) \((i \text{ is an integer defining the number of sub-bands below } E_F)\) in the same manner as in the case of parallel resistors. This is possible because there is no scattering in the conducting sub-bands, that is they do not interact with each other. Consequently the resistance drops in a quantized manner with \( i = 2, 3, 4 \) and \( 5 \) (as seen in fig. 17). In the case of our results, the plateaux of the individual quantum steps are not flat but they have a finite slope, that is the resistance decreases at higher voltages between the individual plateaux. This could be caused by the presence of additional non-ballistic conduction mechanisms. The broken line in fig. 17 indicates the resistance–voltage characteristics expected from the idealised ballistic component. We emphasize that the data presented in figs. 15–19 are raw data, and no attempt has been made in the present work to distinguish between the various contributions to the conductivity. In the above model we have considered the effect of moving the sub-bands through the quasi-Fermi level by changing the dimensions of the channel. In principle, we believe it would be equally possible to explain the data with a model in which the sub-bands are fixed in energy and the quasi-Fermi level is varied by the changing concentration of injected charge in the channel as a function of applied voltage.
When a magnetic field is applied, additional steps occur. This is probably due to the Zeeman splitting of electron energy levels, that is a lifting of the spin degeneracy splits each level into two. Therefore extra steps will be observed at $R = \hbar/2(i + j)e^2$ resistance levels as seen in fig. 18. The magnitude of the effect seen is rather surprising for a temperature of 4.2 K and a magnetic field of 0.2 T. However, this could be explained by assuming an anomalously high value for $g_e$.

The most important feature of the observed ballistic behaviour is that it can be observed up to about 190 K, namely, much higher temperatures than previously observed. We suggest that this might be due to the very small size of the ballistic channel.

The model suggested above describes the possible mechanism for the ballistic transport and provides important information about the structure of the filament, but, as it stands, it does not explain the memory switching. The most important difference between the two phenomena lies in the fact that ballistic transport is observed at applied voltage levels lower than about 1 V, whereas no memory switching occurs. The observed quantized jumps in the device resistance are of the type associated with threshold switching. If the voltage is lowered, the sample resistance reverts to the original low-bias case, that is there is no permanent change in the current–voltage characteristics of the device. On the other hand, memory switching occurs at voltage levels from about 1.5 V to about 4 V, resulting in a permanent change in the current–voltage characteristics of the device, that is a different resistance state of the memory.

In order to explain the different and permanent resistance states a model for memory switching should embody one or more of the following features:

1. changes in the conductivity of the whole filament;
2. local changes in conductivity at a certain point(s) in the filament;
3. changes in the filament geometry.

Model (a) would presumably require an average energy for switching of a similar order to that established for forming, but it has been found that ON→OFF→ON switching can be achieved using single pulse energies of about $10^{-16}$ J (see fig. 10). In addition, if it is assumed that the filament is homogeneous, there is no obvious way of introducing a polarity dependence. However, if there are inhomogeneities in the filament, it may be possible to increase substantially the local energy density and field strength. Consider for example, the simple one-dimensional arrangement shown in fig. 21. The regions A consist of a high-conductivity material separated by regions of lower-conductivity material B. The observed resistance of the filament will depend

Fig. 21

(a) Possible one-dimensional arrangement of the conducting filament; (b), (c), schematic energy diagrams for the tunnelling region.
upon the proportions of A and B and hence the relative contributions to the filament resistance of regions A and B. If \( R_A < R_B \), the properties of the type A material will be observed and, if \( R_A < R_B \), the properties of the type B material will be observed. If the 'gaps' between the highly conducting regions are small, then it is also necessary to consider the possibility of inter-island tunnelling. In order to establish such a model, a more detailed description of regions A and B must be included, for example, as follows.

1. A is metallic or a degenerately doped material whose contribution to the filament is fixed and \( \sigma_A \neq f(V) \) where \( V \) is the applied voltage and \( \sigma \) is the conductivity.

2. B is a 'gap' between type A material which is sufficiently small to allow significant tunnelling to occur; consequently \( \sigma_B \) is some function of the voltage.

One possible explanation is that, in the ON state, \( R_A < R_B \) and the measured conduction properties of the filament are those corresponding to material A, that is Ohmic current–voltage characteristics, and a small thermal activation. In the OFF state, \( R_A > R_B \) and the conduction properties are governed by the behaviour of the tunnelling regions. For the purpose of the model the tunnelling region can be represented in a simplified way, as shown in fig. 21. The current density in such a simplified system is governed by the following factors:

1. the applied bias \( V \);
2. the average tunnelling distance \( d \);
3. the average barrier height \( \Phi \) and asymmetry \( \Delta \Phi \);
4. the density of states \( g(E) \) in the metallic regions;
5. the occupation of \( g(E) \), that is the temperature dependence.

The functional form \( J(V, T) \) of the current density \( J \) against voltage \( V \) at different temperatures can be obtained under limiting conditions.

(a) \( V \ll \Phi \), and \( \Delta \Phi \ll \Phi \), constant \( T \). In this case the transmission probability \( P = c \exp(-d\Phi^{1/2}) \) may be treated as a constant for electrons near the Fermi energy \( E_F \) and provided that \( g(E) \) variations near \( E_F \) are small, \( J \) will increase linearly with \( V \) as more empty states become accessible in metallic region A. Conduction is then approximately Ohmic and symmetric (Giaver and Zeller 1968) in good accordance with fig. 8.

(b) \( V \ll \Phi \), and the temperature is varied. In this case, \( P \) is not very dependent on \( T \) as usually \( kT \approx 10^{-2} \text{ eV} \). Thus only a few electrons are promoted thermally to levels where \( P \) is significantly larger than at \( E_F \). The approximate form of the current density–voltage characteristics can be obtained as

\[
J(V, T) = J(V, 0) \left( 1 + \frac{3 \times 10^{-9} d^2 T^2}{\Phi} \right). \tag{10}
\]

with \( d \) in Ångström units, \( \Phi \) in electronvolts and \( T \) in kelvins (Giaver and Zeller 1968). The \( T^2 \) dependence of \( J \) at a constant \( V \) has been observed (see fig. 14) and the slope correlated with \( d = 50 \text{ Å} \) and \( \Phi = 1 \text{ eV} \).

How might switching be achieved in terms of the tunnelling model? The low-bias isothermal resistance of 'region-B-like' parts of the filament can be expressed as

\[
R = c \exp(d\Phi^{1/2}). \tag{11}
\]
where the constant $c$ is determined by geometric factors and the transmission probability. This expression gives $R$ directly in ohms if $d$ is in Ångström units and $\Phi$ is in electronvolts (with $c$ as a small correction factor). Taking $d = 50$ Å and $\Phi = 1$ eV, then $R_1 = c \exp (50)$ and increasing $d$ by 5 Å gives $R_2 = \exp (55)$, that is $R$ has increased by about a factor of 150. An equivalent increase in resistance is obtained by keeping $d$ constant and increasing $\Phi$ from 1 to 1.2 eV. Therefore dimensional changes in the tunnel barrier on an atomic scale, or small (20%) changes in barrier height, could account for the presence of switching. A change in the particle size and/or their spatial distribution as a result of localized heating or high field effects could explain the differences between high- and low-resistance states. The high switching speeds (about 10 ns) and the insensitivity of switching to temperature suggest that large-scale structural changes are unlikely. The observation of the quantized jumps in resistance and especially the magnetic field dependence of the quantized resistance levels at values of $R = h/(2(i + \frac{1}{4})e^2)$ provide an important experimental proof that the electron transport is ballistic in the formed Cr–p –V memory structures. The most important feature of the observed ballistic behaviour is that it can be observed up to about 190 K, namely much higher temperatures than previously observed (Wharam et al. 1988). This temperature dependence (see fig. 19) suggests that the states responsible for the quantized resistance levels are separated in energy by $AE \approx 1.6 \times 10^{-2}$ eV which is about $kT$, $T$ being the temperature at which the differences between the quantized electron levels are smeared out. This energy indicates a change of about 40 Å in the $L$ and $L'$ dimensions (as follows from eqn. (4)) between the subsequent steps which correspond to the energy separation of the observed quantum steps. Finally we emphasize that the effect has been observed in what is initially a metal–amorphous semiconductor structure. The ‘classical mobility’ of the amorphous silicon is many orders of magnitude lower than is expected for ballistic behaviour. It is not known with any certainty what effect the forming process has on the structure of the conducting channel but the observed behaviour suggests the importance of the very small dimensions of the conducting channel rather than other physical parameters. The significance of using amorphous silicon sandwich structures lies in the forming process which allows the fabrication of such small structures. The strong influence of the metal contact on the observed memory switching behaviour indicates the importance of ‘alloying’ during the forming process. It has been reported for example that solid-phase amorphization or glass formation occurs in V–Si reactions induced by rapid thermal annealing (i.e. at conditions similar to forming) but do not occur in Co–Si and Cr–Si (Nathan 1988). Therefore the presence of a very small tunnelling conduction path in amorphous silicon Cr–p –V structures might be due to a new type of solid-phase reaction (induced by forming) which creates a homogeneous (possibly amorphous) V–Si silicide.

§ 5. Conclusions

The experimental results presented in this paper show that Cr–p –V amorphous silicon structures exhibit polarity-dependent analogue memory switching. The analogue resistance values are stable after removing the programming voltages and up to temperatures of about 160°C, that is the effect is non-volatile. The temperature dependence of the current–voltage characteristics and of the memory switching transients suggest that the effect can be explained by either dimensional changes in a tunnelling barrier present in the structure on an atomic scale or small changes in barrier height.
Perhaps the most important observation is that the conduction in the memory ON state is restricted to a narrow conducting channel through which the electrons can, under certain conditions, travel ballistically. As a consequence, quantized resistance levels at $R = \frac{h}{2ie}$ values are observed where $i$ is the number of occupied one-dimensional conducting channels (sub-bands) and the spin degeneracy is two (in the case when no magnetic field is applied). As the applied voltage is increased, the cross-section of the conducting channel in increased. This results in additional conducting channels (sub-bands) passing through the Fermi energy and consequently the resistance drops by quantized values.

In the presence of a magnetic field, additional steps occur corresponding to the split levels at values $R = \frac{h}{2(i+\frac{1}{2})e}$. The quantized resistance effect can be observed at relatively high temperatures (up to about 190 K), suggesting an energy separation of $kT \approx \Delta E \approx 1.6 \times 10^{-2} \text{ eV}$. This indicates a change of about 40 Å in the dimensions of the cross-section of the conducting channel between the subsequent steps of the quantized resistance.

ACKNOWLEDGMENTS

It is a great pleasure to contribute to this Festschrift in honour of Professor Walter Spear on the occasion of his seventieth birthday. His contribution to the subject of amorphous materials has been outstanding and we are delighted to participate in this acknowledgment of his achievements. We have also had the privilege of collaborating with Professor Spear in the early stages of our work on amorphous silicon memory devices and we are grateful for his valuable contributions and for his continued interest.

REFERENCES


This paper contains details of a multi-layer resist scheme which provides excellent resolution capability as well as reducing many of the problems associated with the use of conventional microlithographic resists. In the "PASS" (planarized arsenic/sulfur/silver) scheme, surface planarization is first achieved by spinning on a layer which is self levelling by spin-casting. A thin film of As33S67 (at.%) is then vacuum deposited on this layer and topped with silver. The planar nature of the resist scheme reduces the problems of focus variations at steps in the circuit topology for optical lithography. During exposure, the silver diffuses rapidly into the As-S with little lateral spread. The As-S compound is soluble in a CF4 plasma whereas the As-S-Ag ternary compound is extremely insoluble under the same conditions. We can therefore dry develop the active layer. The unremoved ternary is then used to selectively protect the planarizing layer during the subsequent dry etch of this material. In experimental studies, the resist system exhibited extremely high resolution; contrast is typically in excess of 13 for optical illumination and electron-beam direct writing has produced 35 nm lines spaced by 35 nm in the active layer.

1. INTRODUCTION
The rapid movement toward higher levels of integration in monolithic circuits has been made possible by increased component packing densities and smaller geometries. The reduced feature size within circuits is largely a result of advances in lithographic techniques.

The minimum feature size which may be defined is determined by the combination of the capabilities of the system optics and the contrast value of the photoresist. This latter factor for conventional optical resists is actually quite poor. An additional resolution limitation stems from the fact that organic resists use wet developers and these will not easily dissolve material in small spaces. A high-contrast resist system which can be dry developed would therefore be desirable.

Unfortunately, the problems of current lithographic techniques do not stop with resolution. Higher numerical aperture (NA) lenses may provide better resolution but also result in decreased depth of focus. The reduced depth of focus is a serious problem for modern integrated circuits with multiple levels of metallization as these tend to have relatively large changes in surface elevation. In addition, the topography creates a problem for conventional spin-on resists as the thickness of the resist will tend to change over the surface, becoming thinner at the edges of raised features. The thin resist areas then become over-exposed and a change in linewidth occurs at the steps. Also, reflections from the substrate can lead to standing waves which reduce photospeed and create uneven exposure throughout the depth of the resist. If the underlying layer is highly reflective, whenever the layer passes over steps, incident light will be reflected not upward but laterally into the resist which leads to "reflective notching". To reduce this effect, dyed resists are frequently used but these are extremely insensitive. This suggests that a resist scheme which has a light absorbing photo-
active region at the surface would be most desirable.

2. THE PASS SYSTEM

In the planarized arsenic/sulfur/silver (PASS) multi-layer resist system, planarization is achieved by spinning on a layer which is not photoactive. This layer could be virtually any material which is self-leveling by spin-casting. For instance, organic materials such as novolac resin, polyimide, or PMMA (poly methyl methacrylate) or inorganic spin-on glasses may be diluted by solvent to the appropriate viscosity so that they are able to effectively planarize a particular circuit topography. The planarizing film is baked to remove the solvent and leave a film with little volatile component. Note that this layer is made thick enough to reduce the possibility of pinholes, particularly if high enough bake temperatures are used to cause flow.

The active layer is a film of As-S chalcogenide glass, topped with a very thin coating of silver. Both these materials may be deposited by a variety of methods but evaporation and sputtering are favored as they are relatively clean vacuum coating techniques in which control of thickness and stoichiometry is good. When exposed to light or other radiation around or above the bandgap of the glass (about 2.5 eV), the silver diffuses into the arsenic sulfide to create a ternary compound which has very different material properties compared to the arsenic sulfide alone. The total thickness of the active layer, arsenic sulfide plus silver, is typically less than 200 nm. The Ag thickness is less than 40 nm to allow sufficient light to reach the glass so that the sensitivity of the system is comparable to normal thicknesses of conventional photoresists.

3. DEVELOPMENT

The As-S compound is soluble in a CF_4 plasma, whereas the ternary compound is extremely insoluble under the same conditions. Therefore, after removing any unreacted silver by sputtering or a wet chemical dip, we can remove the As-S regions by plasma etching or reactive ion etching (RIE) with CF_4 to expose the underlying planarizing layer. The unremoved ternary is then used to selectively protect the planarizing layer during a plasma etch step with a reactant which will remove the exposed underlying material without attacking the ternary, e.g. O_2 in the case of organic materials. We may therefore dry develop our exposed resist system.

The above process yields a negative of the mask pattern. However, it is also possible to produce a positive image by following a different development sequence. The resist is exposed as before but in this case the first RIE development step uses a sulfur gas plasma which dissolves the ternary but not the arsenic sulfide. An O_2 plasma is once again used to etch the underlying planarizing layer.

The resulting structure may be used in two ways: (1) as a surface mask during a subsequent ion implantation step, etch step, lift-off step, etc. and then stripped afterwards, (2) the surface ternary is removed to leave the patterned under-layer as an inter-metal dielectric which will remain as part of the finished circuit.

4. CHOICE OF CHALCOGENIDE

The particular arsenic sulfide compound for this new resist scheme was chosen due to its inherent high resolution capability. As33S67 (at.%) forms an extremely homogeneous non-phase separated ternary glass when combined with the appropriate amount of silver. During exposure, the silver diffuses into the As-S with little lateral spread. In fact, the mechanics of silver dissolution are such that an edge-sharpening effect occurs which enhances resolution beyond what is predicted by diffraction-limited optics. This leads to a practical contrast value which is greater than 10 (the best conventional resists have a contrast of only 3 - 5). Since the active material has no large macro-molecular elements, the theoretical resolution is essentially a few nanometers. Other chalcogenides, such as germanium selenide, exhibit...
more lateral diffusion and other diffusion materials, such as copper, will tend to thermally diffuse with no applied radiation. Therefore, the As-S/Ag system is the most ideal for applications in lithography.

The As-S is almost opaque for wavelengths shorter than 400 nm due to the high absorption coefficient of the material in this region. The net result is that very little exposure radiation reaches the substrate, particularly for I-line and below, to be reflected back to the active layers. A dye element could also be added to the planarizing layer if desired (particularly for G-line radiation) to further reduce substrate reflections. This has obvious benefits for optical lithography as reflective notching will be eliminated. Alignment may still be performed through the resist using illumination between 500 nm and 1 micron. Yellow light is most ideal for this purpose as longer wavelengths will tend to be reflected and shorter wavelengths will expose the resist.

As an added benefit, since the active layer is sensitive to a wide range of wavelengths, we may combine exposure techniques on the same resist layer, e.g. optical and electron-beam, to improve throughput for ultra-small geometry circuits.

5. SUMMARY OF EXPERIMENTAL RESULTS

In general, the As33S67 material is extremely easy to deposit by evaporation with good control of layer stoichiometry as revealed by EDXA. To attain the optimum amount of silver in the arsenic sulfide film after photodoping, the As-S layer thickness has to be between 3 and 5 times that of the Ag layer and the process should be driven to completion. This puts the three component compound in the central glass forming region of the ternary phase diagram.

Although the Ag layer may be placed below the As-S, which is necessary for thicker Ag layers so that the radiation may readily reach the interface, the results in the optical exposure case tend to be disappointing. For instance, for 85 nm As33S67 on 21 nm Ag, I-line (365 nm) illumination followed by a wet development step (NaOH solution with IPA) produced a contrast of less than 1 and poor sensitivity (greater than 40 mJ/cm2) due to high absorption in the As-S. However, even with this relatively crude wet development step, a 70 nm pitch (35 nm linewidth) pattern has been produced for this combination with electron-beam exposure. In this latter experiment, a 5.5 nm diameter, 40keV beam was used to provide a line dose of 2.5 x 10^-8 C/cm.

Dry development using RIE with CF4 is also highly promising. Figures 1 and 2 show the etching characteristics of As33S67 and the photodoped ternary respectively for a reactive ion etch develop step in CF4. The etch rate of the binary is relatively high, approaching 500 A/sec. In contrast, the etch rate of the fully exposed (fully photodoped) ternary is negligible. Early work using 35 nm Ag on 200 nm As33S67 gave a contrast of 5 with a sensitivity of 20 mJ/cm2. The layer thicknesses were subsequently optimized to attain considerably higher contrast. Using 30 nm Ag on 100 nm As33S67, a contrast of 13 is readily achieved with similar sensitivity. This is illustrated in Figure 3 which is a response curve for this resist. This particular scheme is also highly sensitive to e-beam exposure and features less than...
M.N. Kozicki et al. / PASS - a chalcogenide-based lithography scheme for I.C. fabrication

NORMALIZED REMAINING THICKNESS

Fig. 2. Etching characteristics of Ag doped As$_{33}$S$_{67}$ for RIE in CF$_4$. The r.f. power is 50 W and the pressure is 93 mTorr.

NORMALIZED POST-DEVELOP THICKNESS

Log(Dose, mJ/cm$^2$)

Fig. 3. Response curve for an AsS/Ag image layer for RIE in CF$_4$. The r.f. power is 50 W and the pressure 93 mTorr.

100 nm wide may be created by line doses around 5 x 10$^{-10}$ C/cm$^2$. Having the Ag on top has the added benefit of reducing charging during e-beam exposure.

Using a 1 micron thick layer of novolac based material, topped with 120 nm As$_{33}$S$_{67}$ and 25 nm Ag, followed by a CF$_4$ develop and O$_2$ RIE step to etch the novolac, the smallest features on an optical test mask of 0.5 micron were readily transferred for I-line illumination.

6. CONCLUSIONS

A chalcogenide-based resist scheme for high-resolution microlithography has been reported. The scheme exhibits extremely high contrast for optical and e-beam exposure and may be dry developed to allow the formation of very small geometry patterns. It may significantly reduce many of the drawbacks of conventional resists including depth of focus problems and reflective notching while retaining good sensitivity.

ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation under grant number ECS-87-07089. Travel relating to the cooperative program with the Edinburgh researchers was also supported by the NSF under grant number INT-87-22490.

REFERENCES


Photodoped chalcogenides as potential infrared holographic media

C. W. Slinger, A. Zakery, P. J. S. Ewen, and A. E. Owen

The extension of holographic techniques from the visible to the infrared is important. Potentially, holographic diffractive elements have a large range of uses in this wave band. Examples include mirrors, lenses, filters, and beam combiners. All these elements would have similar advantages to those enjoyed by their visible band diffractive analogs. The metal photodissolution effect in chalcogenides shows promise as one of the few techniques for producing low-loss holographic materials for use at any given wavelength from 0.6 to beyond 16 μm. To date, the work has concentrated on the photodissolution of silver into arsenic sulfide glasses. Both bulk and surface relief gratings can be fabricated simply by holographic or mask exposure. In principle, kinoforms (e.g., blazed zone plates) and Fresnel lenses can also be made. The results of material studies show that phase gratings with high modulation and low absorption can be produced. A coupled-wave analysis is used to calculate the likely grating performance, and some initial grating characterization results are presented. The limitations of the medium are discussed and possible solutions are considered.

Introduction

The extension of optical techniques from the visible into other regions of the electromagnetic spectrum is an important goal. In particular, there are civil and military requirements that use the thermal band atmospheric windows of 3–5 and 8–14 μm. Here, in addition to conventional reflective and refractive elements, infrared diffractive optics have a role to play. They offer similar advantages to their visible band analogs. These include being lightweight and having the ability to perform novel beam-forming operations that are difficult to accomplish by using conventional elements. Their dispersive characteristics may well limit their broadband imaging applications, but conversely, enable diffractive elements to be used as filters and other, more complex, spectral domain devices. Hybrid refractive–diffractive elements may also be of use. In addition, the high cost of bulk materials currently used in thermal band refractive elements (e.g., germanium for the 8–14-μm band) may lead to significant price advantages for diffractive elements.

Both surface relief and bulk modulation structures are important classes of transmissive–diffractive elements. Bulk modulation here is taken to mean that phase modulation occurs, by variation of refractive index, throughout the volume of the material, as opposed to phase modulation caused by thickness variation (a surface relief structure). They differ significantly in their angular and wavelength selectivities—bulk structures are generally more selective because of their lower index modulation. Note that a limitation of surface relief structures is that it is difficult to determine how reflection grating behavior can be obtained from them, although there has recently been some progress in this direction.

Whereas a relatively large range of materials is available for diffractive elements in the visible wave band, the same cannot be said of thermal bands. Some surface relief type structures have been fabricated (notably by reactive-ion etching of germanium), but no high-efficiency, bulk-modulated structures have been made for use in these bands. Some workers have used photoinduced phenomena in chalcogenides to demonstrate gratings, but these were restricted to visible wavelengths. The emphasis here, however, is on research toward structures for use in the infrared, in particular, the thermal bands.

Preliminary work is presented on the use of photodoped chalcogenides as an infrared medium that are capable of being used to form both bulk and surface relief diffractive gratings. Blazed zone-plate type structures can also be fabricated. The process of grating formation is discussed first, followed by re-
suits of optical characterization of the materials. Coupled-wave theory is then used to predict the diffraction performance of the sorts of grating likely to be produced. This is followed by some preliminary experimental measurements made on gratings. Finally, the current limitations of the photodoping process are discussed, and some of the approaches by which they may be overcome are listed.

Grating Formation

The chalcogenide glasses are so named because they contain one or more of the chalcogenide elements S, Se, and Te, which may be combined with one or more of the elements Ge, Si, As, and Sb, among others, to form a wide range of binary or multicomponent glasses. These glasses are primarily known for their good infrared transmission characteristics. In addition, they exhibit a remarkably wide range of photoinduced phenomena including photodarkening, photobleaching, photopolymerization, and photocrystallization. Many chalcogenides (possibly all in principle) exhibit a metal photodissolution effect—also known as photodoping—in which illumination causes metal atoms to dissolve into the glass. The mechanism is not fully understood, but it is thought to be a radiation-enhanced solid-state process, involving a three-dimensional intercalation reaction. Initial interest in photodissolution centered on its use as a high-resolution, inorganic photoresist in the semiconductor industry, and indeed, photoinduced effects on chalcogenides are still under study for application in this area. This paper considers infrared grating formation that uses the photodissolution effect. The principle is shown schematically in Fig. 1.

An amorphous chalcogenide layer is deposited by, for example, evaporation or spin coating onto a substrate. A thin metal film is deposited on top of the chalcogenide. Light of a suitable wavelength is arranged to form an intensity pattern in the chalcogenide, corresponding to the profile of the desired grating. This pattern can be generated by exposure through a mask, as shown here—either binary or grey scale—or by conventional holographic techniques. Grey scale or diffracting masks would be particularly useful for blazed structures (although other techniques could also be used). These would allow generation of such structures in a single exposure, rather than the multiple-step procedure that is required when conventional photoresist staircase approaches are used.

Initially, the actinic radiation is absorbed at the chalcogenide–metal interface. This causes growth of metal-doped regions into the chalcogenide. Light of a suitable wavelength is arranged to form an intensity pattern in the chalcogenide, corresponding to the profile of the desired grating. This pattern can be generated by exposure through a mask, as shown here—either binary or grey scale—or by conventional holographic techniques. Grey scale or diffracting masks would be particularly useful for blazed structures (although other techniques could also be used). These would allow generation of such structures in a single exposure, rather than the multiple-step procedure that is required when conventional photoresist staircase approaches are used.

Initially, the actinic radiation is absorbed at the chalcogenide–metal interface. This causes growth of metal-doped regions into the chalcogenide. Light of a suitable wavelength is arranged to form an intensity pattern in the chalcogenide, corresponding to the profile of the desired grating. This pattern can be generated by exposure through a mask, as shown here—either binary or grey scale—or by conventional holographic techniques. Grey scale or diffracting masks would be particularly useful for blazed structures (although other techniques could also be used). These would allow generation of such structures in a single exposure, rather than the multiple-step procedure that is required when conventional photoresist staircase approaches are used.

Finally, the actinic radiation is absorbed at the chalcogenide–metal interface. This causes growth of metal-doped regions into the chalcogenide. Light of a suitable wavelength is arranged to form an intensity pattern in the chalcogenide, corresponding to the profile of the desired grating. This pattern can be generated by exposure through a mask, as shown here—either binary or grey scale—or by conventional holographic techniques. Grey scale or diffracting masks would be particularly useful for blazed structures (although other techniques could also be used). These would allow generation of such structures in a single exposure, rather than the multiple-step procedure that is required when conventional photoresist staircase approaches are used.

Finally, the actinic radiation is absorbed at the chalcogenide–metal interface. This causes growth of metal-doped regions into the chalcogenide. Light of a suitable wavelength is arranged to form an intensity pattern in the chalcogenide, corresponding to the profile of the desired grating. This pattern can be generated by exposure through a mask, as shown here—either binary or grey scale—or by conventional holographic techniques. Grey scale or diffracting masks would be particularly useful for blazed structures (although other techniques could also be used). These would allow generation of such structures in a single exposure, rather than the multiple-step procedure that is required when conventional photoresist staircase approaches are used.
exhibit the same effect. The actinic radiation can be anywhere from x rays through to beyond visible red wavelengths. The choice depends on the chalcogenide-metal system used and the thicknesses of the chalcogenide film. Thus, the reaction can be stimulated by photons with energies above or below the optical band-gap energy, although the process is more efficient when above band-gap illumination is used. Theory suggests that, for silver and As$_2$S$_3$ films of several micrometers thickness, visible green wavelengths are optimum.

**Bulk Material Properties**

For efficient diffraction, the main requirement of the grating material is that it should be of low loss in the wave band of use. For bulk gratings, a refractive-index variation (again in the wave band of use) is also necessary.

The excellent infrared transmission characteristics of chalcogenides are well known and documented (see, e.g., Ref. 9). For amorphous As$_2$S$_3$ in particular, Young$^{19}$ has reported absorption coefficients of $< 1$ cm$^{-1}$ for 0.8 to beyond 10 $\mu$m. He also reported a refractive index of from 2.4 to 2.3 in the 1-20-$\mu$m region.

Proustite (crystalline Ag$_3$AsS$_3$) has been found to be weakly absorbing between 0.65 and 12 $\mu$m and has indices of $n_E = 2.52$ and $n_o = 2.73$ in the infrared.$^{20}$ No data, however, have been published on the optical properties of As-S-Ag glasses. Measurements were therefore made by using a combination of mechanical stylus techniques and transmission as a function of wavelength. The interference-free transmission $T_e$ and the refractive index of undoped and silver-doped As$_2$S$_3$ were determined by using the method developed by Swanepoel, which does not require measurement of the reflection spectra.$^{21}$ The samples were 1.7 and 1.77 $\mu$m thick, respectively, the doped sample being formed by photodissolution of 0.17 $\mu$m of Ag into 1.6-$\mu$m As$_2$S$_3$. The results are shown in Figs. 2 and 3.

Both doped and undoped samples show good transmission from 0.7 $\mu$m to beyond 13 $\mu$m. The doped sample is more absorbing in the visible, probably because of a combination of higher intrinsic absorption and the photodarkening effect. This latter contribution can be removed by annealing. The overall transmission of the doped glass is lower, mainly because of the higher reflection losses caused by its larger refractive index, as shown in Fig. 3. There is a refractive-index difference between undoped and doped samples of 0.4 across most of the transmission region, showing the high modulation that can be achieved by photodoping.

The refractive index of the doped regions has been found to be related directly to the amount of silver present (Fig. 4). Several groups, using a variety of different techniques, have shown independently that the metal concentration in a sample undergoing photodissolution has a steplike profile and not the Fickian distribution characteristic of normal dopant diffusion in semiconductors.$^{11}$ Hence the silver concentrations in the samples that yielded the data in Fig. 4 are expected to be essentially constant with depth. By suitable choice of the silver thickness, relative to that

![Fig. 2. Interference-free transmission $T_e$ of undoped and doped evaporated As$_2$S$_3$ glass material as a function of wavelength: undoped thickness 1.7 $\mu$m, doped thickness 1.77 $\mu$m (from 1.6-$\mu$m As$_2$S$_3$ doped with 1.7 nm Ag).](image-url)

![Fig. 3. Refractive index of undoped and doped material as a function of wavelength. The materials are as in Fig. 2.](image-url)

![Fig. 4. Refractive index versus wavelength for 500-nm thick As$_2$S$_3$ films photodoped with various amounts of Ag. The Ag concentration in the films was controlled by varying the thickness of the initial Ag layer, and these thicknesses are indicated in the figure.](image-url)
of the chalcogenide, the index modulation of the gratings can be varied as desired. The ability to fabricate gratings with specific index modulations may be an advantage in some circumstances.

**Theoretical Analysis**

The performance of the gratings produced by the metal photodissolution effect can be predicted to good accuracy by using a coupled-wave analysis. The underlying theory has been covered elsewhere (see, e.g., Ref. 22), and only the main results are presented here. The replay situation described below is used for simplicity. Other geometries and grating configurations can be analyzed by relatively straightforward extensions of the theory.

The grating is assumed to be lossless, phase modulated, unslanted (i.e., the grating fringes are perpendicular to the substrate surface), and have a profile that can be described by an even function. The slab of thickness \(d\) containing the grating is taken to be infinite in the \(y-z\) plane with a normal in the \(x\) direction and bounded by homogeneous regions. The grating vector \(\mathbf{K}\) lies in the \(y\) direction, where \(K = 2\pi/\lambda\) and \(\lambda\) is the grating period. The grating parameters may generally vary as a function of \(y\) and \(z\), but it is assumed to be locally plane. An infinite, monochromatic plane wave of amplitude \(A_0\) is incident in the \(x-y\) plane.

The coupled-wave equations that describe the propagation of diffraction orders inside the grating (\(0 < x < d\)) can be written in terms of the dimensionless parameters \(\zeta, \Omega\) as a set for \(m = -N, \ldots, -1, 0, +1, +2, \ldots, +N\):

\[
-k_1 \frac{d^2 A_m}{\cos \theta_0 j \beta d} \frac{d A_m}{d \zeta} - j m (1 + P) A_m + \sum_{i+1}^{N} (A_{m+i} + A_{m-i}) \zeta = 0. \tag{1}
\]

where \(A_m\) is the amplitude of the \(m\)th diffraction order, \(k_1 = \beta k_0/\gamma k_i\), \(\beta = 2\pi/\lambda\) is the propagation constant in the grating, \(\zeta\) is the \(\z\)th Fourier harmonic of the phase modulation profile at depth \(x\), \(\zeta = k_1 x/\cos \theta_0\) is a modulation parameter, \(\Omega = K^2/2\beta k_0\) is a volume parameter, \(P = (\sin \theta_0)k_0/k_1\). \(K\) is a Bragg parameter, and \(\theta_0\) is the propagation angle of wave \(0\) inside the grating.

These equations, subject to the grating boundary conditions (matching of the tangential electric and magnetic fields at \(x = 0, d\)), can be solved numerically—only in the limiting case of \(\Omega > 1\) are some relatively simple analytic solutions possible. The amplitudes of the diffracted waves can thus be determined as a function of the grating parameters.

It is important to realize that the true form of the grating produced by the photodissolution process is essentially that of a surface relief structure. Thus holographic or gray scale mask exposures result in forms of grating that are generally different from those produced in other, bulk holographic recording media. For example, a two-plane-wave, holographic exposure produces a sinusoidal intensity distribution in the chalcogenide. The square root time dependence of the photodissolution process (see later), combined with the requirement of growth from a source of metal atoms results in a parabolic–sinusoidal structure. This can be analyzed by slicing the grating perpendicular to the \(x\) direction. In this way, the profile generated by the sinusoidal intensity distribution can be represented as a rectangular grating of variable mark–space ratio:

\[
e_0 = e_{max} + \Delta t \epsilon, \quad \epsilon = \frac{2 \Delta t}{\lambda} \sin (\mu \pi \zeta), \tag{2}
\]

with

\[
\mu = \frac{1}{\pi} \cos^{-1} \left[ \frac{2 (x \zeta/n)^{-2}}{d \zeta} - 1 \right]
\]

where \(\epsilon = e_{max} - e_{min}\) and \(e_{max}\) and \(e_{min}\) are the maximum and minimum relative dielectric constants, corresponding to the doped and undoped regions, respectively, and \(\Delta t\) is a saturation factor.

Further simplifications to Eq. (1) can be made in some cases. If the ratio \(e_i/e_f < 1\) (generally a value of 0.2 is the limit), the second derivatives in Eq. (1) can be neglected with little loss in accuracy. For bulk and surface relief gratings in photodoped chalcogenides, typical values for this ratio are \(0.1\) and \(1.0\), respectively. Thus, for bulk gratings the first-order theory can be used, and amplitudes \(A_m\) can be described in terms of the three dimensionless parameters \(\zeta, \Omega, P\). Equation (1) becomes

\[
\frac{d A_m}{d \zeta} = \mu (1 + P) A_m + \sum_{i+1}^{N} (A_{m+i} + A_{m-i}) \zeta = 0. \tag{3}
\]

with transmission grating boundary conditions of

\[
A_0(\zeta = 0) = 1, \quad A_m(\zeta = 0) = 0 \quad (m > 0). \tag{4}
\]

Analytic solutions of Eq. (3) are possible for the cases of \(\Omega < 1\) (Ref. 25) and \(\Omega \gg 1\) (Ref. 26), corresponding to thin (Raman–Nath) and volume diffraction regimes, respectively. Generally speaking, thin regime behavior occurs to good accuracy for \(\Omega < 0.1\), and volume diffraction for \(\Omega > 10\). Between these two types of behavior, the grating occupies the multivwave diffraction regime and numerical solution is necessary.

Usually, one of the main requirements for a diffractive element is that the first-order diffraction efficiency \(\eta_1\) is high (>80%, say). \(\eta_1\) can be defined as the ratio of power in the \(i\)th diffraction order at \(x = d\) to that in the incident order at \(x = 0\). In the volume regime, for \(m\)th order on-Bragg replay \((m + P = 0)\), near 100% efficiency is possible, regardless of grating profile. In the other regimes, the achievable efficiency does depend on the grating profile.
Figure 5 shows the first-order diffraction efficiency as a function of $\xi$ and $\Omega$ for a square-wave, bulk grating. These results were obtained by numerical solution of Eq. (2) for $P = -1$ (on-Bragg for the first diffraction order). Reflection losses have been neglected. The square-wave grating is equivalent to exposure of the chalcogenide through a binary mask. The efficiencies are seen to agree with the analytic solutions in the thin and volume regimes having maxima of 40.5% and 100%, respectively. In the multiwave regime, it is interesting to note the presence of several high-efficiency regions. For example, $\eta_1 > 96\%$ at $(\xi, \Omega) = (\pi/2, 2.3)$ and $> 90\%$ at $(3.0, 0.5)$. Such operating points may be useful in some applications. For a bulk square-wave grating with refractive indices of 2.3 (undoped) and 2.65 (doped region) at $(\xi, \Omega) = (\pi/2, 2.3)$, a grating thickness of 2.19A is required—around 23 $\mu$m at $\lambda = 10.6$ $\mu$m.

It is of interest to compare the performance of the bulk grating with that of the equivalent surface relief structure. For the latter, the rigorous coupled-wave Eq. (1) must be solved numerically. The lack of analytic solutions for the second-order theory does not mean that trends cannot be identified. It was found that the volume parameter $\Omega$ still influenced the number of significant diffraction orders that were present. Thus, generally speaking, the larger the refractive index of the surface relief structure (and therefore the larger the index modulation), the smaller the grating thickness $d$ needed for peak efficiencies, although these peak efficiencies were lower than surface relief gratings of smaller index modulation. Low-index gratings gave higher peak efficiencies (because of their higher $\Omega$ values) but required greater depths to achieve them. A typical result is shown in Fig. 6, in which first-order diffraction efficiency is plotted as a function of the normalized grating thickness for a square-wave profile grating. Such a grating could be obtained by reactive ion etching of the above bulk structure. The equivalent result for a sinusoidal profile surface relief grating is also shown.

For the cases shown in Fig. 6, peak efficiencies of 63% and 77% occur at $d = 0.76\lambda$ and $1.1\lambda$ for the square and sinusoidal profiles, respectively. Note that these figures should not be regarded as the maximum achievable. They can be increased by suitable coatings or changes in the grating parameters (choice of lower index or decrease of the grating period). A 0.15A thick conformal coating of 1.58 index increased the calculated peak efficiencies to 83% and 87% for the square and sinusoidal profiles, respectively.

For structures other than the relatively high frequency diffractive gratings ($\Lambda = \lambda$) analyzed above, for example, kinoforms (e.g., zone plates) and Fresnel lenses, a coupled-wave analysis is not necessary. Fresnel lenses rely on refractive principles, so are capable of 100% efficiency, although the thicknesses required are comparatively much higher than in diffractive devices. Blazed zone plates can also achieve 100% efficiency, the depth required being given by $d = \lambda/\sin \delta$. The so-called binary optics $^{28}$ are also capable of high efficiencies (over 90% for six-level phase structures, for example).

Thus, for both bulk and surface relief structures, typical of those capable of being produced by photodoped chalcogenides, high efficiencies should be achievable.

Experimental Work

As discussed in the following section, these material systems currently require long exposure times, hence to date, fabrication of gratings has been restricted to chalcogenide thicknesses of up to 2 $\mu$m or so. The
modulation parameter \( \xi \) in the infrared is thus small. Measurements have therefore been at visible and near-infrared wavelengths. Surface relief gratings have been made by using holographic and mask exposures followed by alkali etch. The grating can be of good quality, exhibiting few surface defects.

Several bulk gratings (both thin and volume types) have also been made by using holographic exposures. Recordings were made by using the interference of two plane waves from a single-frequency, argon-ion laser operating at 514.5 nm (theory indicates that wavelengths corresponding to the optical gap energy of the chalcogenide, \( \sim 2.4 \text{ eV for As}_2\text{S}_3 \), are the most effective for inducing photodissolution in the thick samples; shorter wavelengths are highly absorbed in the chalcogenide, whereas longer wavelengths are less efficient at stimulating the effect). Typical power densities were 0.25 W cm\(^{-2}\)/wave. The samples were made by evaporating \( \text{As}_2\text{S}_3 \) (supplied by Servofrax) to a thickness of 1.6 \( \mu \text{m} \) on microscope slides and subsequently coated (without a break in vacuum) with 275 nm of Ag. After exposure, the samples were washed with a solution of iron (III) nitrate, to remove the unused Ag. By varying the interbeam angles of the recording waves, we made both thin, multwave and volume gratings. For example, an interbeam angle of approximately \( \pm 2^\circ \) to the substrate normal produced bulk gratings having a significantly thin character. On replay, using a He–Ne laser at 632.8 nm, we could observe many diffraction orders that were typical of Raman–Nath behavior. Figure 7 is an optical phase contrast micrograph of such a grating of period 22 \( \mu \text{m} \), formed by photodoping 3-\( \mu \text{m} \) \( \text{As}_2\text{S}_3 \) with 1 \( \mu \text{m} \) of Ag. The definition of the grating is reasonable, although the small-scale structure that is apparent may be due to some kind of inhomogeneity in the photodoped material, for example, phase-separated clusters. A change in the arsenic sulfide composition may well remove these clusters, and in any case, scatter from them will decrease at longer wavelengths.

Figure 8 shows a scanning electron microscope image of a cleaved bulk grating. The doped regions can be distinguished from the undoped arsenic sulfide, and the shape of the doped regions appears to follow the parabolic–sinusoidal profile predicted from the kinetics results.

Interbeam angles of \( \pm 25^\circ \) results in gratings displaying significant volume behavior at visible wavelengths. Figure 9 shows the measured 0 and \( +1 \) diffraction orders as a function of replay angle for an early version of one such grating with an exposure time of 120 s. Replay was in air, using a He–Ne laser at 632.8 nm. For the case shown, just under 9% efficiency is achieved, all but 1% of the energy being removed from the incident wave. Reflection losses (nearly 30% at the Bragg angle), absorption, and scatter account for the missing energy. The absorp-
tion may be partly due to the photodarkening of the chalcogenide (a shift in the visible absorption edge). Photodarkening has little effect in the infrared band. In addition, it can be removed by annealing or reduced by using a different chalcogenide composition. Additionally, this early sample suffered from poor film adhesion.

Equations (2) were modified to include bulk losses and absorption modulation (see, e.g., Ref. 29). A theoretical analysis of the grating was made, based on Eq. (4) and numerical solution of the modified equations. The results are shown as solid curves in Fig. 9. A Talystep profilometer was used to determine the thickness $d$ of the sample. The parameters derived from the fit are within the range expected from the bulk material characteristics. The agreement between theory and experiment is generally good, the main features of the curves being well matched.

Figures 10 and 11 show the measured diffraction performance of more recent gratings at 632.8 nm and 1.5 $\mu$m, respectively. The angular response shown in Fig. 10 is from a grating recorded as described above, but using 2.5 $\mu$m of As$_{30}$S$_{70}$ with 490 nm of Ag and exposed at angles of $\pm$30$^\circ$. A peak diffraction efficiency of 34% can be seen. This increase in efficiency over the previous result (Fig. 9) can be attributed to increased film quality and both the lower intrinsic absorption and a reduction in photodarkening of the lower arsenic content film. In fact, photodarkening is negligible near the composition of As$_{30}$S$_{70}$.

The grating whose angular response is shown in Fig. 11 was optimized for performance at near-infrared wavelengths, and recording angles of the plane waves were therefore reduced to $\pm$12.5$^\circ$. 2.6 $\mu$m of As$_{33}$S$_{67}$ coated with 1.0 $\mu$m of Ag were exposed for 240 s. Despite the reduced values of $\xi$ at the longer replay wavelengths of 1.5 $\mu$m, the lower absorption and reduced scatter result in a first-order diffraction efficiency that exceeds 35%. These results are encouraging and confirm that, for greater grating thicknesses (such that $\xi \rightarrow \pi/2$), high efficiencies (>90% neglecting reflection losses) should certainly be achievable at infrared wavelengths. However, both Fig. 9 and 10 show anomalous behavior in their angular responses (examples of which are marked in the figures); their cause is as yet unknown, but they may be caused by noise grating phenomena or other interference effects at recording.

Possible Limitations of the Photodissolution Process

Possible problems that may limit the applicability of the photodissolution process include its speed and the classes of grating that can be fabricated. The rate at which the process occurs determines the exposure time required for a given grating thickness $d$. Coupled wave theory can be used to predict the values of $d$ that are required for a particular type of grating and wavelength of operation. The results for a square-wave grating, assuming undoped and doped refractive indices of 2.3 and 2.65, respectively, are given in Table 1. Volume operation at $(\xi, \Omega) = (\pi/2, 10)$ was assumed for the bulk grating. For the surface relief profile, the grating period $A$ was taken to be equal to the replay wavelength $\lambda$. The thickness required for a blazed zone plate is also shown.

It can be seen that, for far-infrared operation, thicknesses of several micrometers are required. Note that the surface relief structures are approximately a factor of 3 or 4 thinner than the equivalent bulk gratings—an advantage in their favor. Silver photodissolves to a depth of >20 $\mu$m in some chalcogenides but can diffractive structures be produced to the required depths for high efficiency? To answer this question it is necessary to study the factors that govern the speed of the process. The depth to which the chalcogenide has been doped can be written as a function of the wavelength of the

| Table 1. Required Thicknesses for Square-Wave Gratings |
|-----------------|--------|--------|--------|
| Replay Wavelength/µm |        |        |        |
| 0.632            | 4.0    | 10.6   |
| $d$/µm           |        |        |        |
| Bulk             | 1.39   | 8.77   | 23.2   |
| Surface relief   | 0.48   | 3.0    | 8.06   |
| Blazed zone plate| 0.38   | 2.4    | 6.12   |

Fig. 10. Angular response of a more recent grating, again measured in air at 632.8 nm. The details of the grating are in the text. A first-order diffraction efficiency of 34% can be seen. Note the anomalies in the response, examples of which are marked a in the figure.
An increase in photodissolution rate with photon energy is counterbalanced by the increasing absorption of the material (both undoped and doped) with decreasing wavelength. Theory suggests that visible green wavelengths are optimum for thicknesses greater than a micrometer or so. The dependence of the rate on intensity is probably linear for low powers but seems to saturate at higher levels (typically 300 mW cm\(^{-2}\)) at 514.5 nm for 4 \(\mu\)m of As\(_{30}\)S\(_{70}/Ag\). The effect of the exposure time on the photodissolved depth is not straightforward. Initially, the doped thickness increases with time \(t\) as \(k_1\sqrt{t}\) until the Ag reservoir is exhausted. After this occurs, the doped region acts as the source of Ag and \(d\) varies as \(k_2Vt\).

The photodissolution rate also depends on the chalcogenide composition. Experiments have been performed investigating this by measuring the transmission and reflection of a sample while exposure is occurring. The optimum composition for the As\(_{30}\)S\(_{70}/Ag\) system was around As\(_{30}\)S\(_{70}\). Fortuitously, this is close to the composition for which minimum photodarkening occurs. However, the rate does not vary significantly for small variations about this optimum.

To give some idea of the rates found in practice, typically a 700-s exposure with a power density of 300 mW cm\(^{-2}\) at 514.5 nm was necessary for a 1-\(\mu\)m thick As\(_{30}\)S\(_{70}\) sample, with an excess of Ag, at 21°C. Therefore, unless great care is taken (e.g., by using active fringe stabilization techniques; holographic exposures for far-infrared gratings would be difficult. Mask exposures, while overcoming the stability problems, may still be awkward.

However, several avenues of study may yield an increase in the photodissolution rate. These include seeding of the chalcogenide with a small percentage of Ag before doping. A 1% addition of silver to As\(_{30}\)S\(_{70}\) increases the conductivity by a factor of \(2 \times 10^3\). The ionic and electronic conductivities are thought to influence photodissolution, so the rate should increase with seeding, although our results so far have not supported this. It is nevertheless possible that other additives may enhance the diffusion or reaction rates. Application of an electric field during exposure may also increase the doping rate. Other possibilities include the use of a photosensitive silver layer to minimize the exposure requirements, the use of double metal layers (e.g., Ag and Cu) and other chalcogenide-metal systems. Finally, rates at higher temperatures are greater, although the advantages may be limited by lateral doping that is due to the closely related thermal dissolution effect. Other current limitations of the photodissolution process include the restricted class of structures that the process is capable of fabricating. Transmission bulk and surface relief gratings having low-to-medium short angles (<45°, say) in addition to binary forms can certainly be made. But it is difficult to envisage the writing of gratings having larger slant angles, since Ag source must not overlap with each fringe, and fringe growth starts from the source. Thus fabrication of reflection (conformal) gratings would require some effort.

**Conclusion**

A preliminary experimental and theoretical investigation of the metal photodissolution effect in chalcogenide glasses has been presented. In particular, the silver doping of arsenic sulfide glasses has been examined. The bulk optical properties of this material system show low losses from the visible red to beyond 15 \(\mu\)m. Surface relief structures obtained by using this process should be of use at any given wavelength in this wide band. In addition, refractive index measurements show that phase modulation should also be possible over this waveband, enabling bulk diffractive structures to be made.

Theory has shown that high efficiencies should be possible. Experiments on visible band gratings showed reasonable agreement with the theory, when this regime and near-volume transmission gratings were fabricated. Measurements of grating efficiency for near-infrared wavelengths (1.5 \(\mu\)m) have confirmed that high efficiencies should be possible after suitable gratings are fabricated. Although the large material thicknesses required for 8-14-\(\mu\)m band structures may be difficult to fabricate, several approaches could overcome this limitation. Currently, the process is restricted to transmission-type structures although work is under way to try to fabricate reflection gratings.

Although the emphasis in this paper has been for use within the 3-5 and 8-14-\(\mu\)m thermal bands, the large bandwidth optical properties of the material system may enable it to be applied usefully to other spectral regions, for example, in the near infrared for optical communication applications.

The authors thank Abdelkarim Zekak for the scanning electron microscope micrograph used in Fig. 8. This work was carried out with the support of the Defence Procurement Executive, Ministry of Defence.

**References**


10 May 1992 / Vol. 31, No. 14 / APPLIED OPTICS
14 Photo-induced changes in chalcogenide glasses and their applications

P. J. S. EWEN and A. E. OWEN

14.1 Introduction

The chalcogenide glasses are one of the most widely known families of amorphous materials and have been extensively studied over the past three decades, partly because of their interesting fundamental properties and partly because of their applications. Indeed, it was the discovery of electronic switching in chalcogenide systems such as Si–Te–As–Ge (Ovshinsky, 1968) that stimulated the upsurge of interest in amorphous materials in general that occurred during the 1960s. The chalcogenides are so called because they contain one or more of the chalcogen elements S, Se or Te. These elements can be combined with many others, for example As, Sb, P, Si and Ge, to form a wide variety of glass-forming systems, ranging from simple binary combinations to more complex multi-component systems such as As–Te–Ge–S–P. The As–S, As–Se and Ge–Se systems are among the most widely studied of the binary combinations and include the well-known compound glasses \( \text{As}_2\text{S}_3 \), \( \text{As}_2\text{Se}_3 \) and \( \text{GeSe}_2 \). In general, the chalcogenide glasses are relatively easy to prepare in either bulk or thin-film form and their properties are often comparatively insensitive to the details of preparation, which is an advantage both in experimental studies and in applications. Further information on the preparation and general properties of these glasses can be found in various texts and review articles (Elliott, 1988; Mott and Davis, 1979; Madan and Shaw, 1988; Savage, 1985; Zallen, 1983).

The most important applications of chalcogenide glasses are now in the field of optics and arise chiefly from either their infrared-transmitting properties or the many photo-induced effects they exhibit. They have potential uses in integrated optics, optical imaging (e.g. holography and very large scale integration (VLSI) lithography), optical data storage and infrared optics. In this chapter the various types of photo-induced effects that occur in amorphous chalcogenides are classified and described, with particular emphasis on the As–S system, and their applications are reviewed. Further information on the fabrication of chalcogenide glasses and their use as infrared waveguides is given in chapter 13.
14.2 The classification of photo-induced effects in chalcogenide glasses

Chalcogenide glasses exhibit a wide variety of light-induced changes, ranging from relatively subtle effects involving minor atomic rearrangements and manifested mainly by shifts in the optical absorption edge, to more substantial atomic and molecular reconfigurations which cause a variety of physical and chemical changes. These effects are of fundamental interest because of the information they yield on defects and metastable structural states in amorphous solids, and because they are mostly unique to the amorphous state. There are several earlier reviews of this field (Elliott, 1985; 1986; Owen et al., 1985; Tanaka, 1981a,b).

At least seven distinct photo-induced structural or physicochemical changes have been observed in amorphous chalcogenides when samples in a suitable form are exposed to light or other irradiation, viz. photocrystallization, photopolymerization, photodecomposition (in compounds), photo-induced morphological changes, photovaporization, photodissolution (of certain metals), and light-induced changes in local atomic configuration. In general these changes are accompanied by changes in the optical constants of the material and particularly shifts in the absorption edge, i.e. photodarkening or photobleaching. Other characteristics of the material, such as the elastic constants (Kolomiets and Lyubin, 1978; Tanaka et al., 1981), may also be affected. There are two reasons why the chalcogenide glasses are so susceptible to light-induced changes (Elliott, 1985; 1986). Firstly, they have a great deal of structural flexibility, partly because the chalcogen atoms are normally only twofold co-ordinated, and partly because their structures tend to be chain-like or layer-like rather than three-dimensional continuous random networks. The second reason is that the chalcogen atoms possess lone-pair electrons which are normally non-bonding but which can be involved in light-induced reactions to produce structural defects consisting of threefold or singly co-ordinated chalcogen atoms. The highest-lying levels in the valence band of these materials consist of the states associated with these non-bonding electrons and are therefore the states that are preferentially excited by illumination.

The photo-induced phenomena that occur in the chalcogenides may be classified according to whether they are primarily structural or physicochemical in nature and whether they are reversible or irreversible, in the sense that the system may revert partly to its initial state after some annealing treatment at a temperature higher than that at which illumination took place, or will recover completely on annealing at the glass transition temperature, $T_g$.

The reversible effects are generally observed most readily in well-annealed vapour-deposited films, or in melt-quenched glasses, which are also comparatively well annealed as a natural consequence of their
PHOTO-INDUCED CHANGES IN CHALCOGENIDE GLASSES

method of preparation. By contrast, the irreversible changes that are primarily structural are found mainly in poorly-annealed vapour-deposited films, while the irreversible physicochemical phenomena occur whatever the state of annealing of the amorphous sample. The inter-relationships between the reversible or irreversible, and structural or physicochemical effects are illustrated schematically in Figure 14.1.

14.3 Reversible photo-induced effects

14.3.1 Photo-induced changes in local atomic structure

Reversible photo-induced changes in local atomic structure are subtle effects which are induced by light of photon energy greater than or close to that of the optical band-gap, $E_g$, of the chalcogenide glass. They are manifested mainly by small shifts ($<-0.1$ eV) in the optical absorption edge to lower energies, that is photodarkening. These shifts are accompanied by small changes in the radial distribution function determined by diffraction experiments, or in the vibrational spectra, and also, normally, by barely discernible changes in volume (contraction or expansion). Amorphous As$_2$S$_3$ is the best-known example of a chalcogenide glass exhibiting these structural changes (Tanaka, 1981a,b) but they have also been observed in other As–S and As–Se glasses (particularly a–As$_2$Se$_3$ and As$_{40}$Se$_{60}$), and in a–GeS$_2$, amorphous Ge–Se compounds (particularly a–GeSe$_2$) and a–As$_4$Se$_5$Ge (Kolomiets and Lyubin, 1978; Tanaka, 1980). Photodarkening does not occur in Group IV or Group V amorphous semiconductors and appears to be a characteristic of most, though not all, chalcogenides (amorphous Te does not exhibit this
effect, and in the case of As$_2$S$_3$ and As$_2$Se$_3$ it can be destroyed by the addition of Cu (Liu and Taylor, 1987). Since photodarkening is sensitive to pressure (Pfeiffer and Paesler, 1989; Tanaka, 1984; Tsutsu et al., 1984), van der Waals bonding may also have a role in the process (Tanaka, 1983). It should be noted that photodarkening does not occur in crystalline As$_2$S$_3$, indicating that this particular photo-induced effect is probably unique to the amorphous state (Tanaka, 1981a,b) and must involve structural features only present in the disordered material.

These effects are reversible in that annealing restores the initial structure and properties, and in some cases the initial state can also be recovered by exposure to light whose photon energy is less than that used to photodarken the sample (Hamanaka et al., 1981).

It has been suggested that the structural change induced by light consists of minor bond rearrangements involving the chalcogen atom relaxing between the minima of double-well potentials (Averianov et al., 1980; Kolomiets and Lyubin, 1978; Tanaka, 1980). Figure 14.2 shows a schematic model of conjectured bistable local bonding geometries in a-As$_2$S$_3$ and the associated double-well potential: illumination induces a transition from arrangement A to arrangement A', possibly as a result of the strong electron-phonon coupling involved in the recombination.

![Figure 14.2 A schematic model of bistable bonding geometries in a-As$_2$S$_3$ and the corresponding double-well potential. A and A' represent atomic configurations before and after illumination. R is some configurational coordinate and U the potential.](image_url)
PHOTO-INDUCED CHANGES IN CHALCOGENIDE GLASSES

of a photo-excited valence electron (Murayama et al., 1980). X-ray absorption studies suggest that the actual structural change occurring during photodarkening in As$_2$S$_3$ is more complicated than that shown in Figure 14.2 and may involve a twisting of adjacent AsS$_3$ pyramids about their shared S atoms together with an increase of the As–S–As bond angle (Lee et al., 1989; Yang et al., 1987). A photo-induced increase in interlayer separation has also been suggested (Lowe et al., 1986).

Another type of photo-induced change in local atomic structure which may also occur in chalcogenides is that produced by the 'self-trapped exciton' (Begelson and Street, 1980; Elliott, 1985; 1986). This involves the breaking of bonds—for example, in the case of a–As$_2$Se$_3$, an optically excited electron-hole pair would cause an As–Se bond to break and a Se–Se bond to form, resulting in a pair of metastable defects Se$_3^-$ and As$_2$, possibly accompanied by deformation of the local structure. Further work is necessary to confirm the existence of these photo-induced changes.

A third type of local structural change has been proposed to account for the appearance of optical anistropy in certain chalcogenides when illuminated with plane-polarized light (Grigorovici et al., 1983; Hajto and Ewen, 1979; Hajto et al., 1982; Hajto and Janossy, 1987; Zhdanov and Malinovskii, 1977; Zhdanov et al., 1979). This phenomenon, which has been termed the ‘vectorial effect’, is believed to involve the re-orientation of small anisotropic structural units already present in the amorphous material. The principal axes of these units are initially randomly oriented but under illumination with plane-polarized light the units are forced to re-orient themselves in such a way that the axes become aligned with some preferred direction defined by the plane of polarization. In the case of GeSe$_2$, it has been suggested that the anisotropic units are the ‘outrigger rafts’ identified by Phillips (1981). The induced anisotropy can be erased by unpolarized or circularly polarized light (Hajto et al., 1982).

14.3.2 Photodecomposition

As the name implies, photodecomposition is the light-induced dissociation of a compound into its constituents. It has been studied most frequently in amorphous As–S and As–Se compositions and illumination of photon energy greater than $E_g$ is necessary (Asahara and Izumitani, 1975; Berkes et al., 1971; Hamanaka et al., 1976; Keneman et al., 1978; Tanaka, 1981a,b). Absorption of a photon of energy greater than $E_g$ causes a dissociation of the arsenic–chalcogen bond, and this is followed by thermal diffusion of the arsenic to form arsenic clusters,
which are highly absorbing. Photodecomposition causes small but detectable changes in X-ray diffraction and vibrational spectra, but is manifested mainly by a red-shift in the optical absorption edge, i.e. photodarkening. Provided the energy and intensity of illumination are not too great the photodecomposition is reversible on annealing at $T_g$. Clearly the proportion of dissociated bonds must be small.

Small clusters of arsenic atoms can, in fact, be formed without significant diffusion of the arsenic: a local bond re-organization such as that illustrated for $a-$As$_2$S$_3$ in Figure 14.3 will give rise to As$_2$ units (i.e. As–As bonds) and several such re-organizations occurring in the same locality will produce clusters of three or four arsenic atoms. In well-annealed films or bulk samples of $a-$As$_2$S$_3$, heteropolar As–S bonds are favoured, although a small percentage (about 1%) of homopolar As–As and S–S bonds may be present (Ewen and Owen, 1980). It is possible that above-band-gap illumination alters the bond statistics from a distribution close to that for the chemically-ordered-network model towards one characteristic of a random-network model, that is the randomness of the bond distribution is increased.

Isolated As–As bonds are energetically unfavourable and when produced by the process illustrated in Figure 14.3 will soon revert to the initial state (Figure 14.3(a)). Larger clusters, in contrast, are expected to be relatively stable and hence can give rise to the photodarkening. At $T_g$, however, the atoms are more mobile, and annealing at this temperature causes even larger clusters to 'dissolve', so that the effect is reversible. Exposure to below-band-gap illumination also reverses the photodarkening. Direct evidence for the formation of As–As bonds in

![Figure 14.3](https://example.com/figure14.3.png)

**Figure 14.3** Schematic model of local bond reorganizations in $a-$As$_2$S$_3$ leading to the formation of an As–As bond. The transition between the initial structure (a) and the reorganized structure (b) is reversible.
well-annealed $\alpha$–As$_2$S$_3$ after exposure to above-band-gap light has been obtained in Raman studies (Frumar et al., 1984).

The increase in film thickness that accompanies this photodarkening in $\alpha$–As$_2$S$_3$ is accounted for by the increase in the average bond length resulting from the replacement of As–S bonds by the longer As–As bonds. Measurements of the fractional change in film thickness can thus be used to estimate the percentage of As–S bonds converted into As–As bonds under illumination, and a value of approximately 6% has been derived (Frumar et al., 1984).

It is interesting to note that if isolated As–As bonds were stable, the two structures depicted in Figure 14.3 could be described as bistable bonding arrangements, by analogy with the two bistable geometries shown in Figure 14.2. It is thus possible that stable homopolar bond formation in chalcogenides is one origin of metastable states.

14.4 Irreversible photo-induced effects

14.4.1 Photocrystallization

Chalcogenide glasses, by definition, have a non-crystalline structure, lacking long-range order. For some chalcogenides illumination can break interatomic bonds directly, and/or indirectly via local heating, thereby allowing the atoms to re-arrange themselves into an ordered crystalline structure. The photocrystallization of amorphous solids is well known and was first observed over two decades ago in a-Se (Dresner and Stringfellow, 1968; De Neufville, 1974). Photo-induced crystallization and ‘re-amorphization’ has been studied in thin films of As–Te–Ge compositions, using Kr-ion and dye lasers (Von Gutfield and Chaudhari, 1972; Weiser et al., 1973). Thus, in a sense the process is reversible, but ‘re-amorphization’ requires that the temperature of the film first be raised to the melting point of the crystals, and this can be achieved with a laser pulse. Strictly, therefore, this is an irreversible process.

14.4.2 Photopolymerization

When chalcogenide compounds are deposited as thin films by evaporation the as-deposited material often contains some of the molecular species of which the vapour is composed; for example, As$_4$S$_4$ molecules have been identified in As$_2$S$_3$ vapour (Solin and Paptheodorou, 1977) and in as-deposited $\alpha$–As$_2$S$_3$ films (Nemanich et al., 1978). In the case of $\alpha$–As$_2$S$_3$ these molecular species are embedded in an amorphous As–S network and illumination causes them to polymerize and combine
with the network (Treacy et al., 1980). Polymerization causes densification and in general, on prolonged exposure, the density and structure of the thin films become virtually identical with those of melt-quenched glasses and well-annealed films (Chang and Chen, 1978; Chang and Hou, 1978).

Photopolymerization has been proposed as an explanation of reversible photodarkening in well-annealed amorphous As–Se films (Grigorovici and Vancu, 1981). For such a model to be applicable, molecular species must exist in the film prior to exposure and annealing must cause depolymerization. It is not clear how photopolymerization can result in the film expansion that often accompanies reversible photodarkening. Photopolymerization has also been observed in the pseudo-binary system $(\text{As}_2 \text{S}_3)_{1-x} (\text{As}_2 \text{Se}_3)_x$ (Onari et al., 1985).

Photopolymerization of $\text{As}_4 \text{S}_4$ molecules has been reported in crystals of $\alpha$- and $\beta$-$\text{As}_4 \text{S}_4$ (Porter and Sheldrick, 1972), so this particular photo-induced effect is not unique to the amorphous state.

### 14.4.3 Photo-induced morphological changes

'Photo-induced morphological changes' is a term proposed for use to denote effects usually described in the literature as 'photovolumetric changes' or 'photocontraction', since the phenomenon has its origin in changes in the macroscopic structure of chalcogenide thin films, that is in their morphology. Changes in morphology have been observed mainly in thin films of Ge–Se compositions which, during vapour-phase deposition, grow on the supporting substrate in a columnar structure (Singh et al., 1979; 1980a, b; Spence and Elliott, 1987). On illumination, particularly with ultraviolet light, the films contract in volume. The magnitude of the change depends to some extent on composition but mainly on the angle of incidence at which the vapour-deposited film is grown. Contractions in volume of up to 12% can occur. The effect is thought to be due to a collapse of the columnar morphology of the films, precipitated by a photo-electronic excitation of the electronic defect states that are characteristic of amorphous chalcogenides (Singh et al., 1980b).

### 14.4.4 Photovaporization

Photovaporization has been observed mainly in amorphous As$_2$S$_3$ thin films and occurs at elevated temperatures ($>\sim 150 ^\circ C$). It is strictly a photo-oxidation reaction followed by thermal evaporation of the volatile oxidation products (Janai, 1981a; Janai and Rudman, 1974): light causes a surface reaction between the oxygen in the air above the film and the
As$_2$S$_3$ to form arsenic oxide and sulphur, both of which evaporate off, resulting in a thinning of the film.

14.4.5 Photodissolution of metals

Of the various photo-induced phenomena that occur in chalcogenides, the metal-photodissolution effect is probably one of the most useful as far as applications are concerned and is therefore worth describing in more detail. When a chalcogenide layer in contact with a metal layer is illuminated, metal ions dissolve into the glass and migrate through it in the direction of illumination, thereby altering the composition and structure of the glass and changing its physical properties, particularly its solubility and optical constants. This effect is termed 'photodiffusion', 'photodoping' or 'photodissolution', and the structure of the system during the process is shown in Figure 14.4. The most widely accepted explanation of the effect is that it is essentially a light-induced chemical reaction between the metal and the chalcogenide to form a new material, which may be a single compound or a phase-separated mixture. This reaction product separates the metal and chalcogenide layers, so that for the reaction to proceed the new material must be both an ionic and an electronic conductor, i.e. it must allow both the metal ions and the electrons to diffuse through it to reach the product–chalcogenide interface (X–Y in Figure 14.4), where they can continue to

Figure 14.4 A schematic cross-section through a sample during the photodissolution process. Under illumination, metal ions dissolve into the As–S glass to form a new 'photodoped' material. The photodoped layer continues to grow as long as the sample is illuminated. The light stimulating the process is believed to be absorbed in the photodoped layer close to the interface X–Y.
react with the chalcogenide. The process is analogous to the tarnishing of metal films. The purpose of the light is to lower the kinetic barrier preventing the reaction at room temperature, so that the light stimulating the effect is expected to be absorbed near the doped-undoped interface X–Y, which has been confirmed experimentally (Owen et al., 1985; Kolobov et al., 1990).

The effect has been observed in many chalcogenide systems, e.g. As–S, As–Se, Ge–S, Ge–Se, As–S–Te and As–Se–Te glasses, and probably occurs in most amorphous chalcogenides, whether in the form of unannealed vapour-deposited films, well-annealed films or melt-quenched glasses. A variety of metals, metal compounds and alloys have been used as the metal-ion source, e.g. Ag, Cu, Zn, Ag₂S, Ag₂Se, AgCl, AgNO₃ and alloys in the Ag–Cu system. The metallic source is not necessarily an evaporated film, but can be a bulk substrate (Ewen et al., 1983) or even a thin coating produced by dipping the chalcogenide into a solution of a metal compound (Petrova et al., 1984).

The amount of metal that can be introduced into a chalcogenide glass by photodissolution can be considerable, e.g. 30–40 atom% in some cases (Janai, 1981b; Yamaguchi et al., 1982). (For this reason ‘photodoping’ is possibly not an appropriate term since doping generally suggests the introduction of small amounts of extraneous atoms. However, the term ‘photodoping’ is widely used in the literature, the reaction product often being referred to as the ‘photodoped’ material and the initial chalcogenide as the ‘undoped’ glass.) In addition, the metal ions can be driven quite deep into the chalcogenide by photodissolution, for example in the Ag–As₁₆S₉₆Te₄ combination, Ag was found to penetrate at least 20 μm into the glass (Inoue et al., 1974).

One of the characteristics of the photodissolution effect is that the interface between the reaction product and the chalcogenide (X–Y in Figure 14.4) is usually relatively sharp, much sharper than that resulting from the Fickian distribution expected if the metal ions simply diffused into the chalcogenide in the same way as, for example, dopant diffuses into a semiconductor. Also, the metal concentration throughout the reaction product layer is essentially constant (Inoue et al., 1974; Janai, 1981b; Matsuda et al., 1973), so that the metal concentration profile during the process is often termed ‘step-like’. Figures 14.5(a) and 14.5(b) show measured profiles (obtained by electron microprobe X-ray analysis) for Ag in As₂S₃ glass (a) for a photodoped sample and (b) for a thermally doped sample. The step-like nature of the profile in the photodoped case compared with that for the thermally doped material is apparent, but it should be noted that ‘step-like’ does not necessarily imply a perfectly abrupt interface between the photodoped and undoped material, only that the interface is sharper than that resulting from simple diffusion.
PHOTO-INDUCED CHANGES IN CHALCOGENIDE GLASSES

Figure 14.5 Ag and S concentration profiles (relative to As) in Ag-doped a-As$_2$S$_3$ obtained by electron microprobe X-ray analysis (Matsuda and Kikuchi, 1973); (a) is for photodoped material and shows the 'step-like' Ag profile; (b) is for thermally doped material and shows a more gradual variation in Ag concentration in the interface region between the doped and undoped materials.

Photodissolution can be stimulated by photons with energies anywhere between the deep ultraviolet (250 nm) (Polasko et al., 1985) and the infrared (800 nm) (Lis and Lavine, 1983) spectral regions, the exact range of energies depending on the sample geometry and material compositions. (It also occurs under X-ray (Buroff and Rush, 1987; Kolwicz and Chang, 1980; Saito et al., 1988), electron-beam (Bernstein et al., 1988; Singh et al., 1982) or ion-beam irradiation (Wagner et al., 1981).) The chalcogenide–metal combination may be illuminated from the chalcogenide side or, for very thin metal films, the metal side, although in the latter case the maximum thickness of photodoped layer achievable will be correspondingly small. The spectral sensitivity of the process depends on absorption in the chalcogenide/metal for illumination from the chalcogenide/metal side, as well as on the spectral dependence of the basic process. It should be noted that the process can still be stimulated with photons of energy below the band-gap energy for the undoped chalcogenide. The usual sample configuration used in
observing the photodissolution effect consists of a 0.1–1-μm thick amorphous chalcogenide film on (or beneath) which has been deposited a thin metallic layer typically 0.01–0.05 μm in thickness. Illumination from either the chalcogenide or the metallic side with light above or below \( E_g \) for the chalcogenide causes the metal to dissolve rapidly into the amorphous film and migrate through it.

Evidence for the precise structural changes occurring during photodissolution has been obtained by Raman (Firth et al., 1983) and extended X-ray absorption fine-structure spectroscopy (EXAFS) (Greaves, 1990) measurements. It should be noted that photodissolution may also occur in crystals, though less effectively (Imura et al., 1983; Owen et al., 1985).

14.5 Applications of light-induced effects in chalcogenide glasses

14.5.1 Imaging properties of the metal photodissolution effect

Of the various photo-induced phenomena that occur in chalcogenides, one of the most important with regard to applications is the metal photodissolution effect and it is therefore useful to note briefly some of its general imaging properties. The metal photodissolution effect has two important features as far as imaging applications are concerned. Firstly, because the dissolving metal ions do not migrate outside the illuminated area an exact image of the intensity distribution at the surface of the chalcogenide film is created in the interior. Secondly, because such large amounts of metal (e.g. 30–40 atom%) can be driven into the glass, its structure and hence its properties are significantly altered: for example, the difference in refractive index between the doped and the undoped material may be as high as 0.5 and the doped material may be impervious to etchants that remove the undoped chalcogenide. Figure 14.6 shows schematically how images are formed using the effect, and how either embedded or surface-relief structures may be obtained in the film. For most metal–chalcogenide combinations at room temperature the metal ions will dissolve into the glass only under illumination, so that if illumination ceases there is no further movement of the metal ions and no degradation of the stored image. However, at higher temperatures (>175 °C for the Ag–As₂S₃ system) thermally stimulated isotropic diffusion of the metal ions may occur.

14.5.2 VLSI lithography

Resist technology is a key issue in the continuing trend to reduce semiconductor-device dimensions, and the development of new resist
Figure 14.6 Schematic illustration of pattern formation by metal photodissolution using the Ag/As—S system. Ag dissolves into the chalcogenide in the illuminated region but not in the masked areas. Because there is negligible sideways migration of the Ag, a faithful image of the mask is produced in the chalcogenide and the photodoped region has an approximately rectangular cross-section. If illumination is terminated \( (I_o = 0) \) there is no further movement of the Ag and an embedded structure is formed. A surface-relief structure can be produced by removing the undoped material with an alkali etchant.

materials is seen as one of the main routes to advances in this area (Brunsvold et al., 1987). The technology and fundamental science of inorganic resists based on photo-induced effects in chalcogenides (particularly the metal photodissolution effect) have been extensively studied, much of the earlier work being carried out at AT&T Bell Laboratories and at NTT. Reviews of this field have been published by Tai et al. (1982) and Mizushima and Yoshikawa (1982). Interest in these chalcogenide resists as a possible alternative to standard organic-polymer resists was originally stimulated by the fact that they could be used to achieve submicrometre resolution using conventional optical exposure sources (Sze, 1983). Because of their high contrast and high absorption coefficient in the ultraviolet, chalcogenide resists do not suffer problems with standing-wave formation or with matching the modulation-transfer function of the exposure system, unlike earlier organic resists, which
required expensive e-beam or X-ray exposure systems to produce integrated circuits with submicrometre features.

Research in this area has concentrated on resists employing the metal photodissolution effect as the imaging mechanism. Resists based on the many other photo-induced effects that occur in these glasses have also been investigated (Janai, 1981a; Mednikarov, 1984) but seem less promising, in some cases because of poor sensitivity. The most extensively studied chalcogenide resist system based on metal photodissolution is Ag$_2$Se/Ge–Se (Tai et al., 1982), although much work has also been carried out on Ag/As–S resists (Chang et al., 1979; Firth et al., 1985; Kolwicz and Chang, 1980). In terms of characteristics such as resolution, contrast and line-width control, the performance of chalcogenide resists generally matches or surpasses that of organic resists and they are compatible with standard industrial techniques of device fabrication such as plasma-etching and spin-coating. (Working devices have been fabricated using these resists (Mizushima and Yoshikawa, 1982; Utsugi et al., 1984.) Evaporation or sputtering techniques can be used to deposit them as thin, uniform films, which is an important requirement for high-resolution lithography. Contamination problems, with e.g. Ag, can be avoided by using a multilayer resist system in which the chalcogenide resist is used as a thin imaging layer above a thick organic layer which planarizes the wafer topography and acts as a barrier between the wafer and possible contamination sources in the chalcogenide. Such multilayer resist systems also tend to yield the best resolution because they eliminate the effects of substrate topography and reflectivity.

Undoped Ge–Se films are resistant to most of the acid solutions used in silicon processing (e.g. HF, HF–NH$_4$F, H$_3$PO$_4$, HCl and H$_2$SO$_4$) but can be dissolved by alkalis. In contrast, Ge–Se films photodoped with Ag hardly dissolve at all in alkalis, so that an alkaline developer such as an aqueous solution of NH$_3$ can be used to achieve a negative process (unexposed resist removed). The swelling deformation that occurs during the development of polymer resists is not present in these Ge–Se resists. Positive processes (in which the exposed resist is removed) can also be achieved using an additional heat treatment (Mizushima and Yoshikawa, 1982) or other light-induced effects (Mednikarov, 1984). ‘Dry’ development of chalcogenide resists by plasma etching is also possible. In the case of the Ge–Se system, CF$_4$, CHF$_3$ or SF$_6$ can be used as the etchant gas and differential etch rates as high as 500:1 have been achieved (Mizushima and Yoshikawa, 1982).

The theoretical limit to the feature size that can be produced in chalcogenide resists using a standard deep ultraviolet exposure system is ~0.3 μm (Tai et al., 1982), and 0.4 μm lines and spaces have already been produced in these resists over typical wafer topography (Ong and
Hu, 1984). These resists have also been used with excimer laser sources and have demonstrated a resolution capability of 0.2 μm (Polasko et al., 1985) for this type of exposure. Chalcogenide resists are also sensitive to X-rays (Buroff and Rush, 1987; Kolwicz and Chang, 1980; Saito et al., 1988), ion beams (Wagner et al., 1981) and electron beams (Bernstein et al., 1988; Chan et al., 1986; Mizushima and Yoshikawa, 1982; Singh et al., 1982), which are the sources required for very high resolution: feature sizes down to 0.05 μm have been achieved in these resists using X-rays, and down to 0.03 μm using e-beam exposure. For X-ray exposure, chalcogenide resists are expected to achieve better resolution than organic materials because they have a higher density than typical polymer resists.

One reason for the high resolving power of the chalcogenide resists is the so-called 'edge-sharpening' effect they exhibit (Tai et al., 1980). This tends to compensate for diffraction and interference effects in the aerial image and is illustrated schematically in Figure 14.7, which shows the edge of an opaque region of a mask and the corresponding intensity distribution over a Ag₂Se/Ge–Se resist film. Because of diffraction effects some light leaks sideways under the opaque region and so the aerial image is not perfectly step-like. Below the transparent region of the mask the illumination intensity is relatively high and Ag from the Ag₂Se sensitizing layer rapidly photodopes the Ge–Se film. As a result,
in this region the sensitizing layer is depleted of Ag, which causes Ag ions in the weakly illuminated area below the opaque region to diffuse sideways along the Ag₂Se layer to compensate for this deficit. This means in turn that less Ag is available at the edge of the masked region so that photodoping is less efficient there. The effect of the light leakage will therefore be less pronounced than for an organic resist and the edge of the photodoped region correspondingly sharper.

The ultimate resolution capability of a resist is governed by its 'grain' size, which for an organic resist will be of the order of ~100s of Å since these grains are composed of large polymer molecules. The 'grain' size for chalcogenide resists will be smaller than this because these materials are essentially covalent solids, with the largest structural units being only a few 10s of Å in extent. Hence the ultimate resolution capability of the chalcogenide resists will be extremely high, and this has been confirmed experimentally (Mizushima and Yoshikawa, 1982): analysis of transmission electron microscope images of fine metal particles deposited on a Ag₂Se/Ge−Se resist by evaporation indicate that the resolution limit for this system is less than 100 Å.

Another characteristic feature of chalcogenide resists is their high contrast: values of contrast, γ, between 5 and 10 have been reported for conventional ultraviolet illumination (Huggett et al., 1983; Kozicki, 1989; Tai et al., 1980) and greater than 10 for excimer laser exposure (Polasko et al., 1985). High contrast (5 < γ < 9) is also obtained with electron beam sources (Balasubramanyam and Ruoff, 1981; Mizushima and Yoshikawa, 1982; Singh et al., 1982), which is an important advantage for this type of exposure as it enables discrimination against backscattered electrons. A γ value of 3.5 has been reported for X-ray synchrotron radiation (Saito et al., 1988).

The main shortcoming of chalcogenide resists at present is their sensitivity, which for conventional optical illumination is typically 3–5 times poorer than that of organic resists, although good sensitivity (5 mJ/cm²) is achievable with excimer laser exposure (Polasko et al., 1985). Their sensitivity can be improved, though at the expense of additional processing steps (Masters et al., 1980). It is possible, however, that this problem may be overcome in the future by using another of the many chalcogenide–metal combinations exhibiting the photodissolution effect. In the case of e-beam exposure, sensitivities comparable to that of poly methyl methacrylate (PMMA) have been observed (Chen et al., 1986), while for X-ray exposure sensitivities 2–3 times better than that of PMMA have been achieved (Saito et al., 1988). It should be noted that in a production environment, throughput depends not only on exposure time but also on the time spent in wafer-handling and alignment in the exposure system, so reduced sensitivity is not necessarily a serious disadvantage (Utsugi et al., 1984).
Over the period of this research into the lithographic applications of chalcogenide glasses there have been improvements in the performance of organic resists, with the result that some are now capable of submicrometre resolution using conventional ultraviolet exposure systems. Although this means that the chalcogenides may have lost their economic advantage, it is possible that they may yet find uses in either very high resolution lithography, using X-rays, electron-beam or ion-beam techniques, or in certain specialized areas such as photomask production (Mizushima and Yoshikawa, 1982). Research is continuing into both the theory and the practical aspects of chalcogenide resists (Belford et al., 1989; Bernstein et al., 1988; Das and Al-Jishi, 1989; 1990; Hajto et al., 1989), a particularly promising area being the development of an in vacuo lithographic process (Kozicki, 1989). Organic resists, which are solvent-based and therefore have volatile components, are not compatible with this concept. All-vacuum processing is seen as a key step in reducing particulate and gaseous contamination of substrates (Toy, 1989), which is becoming an increasingly important issue as the trend to smaller geometries continues and as new device structures, such as those based on superlattices, become widespread.

14.5.3 Infrared diffractive optics

Diffractive optical elements operate as a result of the diffraction of light rather than its bending or reflection, the best known example being the Fresnel lens or zone plate, which was originally developed in the nineteenth century as a lightweight alternative to planoconvex lenses for use in lighthouse optics but is also used nowadays in other large-aperture applications such as photocopierson and overhead projectors. Apart from reduced weight, these diffractive optical elements can, in many situations, have other advantages over conventional refractive or reflective components, and this is true not only for the visible but also for infrared operation. Because they are essentially planar elements they can be cheaper and easier to manufacture than reflective or refractive components such as Ge lenses. Diffractive elements can be produced to perform the functions of, for example, mirrors, lenses, filters and beam combiners, and they can also carry out beam-forming operations that are difficult to accomplish using conventional elements (Frolova, 1988). Hybrid refractive/diffractive elements may also have uses (Swanson and Veldkamp, 1989). Figure 14.8 shows the diffractive analogues of some optical components. The applications of infrared optics are many and varied and general areas include energy management, thermal fault detection, astronomy, electronic circuit-inspection and night vision (Johnson, 1988).
The range of materials that are suitably transmissive in the infrared is small, most optical elements being made from Ge, Si or ZnSe. Chalcogenide glasses are well-known infrared transmitting materials, having pass-bands (depending on composition) from the visible to beyond 15 μm, and they have been employed to produce infrared windows, filters and fibres (Cimpl and Kosek, 1987). Because the chalcogenides also exhibit many photo-induced effects it is possible to create diffractive elements by using one of these effects to produce surface relief or embedded structures in a chalcogenide film. Photo-induced effects in chalcogenides have been used to fabricate holographic gratings (both phase and surface-relief types) for use at visible wavelengths (Andreish et al., 1986; Bordogna and Keneman, 1977; Kase et al., 1984; Yaji and Kurita, 1983) and are currently being investigated as a technique for producing infrared-diffractive elements (Slinger et al., in press; Zakery et al., 1988). As in the case of the very large scale integration (VLSI) lithographic applications, of the various photo-induced phenomena that occur, the metal photodissolution effect is probably the most useful as far as grating applications are concerned, because it produces the largest
change in the properties of the chalcogenide, particularly the refractive index and etch resistance.

In the case of the Ag/As—S system, both photodoped and undoped As—S films are transparent over the range 2–12 µm and the difference in refractive index between the Ag-doped and undoped material over this wavelength range can be as high as 0.5, so that phase gratings with high modulation are achievable. Phase gratings have an embedded structure and this is created automatically by photodoping. Surface-relief gratings can be produced by removing the undoped chalcogenide in an alkali etchant. Theory suggests that near-band-gap illumination is the most efficient for producing the deep structures required for infrared operation: if the photon energy is too high, the light is absorbed in the undoped chalcogenide before it can stimulate photodissolution; if the photon energy is too low, the photodissolution process proceeds very slowly. For the As—S glasses near As$_2$S$_3$ this means that the exposing wavelength should be in the green, and the 514.5 nm line of the Ar laser is therefore an ideal source.

In general, for efficient operation these infrared elements have to be relatively thick (of the order of several micrometres at least), which is in contrast to the VLSI applications, where much thinner films (~0.2 µm) are used. Coupled-wave theory can be used to predict the required thicknesses for different types of grating element at different wavelengths of operation (Slinger et al., in press) and typical values are given in Table 14.1. Chalcogenide films can be deposited using a variety of techniques (including thermal evaporation, sputtering, spin-coating, plasma deposition and hot-pressing) and thicknesses of up to at least 100 µm are achievable. However, the depth to which metal ions can, in general, photodope chalcogenide films has not been established. In the case of the Ag/As—S system, photodoping to a depth of 4 µm has been observed (Owen et al., 1990) and it is likely that photodoping to a depth of at least 6 µm is possible, so that most of the structures in Table 14.1 are obtainable with this system. As mentioned earlier, Ag has been observed to photodissolve to a depth of at least 20 µm in As$_{16}$S$_{80}$Te$_4$, and if this is typical of most chalcogenides then even the 23-µm structures required for bulk gratings at 10.6 µm operation should be

<table>
<thead>
<tr>
<th>Grating thickness $d$ (µm)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Bulk</td>
<td>1.39</td>
<td>8.77</td>
</tr>
<tr>
<td>Surface relief</td>
<td>0.48</td>
<td>3.0</td>
</tr>
<tr>
<td>Blazed zone plate</td>
<td>0.38</td>
<td>2.4</td>
</tr>
</tbody>
</table>

Table 14.1
achievable. Because the photo-dissolution process is essentially diffusion-controlled the required exposure time increases as the square of the grating thickness, so that long exposures may be required to produce very thick structures. However, several possibilities exist for significantly increasing the speed of the photodissolution process, such as the simultaneous application of heat (so-called photothermal doping).

Coupled-wave analysis of the likely performance of bulk and surface-relief structures typical of those capable of being produced by photodoped chalcogenides (Slinger et al., in press) suggests that high diffraction efficiencies (>80%) should be achievable in the infrared.

14.5.4 Optical mass memories

Because the density of information storage in optical media can be up to two orders of magnitude better than in magnetic discs, there has been extensive research into materials capable of optical recording. Chalcogenide films have potential in this area and have been investigated as optical recording media for mass-memory applications such as data discs for video recording or computer information storage (Madan and Shaw, 1988). In these applications the basic mechanism used to record information is generally either photo-induced crystallization of an amorphous film (alternatively, amorphization of a crystalline film) (Jung et al., 1989; Matsushita et al., 1987; Uchida et al., 1989; Watanabe et al., 1983; Yagi et al., 1987; Yamada et al., 1987) or light-induced morphological changes such as the generation of bubbles or holes in the film (Vriens and Jacobs, 1984). Both of these mechanisms are essentially thermally induced, the heating being provided by the optical excitation, which is usually a focused laser beam.

In the case of the photo-induced phase-transition mechanism, the optical constants of the amorphous and crystalline material are sufficiently different that discrimination between the exposed and unexposed regions of the film can be achieved by monitoring reflectance. In a typical system (Watanabe et al., 1983), light from a diode laser is focused to a spot a few micrometres in diameter: to record, the laser power is increased above the recording threshold for the material (usually a few tens of milliwatts), while to read data, the power is reduced below this threshold and the reflected light from the beam spot is collected and measured. Of the various chalcogenides exhibiting the effect, Sb$_2$Se$_3$, Sb$_x$Te$_{1-x}$ (particularly Sb$_2$Te$_3$) and Te$_x$(Ga$_2$Se$_{1-x}$)$_{1-x}$ glasses are among the most suitable found to date. Optical recording of frequency-modulated video signals has been demonstrated in Sb$_2$Se$_3$ films used in conjunction with a reflective layer of Bi$_2$Te$_3$, which increases energy coupling into the photo-sensitive layer (Watanabe et
PHOTO-INDUCED CHANGES IN CHALCOGENIDE GLASSES

al., 1983). In these experiments the recording material was deposited on a PMMA disc substrate, and the recording threshold was <5 mW at a rotation speed of 1800 r.p.m. This memory effect is reversible and so can be used in either write-once or erasable recording: erase times of 0.5 μs and reversibility >10⁴ write-erase cycles have been realized in Te₀.₈(Ga₀.₀₅Se₀.₉₅)₀.₂ doped with 5 wt% Ge (Matsushita et al., 1987); GeSb₂Te₄ exhibits an erase time of 50 ns and reversibility of >10⁵ cycles (Yamada et al., 1987). In the case of the pseudo-ternary alloy system GeTe–Sb₂Te₃–Sb, a reversibility of >10⁶ cycles has recently been reported (Uchida et al., 1989).

Of the morphological mechanisms, the most important is probably the creation of well-defined holes in thin films of chalcogenide material (Vriens and Jacobs, 1984). Uniformly shaped holes with clean edges and dimensions of the order of 1 μm can be generated in a variety of chalcogenides (for example As₂₀Te₈₀, Ge₁₀Te₉₀, As₁₅Se₄₅, Ge₁₀Se₉₀ and Sb₂₀S₈₀) by using a focused laser beam. The beam is believed to vaporize the centre of the irradiated region and to soften the remainder so that material is pulled to the edges of the exposed area by surface tension.

14.5.5 Integrated optics

Optical fibres are now replacing copper wire as the transmission medium in telecommunications because of their significant advantages in terms of data rate, cost and information-carrying capacity. However, the performance of optical-fibre communication networks is limited, not by the properties of the fibres, but by the speed of the electronics used in switching, routing and signal processing. To exploit the full potential of these optical-fibre systems new optical devices and circuits must be developed to process light signals directly. Such devices require materials with non-linear optical properties, that is materials in which the optical constants can be changed by intense optical radiation. The interconnects in these optical circuits will generally be waveguides and fibres.

To minimize attenuation in optical fibres, the wavelength of the light used in optical communications must be in the infrared: typical commercial fibres operate in the range 1.3–1.6 μm at present. For ultra-long-distance communications, attenuation losses must be reduced still more, which implies shifting the wavelength of operation further into the infrared. The use of chalcogenide glasses as materials for infrared-transmitting fibres and waveguides has been extensively studied and has been reviewed by Andreish et al. (1986). As these applications are covered in detail elsewhere in this book (chapter 13) they will not be discussed
further here, except to note that as well as interconnecting chalcogenide devices, chalcogenide films can be used as coupling elements between active and/or passive optical components made in other materials such as silicate glasses and LiNbO₃.

A wide variety of chalcogenide integrated optical components has been studied, including filters, mixers, deflectors and demultiplexers. (As many of these components contain grating structures, there is some overlap between research in this area and the work being carried out on infrared diffractive optics in chalcogenides.) Figure 14.9 shows a schematic diagram of a demultiplexer fabricated using the metal photodissolution effect (Andreish et al., 1986). Single- and double-throw optical switches have been made in As-S waveguides (Tanaka et al., 1985), based on the reversible photo-induced optical-constant changes described in section 14.3.1. Although these devices have long switching times, it is likely that much faster switches can be produced based on the non-linear optical properties of these materials. GeS₂ and As₂S₃ are known to be highly non-linear, the latter having a non-resonant third-order electronic susceptibility of $2.2 \times 10^{-12}$ esu and a response time of less than 10 ps (Hall et al., 1989; Nasu et al., 1989).

Finally, it should be noted that much of the technology required for the fabrication of optical integrated circuits in chalcogenides has already been developed: it is essentially identical with that used for the chalcogenide resists employed in VLSI lithography. Indeed, the processing requirements may be less stringent than those for VLSI circuits since the dimensions of optical-device features will generally be of the order of the light wavelength, which for infrared operation will be greater than 1 µm. In this case, conventional optical exposure techniques can be used in patterning the chalcogenide film.

14.6 Conclusions

As a result of extensive research over the past three decades, considerable information now exists on the properties and applications of

![Figure 14.9](image-url)  
Figure 14.9 Schematic diagram of a demultiplexer made using the metal photodissolution effect (Andreish et al., 1986).
chalcogenide glasses. They are generally 'well-behaved' materials and can be made in bulk form or deposited, by a variety of techniques, as films with thickness ranging from about 100 Å to 100 μm. The chalcogenides exhibit a wide diversity of reversible and irreversible photo-induced effects which makes them unique amongst glass-forming families and leads to many potential applications in the area of optical imaging or recording. Because such a large number of chalcogenide systems exist there is considerable scope for optimizing the performance of the imaging and recording media based on these photo-induced effects. Research is actively continuing in this field, and with the growing interest in 'nanotechnology' and 'photonics' there is considerable scope for applications of the light-sensitive properties of these materials.
HIGH-PERFORMANCE GLASSES


Chapter 14


REFERENCES


fine-structure spectroscopy study of photostructural changes in amorphous arsenic chalcogenides. Phil. Mag., B54 483–490.


Singh, B., Rajagopolan, S. and Chopra, K. L. (1980b) Photoinduced chemical changes in


References added at proof

The following extensive reviews on photodarkening and photodissolution in chalcogenide glasses appeared whilst chapter 14 was in preparation:


Chapter 14
Electronic Switching in Amorphous-Semiconductor Thin Films

J. Hajto, A. E. Owen, and A. J. Snell
Department of Electrical Engineering,
University of Edinburgh, Scotland, U. K.

P. G. LeComber and M. J. Rose
Department of Applied Physics and Electronic & Manufacturing Engineering, University of Dundee, Scotland, U. K.

The aim of this chapter is to summarize experimental and theoretical work on electrical instabilities such as electronic switching and memory phenomena in amorphous thin films and particularly in amorphous silicon thin-film structures. Section 14.1 is a general survey of the field, comparing and contrasting experimental observations and theoretical models of switching in a variety of amorphous thin films, including a brief review of the early work on amorphous silicon. In Section 14.2 recent work in the authors' own laboratories on digital and analog amorphous silicon switching devices is described and discussed in more detail.

For the present purposes, electronic switches are two-terminal solid-state devices that can be changed from a nonconducting OFF state to a conducting ON state by an appropriate electric signal. Electronic switching is associated, explicitly or implicitly, with negative differential resistance (NDR). Two types of NDR exist: current-controlled NDR (CCNDR) [1] shown in Figure 14.1(a) and voltage-controlled NDR (VCNDR) [2] shown in Figure 14.1(b). Conduction in CCNDR involves the creation of conducting filaments in which the current density differs from that of the surrounding material [3] or circuit-controlled oscillations [4]. VCNDR is associated with stationary [2] or steady-travelling high-electric-field domains [5], or with sustained circuit-controlled oscillations [6]. Both CCNDR and VCNDR can give rise to a range of switching phenomena.
14.1 GENERAL SWITCHING PHENOMENA IN AMORPHOUS SEMICONDUCTORS

14.1.1 Threshold and Memory Switching

Many examples of threshold and memory switching have been reported in initially homogeneous thin films of a variety of materials including simple oxides [7], transition-metal-oxides [8], elemental boron [9], metal-semiconductor-metal structures [10], and Langmuir-Blodgett films [11]. Interest in amorphous semiconductor films was stimulated by the discovery of reversible switching in certain amorphous semiconductors [12]. From 1968 onward, interest concentrated primarily on a class of covalently bonded alloys of group IV, V, and VI elements called the chalcogenide glasses. These materials, with appropriate electrodes, exhibit a transition from a low conductance (OFF) to a high conductance (ON) state under the influence of sufficiently high fields.

To illustrate the switching characteristics that are often observed, we shall first describe some typical results for the chalcogenide glasses. A typical switching device consists of a thin film of chalcogenide glass (thickness approximately 0.1–10 μm) sandwiched between two electrodes of refractory metals. Conduction is ohmic up to fields of about $10^4$ V/cm. At higher fields, nonohmic processes become evident and the current rises exponentially with applied voltage. Switching occurs at fields of about $10^5$ V/cm and schematic current-voltage (I-V) characteristics are shown in Figure 14.2(a). Upon switching (at the threshold voltage $V_{th}$), the voltage across the device drops sharply along the load line until a holding voltage $V_h$ of about 1V is reached. Experimental results show that conduction in the ON state is filamentary in character [13]. The device may be maintained in the ON state as
Figure 14.2 Schematic representation of (a) threshold switching and (b) memory switching.

long as the current does not drop below a critical holding current $I_{H}$ and if that is not maintained the device switches back to the low-conducting OFF state. The switching process is highly reproducible, reversible, and essentially independent of polarity. Devices of this kind are called threshold switches; they are nonpermanent, or volatile, and always revert to the OFF state in the absence of an appropriate bias.

The schematic $I$-$V$ characteristics for a different type of switching are shown in Figure 14.2(b). There is again a critical switching voltage for the OFF to ON transition, but both ON- and OFF-state characteristics extrapolate through the $I$-$V$ origin. In these devices, if the high ON-state current is maintained for about a millisecond after the OFF to ON switching, the material along the current filament is modified and a high density of crystallites forms, creating a permanent bridge of high-conductivity material between the two electrodes [14]. After this modifi-
cation of the original structure has occurred, the device remains in the low-resistance ON state even after the electric field is removed. It is possible to return the device to its high-resistance OFF state by applying a sufficiently high, short, current pulse. This “erase” pulse melts the high conductivity material in the conducting channel. Subsequent rapid cooling of the melt restores the original amorphous phase. Devices of this kind are permanent (nonvolatile) and hence are known as memory switches. The amorphous-to-crystalline transition outlined above is, of course, only one possible mechanism by which memory switching can be implemented. In general, memory switching refers to any mechanism that results in the ON and OFF states persisting more or less indefinitely in the absence of any bias.

To summarize, switching phenomena can be characterized into two main categories:

1. Threshold switching, in which continuous electrical power is required to maintain the highly conducting ON state;
2. Memory switching, in which both ON and OFF states can be maintained without electrical power.

It is common to describe switching phenomena in terms of ON and OFF states, but as we will show in Section 14.2.2, switching devices are not always rigidly bistable in their operation, that is, they are not always digital devices. In some cases, a continuous range of intermediate states is observed between the ON and OFF states, giving an analog memory effect.

The speed of the switching transients varies from device to device and covers a range from $<10 \text{ ns}$ to $>1 \text{ ms}$. The actual switching is extremely fast (nanoseconds) but is preceded by a delay time of the order of microseconds near the threshold. The delay time decreases exponentially with the overvoltage, that is, with the voltage above the threshold. Figure 14.3(a) shows a simple circuit that can be used to measure the voltage and current waveforms during a single switching pulse to obtain the speed of the switching transient, and Figure 14.3(b) shows schematic waveforms of the measured voltage $V_s$ and current $I_s$ across the device during an OFF to ON transition (associated with either threshold switching or a memory “write-in” operation). There is a sudden rise in the voltage across the device at time $t_0$ but the current remains low (i.e., the sample is in its low-conducting OFF state) for a certain time termed the “delay time” ($t_{\text{DW}}$), equivalent to $t_1 - t_0$. At $t_1$, there is a sudden rise in the current and at the same time a similarly fast decrease in the voltage associated with the OFF to ON transition (i.e., a sudden increase in the device conductivity). The duration of the ON to OFF transition ($t_2 - t_1$) is called the “switching time,” and it is usually much faster than the delay time $t_{\text{DW}}$. Figure 14.3(c) refers to the erase operation for a memory device. This situation obviously does not arise in a volatile threshold device. When the voltage pulse is applied at time $t_0$, the current is high at first, as the sample is in its ON state. After
a delay time $t_{DE}$, the sample conductance decreases rapidly during the erase switching time ($t_2 - t_1$), and at the same time the voltage across the device increases.

Most chalcogenide devices, whether threshold or memory, can sustain many hundreds of cycles of ON→OFF→ON transitions. The best threshold devices can switch up to $10^6$ times before they fail. Good memory devices will operate for at least $10^6$ WRITE-ERASE cycles before failure.

14.1.2 The Forming Process: A Precursor to Threshold and Memory Switching

Newly fabricated devices rarely show threshold or memory switching effects without an initial modification of their as-deposited structure, a process usually called forming. Forming is achieved by the application of suitable voltage pulses that are always higher in magnitude than the subsequent programming pulses. This invariably produces an irreversible change in the electrical characteristics of a device, often with a substantial decrease in the overall terminal resistance.

The processes involved in forming depend on the type and quality of the thin film, the geometrical structure of the device and in some cases the electrode material. The changes can be either physicochemical, involving a redistribution of the
constituents, or they can be electronic, wherein a quasi-permanent change in the occupancy of some electronic states takes place. The changes can occur throughout the bulk of the film or in a localized region. The most common localized effect is the formation of a filament of highly conducting material that extends completely or a part of the way through the device. Such permanent filament formation is a consequence of temporary filamentary breakdown usually observed in metal-insulator-metal and metal-semiconductor-metal sandwich structures under double-injection conditions [3]. Barnett and Milnes have considered current and voltage instabilities in semiconductors and have shown that with current-controlled negative resistances, such as those expected in double-injection structures, energy considerations imply the formation of a filament. The filament develops first at some device inhomogeneity, is stationary, and grows in size about the nucleation region as the current is increased. This filament should not be confused with the moving transverse wave of a voltage-controlled instability (Gunn effect), which was also considered by Barnett and Milnes. Rather earlier, Ridley [15] had also shown by general thermodynamic arguments that CCNDR implies the formation of current filaments, while field domains are involved in VCNDR.

Filamentary conduction has been experimentally observed in a wide range of materials and sandwich structures. These include single crystal silicon [16], gallium arsenide [17], zinc telluride [18], compensated germanium [19], and polycrystalline silicon [20], all of which show I-V characteristics associated with threshold switching of the kind shown in Figure 14.2(a). In these cases the filament collapses and disappears as the voltage is removed, and there is no permanent change in the device structure or in its electrical characteristics.

In the case of amorphous thin-film structures, however, the development of the filament during the forming process is always accompanied by an irreversible modification of the original structure. This means that the filament—or part of the filament—is permanently "written in" to the originally uniform amorphous structure (i.e., it will not disappear after the voltage is removed). Permanent filament formation can occur by crystallization of an amorphous film [14], stoichiometric changes [21], diffusion of the electrode material into the film [7], or ionization of deep traps [22]. All of these changes usually refer to a localized modification of the structure in the area of the filament. Reports of bulk forming effects are much rarer, one example being bulk diffusion of electrode material into the film [23].

The modification of the initial structure during the forming process is the most important factor determining the subsequent switching operation. This is because in most cases forming creates a new device within the originally deposited structure, whose characteristics determine the ensuing threshold-switching or memory-switching operation. For example, a permanent filament may have a cross-sectional area many times less than the device area, but its effective conductance is often so much higher than that of the surrounding material that it becomes the dominant conducting path.
It is important to emphasize that in both the threshold- and memory-type devices, the initial switching mechanism appears to be the same [24]. It is often initiated by field-dependent nonohmic conductivity and consequent instability. Whether what follows is threshold switching, memory switching, or, in some cases, destructive electrical breakdown depends on the properties of the material and the presence or absence of suitable feedback in the system.

14.1.3 The Mechanism of Threshold Switching

The detailed \( I-V \) characteristics of chalcogenide-based threshold-switching devices have been widely investigated since the original publication by Ovshinsky in 1968 [12], and a general survey of the most significant results has been provided by Adler, Henisch, and Mott [25]. The suggested physical models can be divided into two categories; the first assumes that threshold switching is due to thermal processes [26], and the second considers that it is associated purely with electronic phenomena [27, 28]. Thermal or, more generally, combined thermal and electrical mechanisms are sufficient to explain threshold switching. In the thermal models, the increase in conductivity under the influence of an electric field exceeding the switching threshold \( V_{th} \) is controlled only by Joule heating by a "thermal runaway." This is regarded as the principal driving force of the switching mechanism. The explanation of threshold switching in terms of thermal processes involves a feedback loop, illustrating how, in materials with a thermally activated conduction process, thermal runaway can occur and cause a sharp drop in the device resistance. The feedback loop is:

- high field $\rightarrow$ increased current $\rightarrow$ increased power dissipation
- greater conductivity $\leftarrow$ internal temperature rise

The essential link is the sensitivity of the conductivity to temperature. Whether the final result is threshold switching or memory switching, the conditions leading to the electrical instability can usually be formally described by an energy-balance equation [26]:

\[
A(T, E, \alpha) = B(T, \alpha)
\]  

(14.1)

where \( A \) is the rate at which energy is gained from the field \( E \) at temperature \( T \), and \( B \) is the rate at which it is absorbed or dissipated. The parameter \( \alpha \) is introduced to denote any other relevant property of the current carriers depending on the particular experimental conditions. The variation of the two sides of this energy-balance equation will be generally as shown in Figure 14.4, in which the abscissa
represents some appropriate parameter such as temperature, energy $\xi$, or injected charge $Q$. The rate of gain of energy from the field will normally be given by

$$A = \sigma(T, E)\xi^2$$

(14.2)

or some variant of this. The conductivity $\sigma(T, E)$ is generally a function of temperature and field. With a sufficiently large conductivity that is temperature-dependent, the material will heat up through Joule heating, and, if thermal processes are dominant, then

$$B = \left(\frac{\rho C}{K}\right) \frac{\partial T}{\partial t} - \nabla^2 T$$

(14.3)

The characteristics of purely thermal threshold switching can be accurately predicted from the solution of (14.2) and (14.3) in a one-dimensional form with conductivity expressed only as a function of temperature, that is, $\sigma = \sigma(T)$. The general solution is shown in Figure 14.4; the abscissa in this case represents temperature and the intersection of $B$ and field $E_2$ determines the point at which instability sets in. It should also be noted that even in the case when the isothermal conductivity is ohmic, the dynamic conductance will not be constant because of
Joule heating. The most detailed calculations of the thermal, one-dimensional model have been reported by Warren [26]. The general conclusion, summarized more fully at the end of this section, is that such a simple model is not sufficient to explain switching in relatively thin films (e.g., <8 μm), although the evidence shows it is adequate to account for the switching characteristics of thicker films.

Kroll and Owen et al. have provided the most detailed mathematical description of threshold switching phenomenon in amorphous semiconductors based on a combined thermal and electrical model (i.e., $\sigma = \sigma(T, E)$) [29, 30]. In Kroll's calculation, the $I-V$ characteristics were obtained from very general mathematical considerations concerning stability and bifurcation of solutions of the nonlinear equations for temperature and field, and he showed that switching occurs by the nucleation and growth of a hot spot approximately in the center of the device interior. The primary results of Kroll's analysis are shown in Figure 14.5. The model predicts the formation of the S-shaped CCNDR associated with high-current filaments (Figure 14.1(a)). The existence of a region of negative differential resistance on the $I-V$ characteristics was determined by radially uniform solutions of the thermal balance equation. It was also shown that the dashed portions of seg-

![Figure 14.5 $I-V$ characteristics: solid curve, stable solutions; dashed curve, unstable solutions; dotted line, load line. From: [29].](image)
ments (b) and (c) are locally unstable against temperature and field perturbations which keep the total current through the device constant, and that branches (a) and (d) are stable with respect to such perturbations. Switching is initiated at some point not too far below turnaround (see point 1 in Figure 14.5) as the result of a macroscopic critical fluctuation. Once the switching is initiated, the device discharges through the embryonic channel and switches along the load line to the stable vertical portion of branch (d). Here all the current flow is carried in a hot channel. The possibility of thermally induced negative differential conductivity was suggested and described in other publications [31, 32]. Experimental evidence of high filamentary temperatures during threshold switching has also been reported [33].

Several authors have put forward alternative nonthermal electrical models for creating conditions of instability leading to switching. One of the most frequently used electrical models is based on double injection with recombination in which, in the ON state, carriers are injected at both electrodes, giving rise to a high density of carriers in the valence and conduction bands. In this case the energy input is either stored in the dielectric as injected charge $q_{in}$ or is lost through recombination. An energy-balance equation (14.1) could therefore be written as

$$\sigma(T, E)\varepsilon^2 = \frac{q_{in}}{D\varepsilon \varepsilon_0} \frac{dq_{in}}{D\varepsilon \varepsilon_0} dt - j_{re} \Phi_m$$

(14.4)

where $\varepsilon$, is the relative permittivity, $D$ is a suitable geometrical factor, and $j_{re}$ is the recombination current driven by some appropriate internal potential $\Phi_m$. In this case the abscissa of Figure 14.4 would represent $q_{in}$. A qualitative explanation along the above lines was suggested first by Henisch [27]. Similar but more quantitative models have also been proposed by Lucas [28] and Mott [34]. The conductive ON state is sustained by double injection, provided the applied voltage exceeds the mobility gap. It is assumed that the tails of localized states (typically present in amorphous semiconductors [35]) extend from the conduction and valence bands and overlap somewhere near the center of the energy gap (or mobility gap). This provides a completely compensated set of positively and negatively charged states above and below the Fermi level pinned at or near the center of the gap. These features are inherent in the Cohen-Fritsche-Ovshinsky (CFO) model of the electronic band structure of chalcogenide glasses [36] and are further supported by more recent theoretical interpretations involving the concept of negative-U states [35]. The injected electrons and holes will recombine with and neutralize the positively and negatively charged states, setting up a negative and a positive space charge adjacent to the cathode and anode, respectively. The regions of space charge will limit the current flow in the vicinity of the electrodes, and the field will be redistributed in a way that would be decreasing near the electrodes and increasing
in the center. As the applied voltage is increased, more charge is injected and the space-charge regions grow until eventually they meet and overlap. The physical situation then rapidly becomes unstable. When the space-charge clouds overlap they neutralize each other, causing the field in the interior to collapse and allowing the same or larger current to flow with a lower applied voltage, causing a negative resistance characteristic. The conductivity in the center increases and the field decreases, while the field near the electrodes increases. Electrons and holes are accelerated rapidly across the neutral region and, because of the increased field, injection of carriers at the electrodes increases. Both effects increase the rate at which space-charge overlap occurs (i.e., there is a positive feedback and hence an unstable situation). The stable state, corresponding to the ON state, is achieved when the space charge has spread right across the structure and the bands are practically flat. Schottky-type barriers are established at the electrodes, but because of the high density of traps and because they are thin (about 1 nm), electrons can easily tunnel through from the Fermi level of the metal into the conduction band of the chalcogenide glass and, similarly, holes from the anode tunnel into the valence band. Because the space charge has been neutralized and the traps are filled, the conductivity in the device is now high (i.e., the drift mobility is no longer limited by trapping).

Homma [37] published a systematic and detailed comparison between the threshold switching properties of a chalcogenide material (Te0.25As0.75GeSi1) and a nonchalcogenide alloy (Cd1.87Ge0.87As2) of nearly equal forbidden gap but higher conductivity. The results showed that two alloys of rather different resistivities have the same threshold voltage and also that symmetrical threshold voltages may be associated with highly asymmetric threshold currents. Furthermore, in both cases, the threshold current is light-sensitive (both materials are photoconductive) but the threshold voltage is not but remains practically unchanged by illumination despite a power increase by a factor of three or more. This indicates that a critical field is involved in the circumstances and determines the onset of switching. These results provide evidence against thermal interpretations and thereby support electronic models. Microwave noise and transient ON-characteristics (TONC) measurements [38] also suggest that the Joule heat is not the main cause of the formation of the current filament. Also, relaxation processes and polarization effects [39] observed at low temperatures show an effect on threshold switching opposite to that expected from the thermal theory. It is found that if, during double-pulse experiments, the first pulse is of insufficient height to cause switching, the magnitude necessary for the second pulse to cause switching is lowered if the pulses are of the same polarity and increased if they are of opposite polarities. It appears that contact, injection, and trapping effects must play some role in explaining certain of the asymmetric effects described above. Therefore, although the electrothermal theory describes well the observed initial switching instabilities, it is also clear that contact, injection, and trapping effects must play some role in explaining certain
of the asymmetric electrode effects and polarization and relaxation phenomena described above.

Many other conditions of electrical instability leading to switching have also been investigated. For intrinsic electronic breakdown in the single-electron approximation, for example, the function $A$ in (14.1) is of the form [24]

$$A = e\mu(\xi)E^2 = \frac{e^2E^2\tau(\xi)}{m}$$

(14.5)

where $\mu$, the mobility, and $\tau$ the relaxation time, are functions of energy. The electron loses energy to the lattice through phonon interactions, and the value of $B$ (14.1) can be written as

$$B = \frac{h\omega}{\tau(\xi)(2N_{\omega} + 1)} \left[ 1 - \frac{kT}{\xi} \right]$$

(14.6)

where $\tau$ is the mean time between collisions and is a function of energy, and $N_{\omega}$ is the average number of phonons of frequency $\omega$. Thus, in Figure 14.4, the abscissa represents energy $\xi$, and the field $E_2$ again corresponds to the onset of instability.

Electron-electron collisions may also lead to electrical instabilities [24]. With a high collision rate and a rate of energy exchange due to electron-phonon interactions, the electron energy-distribution function is Maxwellian with a mean temperature $T_\text{e}$ above the ambient temperature $T$ or the lattice temperature $T$. This means that the energy rate of change terms $A$ and $B$ (14.1) can be averaged to $\overline{A}$ and $\overline{B}$, having a similar functional dependence so that the general picture of Figure 14.4 remains valid at least up to energies $\xi$, at the maximum of curve $B$. This collective description has a strong analogy with thermal runaway, and the abscissa corresponds to the electron temperature $T_\text{e}$. In this case, therefore, it is the steady-state electron temperature that rises until, at a critical value, no equilibrium is possible and instability sets in. Near $\xi$, however, the electron-phonon energy transfer increases, implying an increase in the internal temperature of the material, with consequential thermal effects.

Impact ionization and avalanche breakdown can also produce instabilities [22]. If the electron density is low, a high-energy electron may collide with an atom instead of another electron and thus ionize it, producing a hole and two low-energy electrons. The two electrons, in turn, will be accelerated to high energies and ionize more atoms; thus the process can cascade to cause an electron avalanche.

Summarizing these results, there is no doubt that, besides thermal effects, there are significant electronic effects causing nonohmic conduction and CCNDNR in the preswitching region of chalcogenide glass switches. A comparative study of their relative importance has been given by Owen and Robertson [24] and Adler...
The main point of discussion has concerned the extent to which switching can be attributed to thermal or electronic processes.

In either case (thermal or electronic), heating may be the initial process, but the critical field $E_c$, at which switching occurs may behave differently with different thicknesses of the films. For thin films ($< \sim 1 \mu m$) conduction at breakdown is nonohmic. Therefore it is expected in this case that $V_s$ will be independent of the film thickness and will decrease linearly with temperature \([34]\). The nonohmic field-dependent conductivity implies electronic processes such as double injection or avalanche breakdown, as discussed above. Furthermore, the constant value of the threshold voltage in the presence of illumination \([37]\), the microwave noise and TONC measurements \([38]\), and the polarization effects \([39]\) also provide evidence for electrical mechanisms. Based on these experimental results, there are strong reasons for believing that in chalcogenide thin-film ($d \leq 1 \mu m$) threshold-switching devices ("Ovonic switches" \([12]\)), thermal considerations do not play a significant role.

However, for thick films ($\geq 8 \mu m$) for which current at breakdown is ohmic, $E_c$, should be proportional to the inverse of the thickness and decrease exponentially with increasing temperature \([34]\). Kolomiets \([40]\) found a dependence on the thickness showing that in thin films ($\leq 8 \mu m$), the breakdown field was independent of the thickness and depended only weakly on temperature, suggesting that the switching mechanism in thin films is not associated with simple thermal breakdown. On the other hand, in thick films ($\geq 8 \mu m$), the breakdown field is inversely proportional to the thickness and shows strong temperature dependence, confirming the essentially thermal nature of the breakdown in thicker films. Stocker \([41]\) described a quantitative model for threshold switching in semiconducting films $\geq 8 \mu m$ thick based on thermal considerations only (i.e., no field dependence of conductivity) and has shown that the switching behavior of these thicker films can be described adequately by a simple thermal model.

The most useful approach seems therefore to be to compare, wherever possible, the functional dependence of parameters such as the threshold voltage and delay time on, for example, temperature, field, and geometry with the predictions of the related models. With several possible mechanisms competing, the dominant process can change drastically as a function of such factors as temperature or the geometry of the specimen.

14.1.4 Mechanisms for Memory Switching

As described in the previous paragraph, the precursor to memory phenomena in thin ($\sim 1 \mu m$) film is the onset of nonlinear field-dependent conductivity leading to instabilities and threshold switching. For threshold switching, the material must be capable of carrying a much increased current, either uniformly or locally in a
filament, but reversing spontaneously to the original nonconducting OFF state when
the holding voltage is removed. For memory switching, the dielectric material must
be capable of changing into a permanent conducting state in some way—an overall
or localized change in the atomic or microscopic structure for example—but one
that can be reversed to the OFF state by a suitable current (energy) pulse. Obviously
the system must also be able to absorb the reversing pulse without destructive
breakdown. Alternatively, the memory-switching process may be based on charge
storage (i.e., without a modification of the structure).

Models for memory switching can therefore be divided into two broad cat-
egories: electronic or structural. The first of these is based on the long-term storage
of charge, with no major structural modification, to account for the nonvolatile
nature of the switch. The most commonly proposed charge storage sites are traps,
either in the bulk or at an interface between two dissimilar materials. The necessary
characteristic of such traps is that they have a release time comparable to the
retention time of the memory, which may range from a few hours to many years.
If the charge is stored in the amorphous thin-film sandwich structure, it can affect
the conductance in a number of different ways. For example, when stored in the
bulk or at interfaces, it can cause band bending, which in turn modifies the con-
ductance. Simmons and Verderber have used an electronic model to explain their
observations of memory switching in thin silicon monoxide (SiO) films fitted with
gold electrodes [2]. The I-V characteristics of their device (Figure 14.6) show a
VCNDR, and the device could be switched between several memory states (A, B,
C, D) that persisted for a few weeks. The forming process in these devices is due
to gold ion injection into the bulk of the silicon monoxide film, the source of the
ions being the positively biased gold electrode. The sample cannot be formed at
liquid nitrogen temperatures but forms more readily at elevated temperatures,
indicating that ion injection with subsequent ion migration is the origin of the
forming process. The current-voltage characteristics and the weak temperature
dependence of the conduction process suggest that the main factor determining
the current flow in this system is tunneling. For a reasonable current flow in the
system, adjacent ionic sites must be positioned within ~30Å of each other, which
means that the injected ion density must be at least of the order of 10¹⁹ cm⁻³. The
authors concluded that in these devices the conduction and memory processes are
electronic in nature and that the memory effect is due to charge trapping in the
bulk of the silicon monoxide, which in turn modifies the conductance of the metal-
silicon monoxide barrier.

Hovel and Urgell have also suggested an electrical model to explain their
observations of memory switching in epitaxial ZnSe films grown on crystalline Ge
[42]. In the OFF state the current flow through the device is limited mainly by the
high-resistance ZnSe layer. In the ON state, the barrier at the ZnSe/Ge interface
has been narrowed by the ionization of traps at the interface from neutral to
positive, thereby permitting easy tunneling through it. The same general type
of phenomenon has been observed in ZnSe-GaAs [43], GaP-Ge [43], AlN-Si [44], and SbSi-SnO₂ heterojunctions [45].

Switching in chalcogenide glass memory devices is based on a rearrangement of the structure in the ON and OFF states, as seen in Figure 14.7. During initial forming pulses, a filament of material crystallizes in the amorphous structure [14]; this normally happens near the center of the device where the highest temperatures are attained. This modified material has a higher conductivity than the bulk and therefore becomes the preferred current path during subsequent voltage pulses. In the erase operation (i.e., switching from ON to OFF), large voltage pulses with steep trailing edges are applied. These melt the crystalline filament, causing it to solidify in an amorphous form. To write in again, switching from OFF to ON, pulses with more gradual trailing edges are used. These pulses also melt the filament but allow the material to solidify in a crystalline form.

Manhart has also proposed a structural model to explain memory switching in silicon monoxide sandwich structures with silver electrodes [21]. In these structures, a forming pulse was again used to reduce the device resistance from its initial value of \(\sim 10^9 \Omega\) to \(\sim 300 \Omega\). According to Manhart's model, the forming process gives rise to a temperature increase sufficient to induce silver from the electrodes

---

Figure 14.6 Schematic diagram of dynamic I-V characteristics illustrating several memory states (A, B, C, D) and the threshold voltage \(V_T\). From: [2].
to migrate into the silicon monoxide film to produce a metallic filament. This explains the subsequent linear $I-V$ characteristics and positive temperature coefficient in the ON state. Two different mechanisms were suggested for the switching. The first assumes that the erase pulse (10V, trailing edge 1000 ns) melts the metal filament only at the narrowest point. The metal disperses into the silicon monoxide, breaking the link. To write in again and hence reconstruct the filament, a write pulse (6-8V, trailing edge >0.1 ms) encourages more metal from the electrodes to enter, that is, a long pulse is able to heat the boundaries sufficiently to cause new diffusion from the electrodes. The second model also suggests that the erase pulse melts the metal filament. After the removal of the pulse, rapid cooling leads to the formation of a high-resistance amorphous mixture of metal and silicon monoxide phases. To write in again, this material is melted again but then cooled more gradually. The difference in the freezing points of the metal and the silicon monoxide cause them to separate, thus restoring the filament.

Memory switching in thermally grown microcrystalline NiO has been reported by Gibbons and Beadle [7]. They too suggest that a metallic filament is formed in their structures, not by the diffusion of metal from the electrodes, but by the collection of nickel atoms at a structural defect in the oxide. This occurs during the forming stage of the device where a filament of material becomes sufficiently hot for this stoichiometric change to take place. The device can have an ON-state resistance of approximately 100$\Omega$ and an OFF state resistance of 25 M$\Omega$. In the ON state a nickel filament is formed, which connects the ohmic contacts. Once the device is switched ON, the device can be very rapidly switched OFF by essentially burning out a small section of the conducting filament. Once the filament is

Figure 14.7 Memory-switching process in metal-chalcogenide-metal devices.
formed, only the section removed or reoxidized during the OFF state need be replaced. The main body of the filament, which remains more or less intact during the OFF-state transient, provides a large nickel reserve for this replacement. This will obviously limit the lifetime of the device.

14.1.5 Electrical Switching in Amorphous-Silicon Structures

The earliest reports of switching in amorphous silicon (a-Si) were published in 1970 [13, 46] contemporaneously with some of the early literature on switching in chalcogenide glasses. More detailed experiments on the same structures were reported later [9, 47]. These investigations studied vacuum-evaporated films of a-Si in the range 0.3–2.0 \( \mu \text{m} \) thick, fitted with titanium electrodes. Similar observations were made on evaporated films of germanium, boron, and boron plus carbon. As threshold devices, these a-Si structures had threshold voltages \( V_\text{th} \) of 5–10 V, OFF-state resistances in the range of 1–30 k\( \Omega \), and an ON resistance of about 100\( \Omega \). In common with the chalcogenide glasses, there was a delay time \( t \), before switching of 20–50 \( \mu \text{s} \) or more at room temperature, and the actual switching time was at least several microseconds. Feldman and Charles [9, 47] did not, however, report any initial forming process, unlike the situation in chalcogenide glass switches (Section 14.1.1). There was also some tentative indication of memory switching but this apparently was not substantiated. They interpreted their results in terms of a simple and qualitative electrothermal model involving the formation of a conducting filament.

The work of Feldman and his colleagues [13, 46], which originated in the early 1970s, seems to be the only investigation of switching in a-Si until the rather later studies of Dey and Fong [48, 49]. These authors reported results very similar to those of Feldman. They studied thin films of a-Si in the range 0.3–1.5 \( \mu \text{m} \) thick, deposited by electron-beam heating in a vacuum evaporator. Titanium contacts were again used, either in the form of evaporated films or as probes. Dey and Fong reported only threshold switching, with \( I-V \) characteristics similar to those in Figure 14.3; they did not mention any evidence for memory behavior. In contrast to Feldman, however, Dey and Fong did observe forming effects, that is, the initial threshold voltage was relatively large but decreased to a more or less constant value after a number of switching cycles. In Dey and Fong's devices the threshold voltage varied systematically from about 6 V for the thinner films (\( \sim 0.3 \mu \text{m} \)) to about 9 V for the thicker films (\( \sim 1.2 \mu \text{m} \)). The delay time before switching was in the range 2–60 \( \mu \text{s} \), varying in a systematic way with pulse height, pulse duration, and repetition rate, again in a manner very similar to threshold switching in chalcogenide glasses. Dey and Fong also interpreted their results in terms of a simple one-dimensional electrothermal model, although developed a little more quantitatively than that by Feldman and Charles [9, 47]. It should be noted that both teams (Feldman and Charles, and Dey and Fong) used a-Si films deposited by vacuum evaporation, which probably accounts for the relatively low OFF-state
resistances (~100 kΩ) they both found. It is now well established that vacuum-evaporated a-Si is a very different material from the hydrogenated form of a-Si obtained, for example, by glow-discharge deposition of silane [50].

Three papers, concerned specifically with switching in amorphous silicon, by Gabriel and Adler [51], den Boer [52], and Owen [53] appeared almost concurrently early in 1982, each reporting very different effects observed in different amorphous silicon structures. Our own work [54–58], including recent results, will be described in detail in the following sections.

Den Boer studied n+–i–n− structures of a-Si:H prepared by the glow discharge decomposition of SiH4 (i stands for near-intrinsic, or undoped, material). The n+ layers were 0.05 μm thick and were prepared by adding 1% PH3 to the SiH4 gas flow; the i layer in different devices ranged in thickness from 2.5 to 5 μm. Den Boer found that the n+–i–n− devices functioned as threshold switching devices with nonpolar characteristics similar to those in Figure 14.2(a). For the first switching cycle, the threshold voltage was in the range 40–100 V, but for all subsequent operations it was only 10–35 V, depending on the i layer thickness. (As the i layer thickness increased, the threshold voltage also increased.) The OFF-state resistance of the n+–i–n− devices was about 1 MΩ, and the ON-state resistance was about 1 kΩ. There was an observable delay time before switching, ranging from a few microseconds when the applied voltage was about 8 V greater than Vth to about a millisecond for voltages within 1 V of Vth. The n+–i–n− switches could be cycled through at least 109 stable switching operations. Den Boer also compared structures with chromium or a combination of chromium and n+ contacts (i.e., Cr–n+–i–Cr and Cr–i–Cr). The Cr–n+–i–Cr contacts had rectifying characteristics, while the Cr–i–Cr contacts switched but were very unstable.

Gabriel and Adler [51] prepared their films by sputtering from a polycrystalline silicon target in an argon-hydrogen plasma. In all cases their devices were notionally homogeneous thin films of intrinsic a-Si:H with molybdenum contacts. The samples were fabricated under a wide range of deposition conditions in two sputtering systems, and although results from some of the devices were rendered rather doubtful because of contamination problems, in no case did Gabriel and Adler observe any evidence of reversible switching. They concluded that, in contrast to the chalcogenide glasses, amorphous silicon does not have the electronic and structural properties required for reversible switching.

14.2 RECENT DEVELOPMENTS ON AMORPHOUS-SILICON SWITCHES

14.2.1 Digital Switching in a-Si:H p+-n–i Devices

We now turn to the work on electrical switching carried out in the authors' laboratories. Although a number of different a-Si:H multilayer structures have been investigated, all the earlier results discussed in the following refer to metal-
In the early stages of the work, the a-Si:H layers were deposited onto a stainless-steel substrate. After completion of the a-Si:H deposition, a series of gold (Au), aluminum (Al), or nichrome (NiCr) dots, up to approximately 1 mm in diameter, was evaporated onto the surface of the samples. The top contact was completed by a probe or by a thin wire attached to the metal dots with conducting silver paste.

In order to facilitate the investigation of device properties on a more reproducible basis, a "pore" device structure was designed, with device areas down to a few µm in diameter. A set of photomasks was produced that allowed modern photolithographic technique to be combined with the normal a-Si:H deposition process. Figure 14.8(a) shows a cross section of one of the pore structures produced by photolithographic techniques, and Figure 14.8(b) represents a plan view of the same structure. The a-Si:H films were deposited on Corning glass substrates previously patterned with chromium bottom contacts. The a-Si:H was prepared in a layer sequence of \( p^+-n-i \) by the glow-discharge decomposition of silane using gas-phase doping. The \( p^+-n-i \) films were then patterned, and an insulating layer was used to define an active device area of \( 10^{-6} \) cm\(^2\). The metal used for the top contact was normally chromium but a number of other metals were also used, and their influence on the memory operation was thoroughly investigated.

\begin{figure}[h]
  \centering
  \includegraphics[width=\textwidth]{image}
  \caption{Figure 14.8 Structure of metal-amorphous silicon-metal switch device.}
\end{figure}
14.2.1.1 Static Current-Voltage Characteristics

Typical $I-V$ characteristics for a freshly prepared (unswitched) device are illustrated in Figure 14.9(a) in both the forward and reverse directions; the forward direction is defined so that the substrate (and hence the $p^+$ region) is positively biased. It must be emphasized here that these measurements were taken "by hand," point by point, in a manner that required a few seconds for each measurement. (The significance of this remark will become apparent in the next sections.) In the forward direction there is a region of ohmic behavior over a limited voltage range followed by an abrupt change to a markedly nonohmic region until, at the point indicated by the arrow, the device is unstable and it becomes impossible to continue with point-by-point measurement. The change from ohmic to nonohmic behavior is more

Figure 14.9 (a) $I-V$ characteristics of unformed (unswitched) amorphous silicon device; (b) conductivity versus voltage plot.
clearly apparent in the effective-conductivity-versus-applied-voltage plot of Figure 14.9(b). As the temperature increases, the onset of nonohmic behavior moves to lower voltages. In the reverse direction, corresponding to a negative potential applied to the $p^-$ side, there is an initial ohmic region that is symmetrical for positive and negative voltages. However, the change to nonohmic behavior is much more gradual in the reverse direction, leading to eventual breakdown of the device.

As noted above, during point-by-point measurements under forward bias, the a-Si:H $p^--n-i$ device tends to become unstable when the applied bias is about 24V at room temperature. At higher temperatures, the instability, indicated by the arrows in Figure 14.9(a), occurs at lower voltages. When we attempt to increase the voltage still further, the device switches into a low-resistance ON state. Typical $I-V$ characteristics for both polarities in the ON state are shown in Figure 14.10. The $I-V$ curve is ohmic; it extrapolates through the origin (i.e., the ON state is a permanent memory state); and it is slightly asymmetrical. Note that the current is now measured in milliamperes and the voltages across the device are small. On increasing the voltage in the forward direction, we find that the ON-state current continues to increase apparently indefinitely, subject only to any current-limiting resistor, and the device is eventually destroyed, presumably by Joule heating. In the reverse direction, however, another instability is observed, and at about $-1V$ (typically) the device switches back into a high-resistance OFF state, as shown in Figure 14.11. The OFF-to-ON transition may now be repeated by biasing in the

![Figure 14.10 I-V characteristics of memory ON states.](image-url)
forward direction, but on the second and all subsequent switching operations the forward threshold voltage $V_{ThF}$ occurs at a much lower voltage than the first operation (e.g., at $\sim 5V$ compared with the $25V$ observed under the conditions of the measurements shown in Figure 14.9). The first OFF-to-ON transition, occurring at a relatively high voltage, seems unique and by analogy with the terminology in chalcogenide glasses, it is referred to as "forming."

The formed $p^+\cdot n\cdot i$ device may be cycled through ON and OFF states by a biasing in forward and reverse directions with critical points at $V_{ThF}$ and, in the reverse direction, $V_{ThR}$, as illustrated in Figure 14.11. It is important to emphasize that these devices exhibit not only nonvolatile but polarity-dependent memory switching. On occasions the device appears to go through a number of intermediate states during the OFF to ON transition, and this is indicated in the figure. In addition, there is often an observable and appreciable region of negative resistance in the reverse-biased OFF-state characteristics of the formed device.

A number of other dc experiments have also been carried out on formed devices in an attempt to provide additional information, primarily about the nature of the ON state. Figure 14.12(a) compares the area dependence of the OFF- and ON-state resistances for samples of different area (i.e., different pore sizes with diameters from 5 to 300 $\mu$m) from a single $p^+\cdot n\cdot i$ deposition run [54]. These results were obtained on specimens that had previously been switched many times. Within
the experimental scatter, $R_{\text{OFF}}$ clearly scales with the reciprocal of the area $A$, demonstrating that in the OFF state the current flows throughout the whole area of the specimen. In complete contrast, the values of the ON-state resistances $R_{\text{ON}}$ for the same specimens show no area dependence at all. This result can be understood only if the ON state has its origin in a highly conducting filament, less or equal to a few $\mu$m in diameter, that extends through at least part of the specimen thickness.

The temperature dependence of both the ON- and OFF-state conductance has also been measured over the temperature range from about 230K to 400K. Results for a typical $p^+ - n - i$ specimen are shown in Figure 14.12(b). The OFF-state conductance $G_{\text{OFF}}$ varied slowly with temperature, increasing by less than a factor of three between 230K and 360K. The ON-state conductance was even less temperature-dependent, increasing by only 10% over this temperature range. The insensitivity of these device parameters to temperature is also observed in other properties. For example, the magnitudes of the voltages required to switch the device ON and OFF, measured under pulsed conditions, both increase by only a factor of 2.5 as the temperature is decreased from 400K to 200K. Clearly, the general temperature insensitivity of the switching must also be a feature of any theoretical model.
Transverse magnetoresistance $\Delta \rho/\rho$ measurements of the ON state of a number of specimens have also been carried out at room temperature up to magnetic fields of $B = 0.5 \, T$ [55]. Within the experimental error, $\Delta \rho/\rho$ was proportional to the square of the magnetic field strength $B$ and was found to be positive. The values of $\Delta \rho/\rho(B^2)$ ranged from $0.5$ to $2.0 \times 10^{-4} \, T^{-2}$. Experiments on phosphorus-doped a-Si:H after thermal crystallization also gave a positive magnetoresistance with the same $\Delta \rho/\rho(B^2)$ dependence. Unfortunately, no results have been reported on homogeneous undoped glow-discharge a-Si films; it is therefore difficult to draw any definite conclusions about the amorphous or crystalline nature of the filament from the present results. However, it is probably correct to associate the $B^2$ dependence of the memory ON state with a longer mean free path than is normal in amorphous solids. By conventional crystalline theory [58], the magnitude of the measured $\Delta \rho/\rho(B^2)$ would lead to an effective mobility $\sim 100 \, \text{cm}^2/\text{V} \cdot \text{s}$ supporting of this suggestion. The importance of this measurement will be further discussed in Section 14.2.2.1 in connection with the very recent observation of electron quantization phenomena observed at low temperatures.

Direct evidence for the existence of a filamentary ON state was obtained from thermal imaging [59]. The a-Si:H $p-n-i$ device was covered with a thin layer of thermochromic liquid crystal. By passing current through the device in the ON state, it is possible to observe the current path from the changes produced by Joule heating in the reflected color of the liquid crystal. The resulting features for 20-μm pore diameter structures were viewed through a high-powered optical microscope. The change in the liquid-crystal appearance produced by the filament in the ON state could be seen clearly as a small circular “hot spot” approximately in the center of the pore. The visual observation of the current filament has also enabled us to establish that switching a device OFF and ON again produces the current filament in the same place and this implies that the switching processes are not destructive. In these experiments the specimens were covered by a thin layer of a liquid crystal, 4-cyano-4-alkylbiphenyl, which undergoes a nematic-liquid phase transition at 35.3°C. The phase boundary may be observed quite easily in cross-polarized light, and the transition was found to be fast and without hysteresis. If the sample temperature is fixed using a thermostatically controlled stage, the difference in temperature between a region of local heating (at temperature $T$) and the surrounding film (at $T_0$) may be determined. The results indicate that no observable temperature rise occurred as a result of applying electrical power to the pore in either the unformed state just prior to forming or in the formed OFF state just prior to switching. However, as described above, in the ON state, local heating (which results from applying continuous power) could be clearly seen. The effect of changing the rms power applied to a 50-μm-diameter pore was studied using a continuous train of 300-ns pulses. The stage was maintained at 21°C; thus the phase boundary represented the locus of points $(35.3° - 21° = 14.3°C)$ above the film.
temperature. These isotherms were circular and in the particular case studied were symmetric about the center of the pore.

A plot of the phase boundary diameter \(d_1\), versus rms power \(P_{RMS}\) is shown in Figure 14.13(a). Although there is considerable scatter, it can be seen that the relationship between \(d_1\) and \(P_{RMS}\) is substantially linear for \(d_1 > 2 \mu m\). Below this the accuracy of the measurements is limited by the resolution of the microscope used; the onset of observable effects occurs at \(P_{RMS} = 2\) mW. A linear dependence of \(d_1\) on \(P_{RMS}\) is obtained as a solution of the heat conductivity equation for an idealized system of this kind, in which the source of local heating is assumed to be much smaller than \(d_1\). Thus these data indicate that the diameter of the ON-state filament \(d_1\) must be less than about 2 \(\mu m\).

![Graph](image-url)

**Figure 14.13** (a) Liquid crystal phase boundary diameter versus RMS power; (b) phase boundary diameter versus RMS power at various values of excess temperature \(\Delta T\).
The steady-state thermal properties of the 20-μm pore structures with different top electrodes (Al, Cr, Ti) have also been studied, and the results are represented in Figure 14.13(b), where rms power is plotted versus the phase boundary diameter for different local temperature increases. It is found that the gradients $G_A$ and intercepts $P_0$ scale linearly with the local temperature rise $\Delta T$, with the constants of proportionality $M_G$ and $M_{P_0}$ depending on the top electrode metal (and possibly its thickness, although this has yet to be established). Therefore, the following relationships can be used to characterize the steady state thermal properties:

\[
G = M_G \Delta T \tag{14.7}
\]

\[
P_0 = M_{P_0} \Delta T \tag{14.8}
\]

and therefore

\[
P = M_{P_0} \Delta T + M_G \Delta T d \tag{14.9}
\]

The values of $M_G$ and $M_{P_0}$ are summarized in Table 14.1. In the case of Al and Cr top electrodes, several different samples have been investigated, with the values obtained being similar to those in the table.

<table>
<thead>
<tr>
<th>Top electrode</th>
<th>$M_G$(Wm$^{-1}$K$^{-1}$)</th>
<th>$M_{P_0}$(mWK$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al</td>
<td>3</td>
<td>0.2</td>
</tr>
<tr>
<td>Cr</td>
<td>6</td>
<td>0.02</td>
</tr>
<tr>
<td>Ti</td>
<td>7</td>
<td>0.007</td>
</tr>
</tbody>
</table>

The physical significance of $M_G$ is that it represents the surface thermal conductivity of the pore, and thus it is not surprising that the values obtained experimentally lie between the bulk values for glass (1 Wm$^{-1}$K$^{-1}$) and metal (100 Wm$^{-1}$K$^{-1}$). The significance of $M_{P_0}$ is less clear. In the ideal case of a point heat source located on one thermally insulated surface of an infinite plate, the other surface being isothermal, $M_{P_0}$ should be zero. Thus it may be that the nonzero values in some way reflect the extent of the heat source. If this is the case, it would appear that the filaments in pores with Cr and Ti top electrodes are smaller than in those with Al electrodes.
14.2.1.2 Current Instabilities in Unformed State

Unusual current instability phenomena have been observed in a-Si:H $p^*-n-i$ structures with Al top electrodes in which the $i$ layer is either thin (100Å) or lightly $n$-doped (1–5 vppm) [56]. The principal features associated with the current instability are shown in Figure 14.14. A voltage pulse $V$, greater than a minimum threshold $V_c$ (specified later) is applied so the $p^*$-layer is positively biased. A displacement current occurs, which decays to a steady current $I_s$. Some time $t_d$ after the voltage pulse is first applied, the current rises abruptly. After reaching a maximum value $I_{max}$, the current decays. A second pulse may be applied at time $t_r$ after the first pulse, which may or may not give rise to a second instability, depending on the values of $V$ and $t_r$. The dependence of the onset time $t_o$ on pulse height $V$

![Figure 14.14 Current instability in metal-amorphous silicon structure.](image)

for three different temperatures is shown in Figure 14.15. The experimental data are summarized by the following empirical relation:

$$t_o = t_{o0} + t_o \exp \left[ \frac{-(V - V_o)}{V_0} \right]$$

(14.10)

where $V_o \sim 0.6-0.8V$, $V_c \sim 10-15V$, $t_o \sim 10-100 \mu s$, and $t_{o0} \sim 100$ ns. We note that $V_o$ is temperature independent over the range studied, but some or all of the other parameters clearly show a temperature dependence.
Figure 14.15 Onset time versus pulse height (current instability) at different temperatures.

The above relation applies only to the region where $V$ is somewhat larger than $V_i$. The detailed relationship between $t_d$ and $V$ at $V = V_i$ is difficult to establish; there is a tendency towards erratic behavior, but no current instabilities have yet been recorded with $t_d > 100 \mu s$. It appears that $t_d$ either increases almost discontinuously at $V = V_i$ or the effect simply does not occur under these conditions. As $V$ is increased, several volts above $V_i$, $t_d \rightarrow t_m$, but further increase in $V$ causes an irreversible change in the device characteristics, culminating in a permanent low-resistance ($10^3 - 10^6 \Omega$) state when the voltage pulse is removed.

The observed minimum delay time $t_m$ may correspond to the transit time of an injected pulse of holes through the n-region. Taking typical values of $t_m$, n-layer thickness, and maximum applied voltage ($0.1 \mu s$, 0.1–0.3 $\mu m$, and 10–20V, respectively) and assuming that a large proportion of the applied bias under these conditions appears across the n-layer, the hole-mobility values can be estimated to lie in the range $10^{-4}$ to $10^{-3}$ cm$^2$V$^{-1}$s$^{-1}$, which is not unreasonable. Measurements of the dependence of $I_{max}$ at constant $t_m$ on device areas over the range $10^{-4}$–$10^{-7}$ cm$^2$ have failed to reveal a direct proportionality; $I_{max}$ ranges over an order of magnitude in the various samples, but this seems to be primarily a random variation. As the smallest device examined in this study was about 10 $\mu m$ in diameter, this suggests that the current during the instability is transported through a region of these dimensions or less. This is an important observation, as it has been established
(Figure 14.13(a)) that conduction in the memory-ON state is localized within a filament of 1-μm diameter or less. Thus it seems that the localized conduction occurring during the current instability may signify the formation of an incipient filament and is therefore a precursor to the memory-forming process. Additional indirect evidence supporting this view is that the forming delay time versus applied voltage relationship can be described by an expression similar to that given in (14.10).

Certain features associated with the current instability are found to occur in the analogous crystalline-silicon metal/n-p+/metal (MISS) devices [10]. In particular, the relationship between \( t_d \) and \( V \) is similar. This suggests that the principles on which the theory of operation of the crystalline device is based may be applicable here. Although such theories differ somewhat in detail, the central concept is that sufficient holes, injected from the forward-biased p-n junction, accumulate to form an inversion region at the n-i interface. Buoxo [59] demonstrated that the theoretical expression for the \( t_d \)-versus-\( V \) relationship, based on the establishment of an inversion layer, is in good agreement with experiment. However, the behavior of the amorphous device differs in two important aspects. First, the current decays after reaching a maximum value, even when the voltage is maintained. The crystalline MISS device remains in its high-conductivity state provided a "holding" voltage is present. Second, there exists a voltage-dependent recovery time \( t_r \), as described earlier and shown in Figure 14.14. This effect has no analog in the crystalline MISS device. The current decay may occur as a result of electron-hole recombination close to the n-i interface, which will increase the potential barrier to electrons tunneling from the metal into the conduction band in the n-layer. This negative feedback will decrease the tunneling contribution to the total current; however, once the initial conditions are reestablished, the current should again rise, and this is not observed experimentally on a time scale of μs. The recovery effect may be the consequence of some longer-term structural change caused by recombination, or by Joule heating, as the local power density will be quite high.

14.2.1.3 Forming

The forming process does not occur instantaneously when a voltage step or pulse is applied to the device. Initially there is a delay time \( t_0 \) during which the device current remains essentially constant at the OFF-state value appropriate to the voltage across the device. Only after this delay does the current begin to increase, and it then rises almost instantaneously to its ON-state value. The forming delay time is an extremely sensitive function of the applied forming voltage \( V_F \), and typical data, obtained at three temperatures, are given in Figure 14.16. The forming delay time varies over nearly 10 orders of magnitude, from a few hundred seconds at low forming voltages to about 10 ns at high \( V_F \). In particular, at a temperature-dependent critical forming voltage \( V_F \), there occurs a virtually discontinuous change
in \( \tau_0 \). The critical voltages \( V_c \) indicated in Figure 14.16 are approximately the same as the voltages at the points of instability marked by the arrows in Figure 14.9(a); \( V_c \) also corresponds to the forming voltage obtained in experiment with a curve tracer operated in ac mode at a frequency of 100 Hz. It can also be seen in Figure 14.16 that above and below \( V_c \) the delay time tends to a value that seems to be approximately independent of both voltage and temperature; for the particular results illustrated, \( \tau_0 \) is in the range \( 10^2-10^3 \) s for \( V < V_c \) and lies between 10 and 100 ns for \( V > V_c \).

The results plotted in Figure 14.16 for \( V < V_c \) correspond, of course, to voltages less than the point of instability indicated in Figure 14.9(a). There does appear to be a lower limit to the forming voltage, however, and present results indicate that the limiting voltage coincides with the bias at which the \( I-V \) characteristics change from their ohmic to nonohmic behavior (Figure 14.9(b)). Several experiments have shown that virgin devices fail to switch (form), even if held for many hours at a forward bias only slightly below the nonohmic region. In other words, forming occurs at any forward bias within the nonohmic region of the \( I-V \)

---

**Figure 14.16** Forming delay time versus voltage height.
characteristics, but at voltages below the point of instability, \( \tau_d \) is comparatively long. It must also be reemphasized that the device current remains constant at its preformed magnitude during the delay time, even when \( \tau_d \) is 100 s or more.

Experiments have also been carried out to determine the effects of \( p^+\text{-}n\text{-}i \) device geometry on the forming voltage \( V_f \). It was found that \( V_f \) increases linearly with the thickness \( d_n \) of the \( n \)-layer [54]. The charge \( Q = \int I \, dt \), which flows through or into the device during the forming delay time, has also been determined for \( V_f > V_c \). In this range the ratio \( (Q/d_n) \) is approximately independent of \( V_f \) and \( d_n \) for \( n \)-layer thicknesses between 0.2 and 0.8 \( \mu \text{m} \). This could mean that forming occurs when a critical volume charge has accumulated in the \( n \)-region.

The current \( I-V \) characteristics in the high-field preswitching region are plotted as \( \log \sigma \) versus \( E^{1/2} \) in Figure 14.17, showing that the logarithm of conductivity prior to forming is proportional to the square root of the applied electrical field. There are at least two conduction mechanisms that can give rise to nonohmic behavior of this kind:

1. The Poole-Frenkel effect, in which the potential barrier for the thermal excitation of trapped electrons into the conduction band is lowered by the applied external field;
2. The Richardson-Schottky effect, which is associated with the lowering of the effective work function (or barrier) for charge-carrier emission from a metal electrode when an electric field is applied. The lowering is due to combined effects of the field and the image force.

Assuming that the conductivity is bulk limited, the high frequency dielectric constant obtained by fitting the experimental data to the Poole-Frenkel model is 48.7, whereas a value of 12.7 is obtained using the Richardson-Schottky model. The dielectric constant for crystalline silicon is about 12, and this value does not change very much with frequency. The dielectric constant for amorphous silicon is probably similar. It is more likely therefore that the current flow in the high-field preswitching region (just prior to forming) is determined by the Richardson-Schottky effect at the metal/\( i \)-layer contact. This conclusion is further supported by the observation that the forming voltages are different in devices with different metal electrodes but otherwise identical thicknesses, indicating the importance of the Schottky barrier in the structure. However, it should be kept in mind that the values of \( E \) in the above analysis presuppose a uniform field whose magnitude is proportional to applied bias. It was shown in Section 14.1.1 that, under certain circumstances, conduction in \( p^+\text{-}n\text{-}i \) samples in the high-bias regime is spatially nonuniform, which suggests that the assumptions made concerning the field within the conducting region may become less valid as the bias is increased. The rapid rise in conductivity under these conditions, as shown in Figure 14.17, may correspond to the onset of electron tunneling from the metal through the \( i \)-layer under a very high field.
14.2.1.4 Dynamic Switching Behavior

The principal features of the pulsed operation of formed a-Si:H $p^+-n-i$ devices have been described by Owen [53] and LeComber [54]. A representative diagram from an oscilloscope trace of the OFF $\rightarrow$ ON (WRITE) and the ON $\rightarrow$ OFF (ERASE) transitions on applying a voltage pulse is shown in Figure 14.18. The oscillations on these traces are caused by ringing effects in the rather poorly matched electrical setup. The main points to note are:

1. When biased with a pulse in the forward direction, the device switches from OFF to ON (WRITE), provided the pulse height exceeds the static threshold voltage $V_{th}$, as defined in Figure 14.11.

2. There is a delay time in the WRITE operation (Figure 14.18(a)) which is a strong function of the WRITE pulse magnitude as shown in Figure 14.19. These delay times are significantly faster than the delay times reported for chalcogenide switches (see for example [12, 14]). If the results in Figure 14.19 are expressed in the form $t = t_0 \exp(-V/V_0)$ then $t_0 = 335$ ns and $V_0 = 4.5$ V. Similar results have been obtained for all the specimens investigated although the $V_0$ values ranged from about 0.5 to 13 V.

![Log conductivity versus square root of electric field.](image)
Figure 14.18 (a) WRITE and (b) ERASE switching transient waveforms.

Figure 14.19 WRITE delay time versus voltage height.
3. Provided the pulse is long enough, the ON state is permanent and the pulse duration required for switching to a memory state increases as the pulse height decreases toward \( V_{th} \). In typical cases a permanent ON state is obtained with pulse durations of a few tens of nanoseconds and magnitude \( \sim 5 \text{V} \) in excess of \( V_{th} \).

4. Similarly, on biasing in the reverse direction with a pulse of height \( > V_{th} \), the device switches from ON to OFF (ERASE) and again there is a delay time of the order of nanoseconds in the response (Figure 14.18(b)).

5. Both the OFF and ON states appear to be truly permanent. No detectable changes have been observed in devices stored at room temperature, either in the OFF or ON state, for a number of years.

14.2.2 Analog Memory Effects in a-Si:H Metal-p'-Metal Structures

More recent experimental results [57, 60] have demonstrated a new metal-p'-metal amorphous silicon device which, rather than exhibiting a two-state digital operation, has a continuum of stable states that are nonvolatile and fully programmable by single 10-ns voltage pulses. It has also been suggested that the new analog memory devices can be used as nonvolatile and reprogrammable memory elements in analog neural networks [61]. In this section we present a summary of the new results obtained on nonvolatile analog switching effects in amorphous silicon metal-p'-metal devices and discuss the possible physical mechanisms responsible for the phenomena. The samples used for this work were amorphous silicon Cr-p'-V thin-film structures. The p' layer was prepared by RF glow-discharge decomposition of SiH\(_4\) containing 10 ppm of B\(_2\)H\(_6\). Films of 1000Å thickness were deposited onto Corning glass substrates previously patterned with chromium bottom contacts. The p' amorphous silicon was then patterned, and an insulating layer was used to define an active device area of 10\(^{-6}\)cm\(^2\). The metal used for the top contact was normally vanadium. However, a number of different metals were also used, and their influence on the memory operation will also be described.

In accordance with our previous results [53], memory devices prepared in this way require an initial forming process. This means that the resistance of the as-deposited (unformed) device has to be lowered from \( R \sim 10^{10}\Omega \) to \( R \sim 10^3 \sim 10^4\Omega \) (i.e., the typical value of an ON state). The forming can be achieved by biasing the sample with a single voltage pulse (duration 300 ns, magnitude \(-12\text{V}\)) with positive polarity applied to the top V contact. The metal-p'-metal memory structures exhibit a forming step that is different from the previously investigated metal-p'-n-i-metal structures [55]. The differences are demonstrated in Figure 14.20. In the case of metal-p'-n-i-metal structures, the resistance suddenly drops from \( \sim 10^{10}\Omega \) (virgin state) to \( \sim 10^2\Omega \) after the critical voltage (forming voltage \( V_f \)) has been applied. No change in the virgin resistance occurs when the sample is biased
Figure 14.30 (a) Waveforms of a single forming pulse applied to amorphous silicon Cr-p'-n-i Al structure showing hard forming; (b) waveforms of single pulses of increasing magnitude applied to amorphous silicon Cr-p'-V structure showing soft forming (V and I represent voltage across device and corresponding device current, respectively).

With voltages less than $V_r$, this type of forming is termed hard forming. In contrast to this, the resistance of the unformed metal-p'-metal structures can be lowered gradually by applying voltage levels with progressively increasing magnitudes. In this case, no sudden change of the current or voltage signal can be detected when the sample is biased with a voltage pulse, as seen in Figure 14.20(b). This process is called soft forming. Figure 14.21 shows the device resistance as a function of the soft-forming voltage (pulse duration = 300 ns). On reaching a critical voltage (~14 V in Figure 14.21) the device resistance suddenly drops from $\sim 10^4 \Omega$ to $\sim 10^2 \Omega$. This is the memory-ON state of the device. Once the memory device had reached its first (nonvolatile) ON state, all subsequent switching operations were performed with 10-100-ns pulses of 1-5 V.

An example of the analog switching effect is shown in Figure 14.22(a), where the sample resistance is plotted as a function of applied alternating WRITE and ERASE pulses (100 ns pulse duration each). It is important to emphasize the polarity dependence of the analog memory behavior. In the case of the WRITE pulses, positive polarity is applied to the Cr track (bottom contact), whereas ERASE pulses have opposite polarity. The sample was first switched to an ON state ($R_{ON} = 2 \times 10^2 \Omega$) and then a series of alternating WRITE and ERASE pulses were applied. The WRITE pulses were kept at a constant magnitude of 3.4 V but the ERASE pulses were incremented by 0.05 V steps from 1.2 to 3.4 V after each WRITE pulse. It can be seen from Figure 14.22(a) that the sample resistance
changes in an analog manner as the magnitude of the ERASE pulse increases, that is, the difference between \( R_{ON} \) and \( R_{OFF} \) is a function of the magnitude of the ERASE pulses. A voltage range of \( \Delta V \) (ERASE) = 1.6V resulted in a change in resistance from \( R \sim 2 \times 10^5 \Omega \) to \( R \sim 6 \times 10^4 \Omega \). The shaded area in Figure 14.22(a) indicates the reproducibility of the analogue memory switching, i.e., the scattering in the resistance during repeated experiments (data from 100 cycles are shown).

Figure 14.22(b) shows another case where the ERASE pulses were maintained at a constant value of \( V = 3.4 \text{V} \), but the WRITE pulses were incremented from 1.2V to 3.4V in 0.05V steps. The value of the OFF-state resistance remained constant at \( ~6 \times 10^4 \Omega \) that is, it changed back to this constant level from every ON state, while the ON-state resistance decreased through a continuum of intermediate states over a similar \( \Delta V \) to the ERASE operation. The shaded area in Figure 14.22(b) again represents the reproducibility of the analog switching for 100 complete cycles.

It is emphasized that the device will switch between any two resistance states within the range from about 1 k\( \Omega \) to 1 M\( \Omega \) by selecting the correct polarity and magnitude of the WRITE and ERASE pulses. For all devices with a vanadium top contact, the values of \( \Delta V \) range from 1.5 to 2.0V for both the WRITE and ERASE operations.
We have repeated the above experiments on metal-p*-metal devices with Cr as top metal and observed similar polarity dependent changes in the memory state resistance. However, in the WRITE and ERASE experiments, intermediate states were found to exist over only a narrow $\Delta V$ of about 0.2V, as shown in Figure 14.22(c). Therefore, these devices are considered as digital devices. It is important to emphasize that both the analog (Figures 14.22(a) and (b)) and digital (Figure 14.22(c)) memory switching effects are nonvolatile. Devices set to ON or OFF states have been monitored for more than two years without any significant change.

Figure 14.22 (a) Memory resistance as a function of ERASE voltage in a Cr-p*-V structure; (b) memory resistance as a function of WRITE voltage in a Cr-p*-V structure; (c) memory resistance as a function of WRITE and ERASE voltages in a Cr-p*-Cr structure.
their resistance. Also, operation at temperatures up to 160°C shows little change in the threshold voltages or device stability. However, it is also found that devices with certain top metal contacts, such as Mo and Pd, show a volatile memory-switching effect. This is illustrated in Figure 14.23, where the current signal through the device is continuously monitored at a low voltage level (0.5V), which is below the voltage level of the programming pulses. The current signal decays rapidly after the end of the programming pulse, indicating that the memory state is volatile. The role of the top metal contact, therefore, has been thoroughly investigated by fabricating devices with a range of different top metals but with otherwise identical physical parameters (i.e., ~1000Å thickness of p'-layer, Cr bottom electrode). With the analog switching voltage window $\Delta V$ as a guide, its value is found significantly dependent on the top metallization contact. This is illustrated in Table 14.2.

![Figure 14.23 Volatile memory effect in a Cr-p'-Mo structure.](image)

It can also be seen that the definition of analog ($\Delta V \approx 1$V) or digital ($\leq 0.5$V) memory switching is somewhat arbitrary, because there is no sharp boundary between the two types of operation, that is, they are almost certainly associated with the same underlying physical phenomena. However, in the cases of Mo and Pd top contacts, a new type of volatile switching effect is observed and in the cases
of Ti, Au, and Cu, no reproducible switching effects can be observed. These results suggest that the top metal contact plays a crucial role in determining the type of memory-switching phenomena observed in these devices. Here we will concentrate on the p⁺ memory devices with vanadium top contact, because these show typical nonvolatile analog memory switching.

The I-V characteristics of analog memory-resistance states were investigated systematically, both at room temperature and lower temperatures, and the following results were obtained. It was found that all the room temperature current-voltage characteristics show a linear-plus-power-law behavior in the various analog memory resistance states as shown in Figure 14.24. The I-V curves can be described by a simple nonlinear relationship:

$$I = C_1V + C_2V^n$$

(14.11)

where $C_1$ and $C_2$ are constants and the exponent $n$ increases with the low bias (linear region) resistance according to the relationship $n = A + B \log R$. The observed power-law behavior could indicate the possibility of space-charge limited conduction at higher bias, although investigation of the thickness dependence has shown this to be unlikely. In accordance with the pulsed analog memory switching results, a continuous transition of states can be found between the ON ($\sim 10^9\Omega$) and OFF ($\sim 10^4\Omega$) states. The terms ON and OFF seem to be somewhat arbitrary, therefore, and are used in this chapter only for practical reasons. If the curves in Figure 14.24 are extrapolated above 1V, they meet in the region of 3 to 4V (i.e., at typical programming levels—see Figures 14.22(a)-14.22(b)). It was also found that the exponent $n$ does not depend on the thickness or the diameter of the active device but it is primarily determined by the low bias (or linear region) resistance of the analog memory state. This indicates that the nature of the electrical conduction is quite similar in all memory states. However, repeatedly switching the device into the same resistance state need not always result in an identical value
of the exponent. This appears to indicate that the "same" resistance state can be achieved through different conduction paths within the same device. This is consistent with the suggestion that the conduction path (the filament) might have an inhomogeneous structure consisting of a small-scale dispersion of metallic-like particles embedded in an insulating matrix; this structure may provide a variety of conduction paths.

It is possible to extend the range of the $I$-$V$ characteristics up to the critical field where switching occurs using very short, single voltage pulses of varying polarities and pulse heights. Figure 14.25 shows the room-temperature "pulsed" current-voltage characteristics of an analog V-$p^+$-Cr memory device. These characteristics were obtained using 400-ns pulses of progressively increasing magnitude and of both polarities. The 400-ns pulse length is long enough to observe a plateau in the pulse signal (i.e., RC effects are avoided). Positive polarity means that the Cr bottom contact is more positive than the top contact. Starting from a $9.5 \times 10^5 \Omega$ OFF state (measured at 0.5 V, curve (a) in Figure 14.25), the onset of the strong nonlinear rise in the current occurs at about $+1.1$V. Curve (a) is reproducible in that it can be repeated many times without a change in the device resistance, up to a voltage level of about $+1.7$V. Further increase in the voltage height will result in a permanent decrease in the device resistance and, consequently, a change in the characteristics. The decrease of resistance is determined by the magnitude of the maximum voltage applied.

Figure 14.24 $I$-$V$ characteristics of analog memory states plotted on log-log scale; characteristics are symmetrical about origin.
Figure 14.25 Pulsed I-V characteristics of Cr-p-V analog structure.

Curve (b) represents a memory-ON state \((R = 5.6 \times 10^4 \Omega \text{ at } 0.5V)\) of the device. Curve (b) is reproducible up to a voltage level of \(-5V\) if positive polarity is applied. Further increase in voltage might destroy the device. However, if a negative voltage is applied to the same ON state (curve (c)), an ERASE process is observed at voltage levels greater than about 1.7V. On the other hand, if an OFF state \((R = 9 \times 10^4 \Omega)\) is negatively biased (curve (d)), no change in the OFF-state resistance can be observed up to a voltage of about \(-5V\). The apparent polarity dependence suggests that the analog memory switching is not determined simply by the magnitude of the applied power or energy. This is further supported by a comparison of individual switching transients with opposite polarity. Figure 14.26 shows the waveforms of a single WRITE pulse (a) from OFF \((R = 2 \times 10^4 \Omega)\) to ON \((R = 3.8 \times 10^4 \Omega)\) and of an ERASE pulse (b) from OFF \((R = 2 \times 10^4 \Omega)\) to a slightly higher OFF state \((R = 2.4 \times 10^4 \Omega)\). In the second case there is no change in the memory state resistance although similar voltage and current levels are applied. The calculated total charge flowing through the sample is also similar: \(Q_w = 8.53 \times 10^{-11} \text{C}\) for the WRITE and \(Q_e = 7.24 \times 10^{-11} \text{C}\) for the ERASE pulse.

These measurements suggest that the memory-state resistance is determined by a combination of applied voltage level and appropriate polarity. On the other hand, the fact that the memory resistance is not determined by the applied power or energy suggests that the memory switching does not depend significantly on the internal temperature of the device. Figure 14.27(a) shows a switching transient of
Figure 14.26 Waveforms of single-pulse memory switching: (a) OFF→ON transient; (b) ON→OFF transient.

Figure 14.27 Waveforms of single-pulse memory switching: (a) ERASE transient at 300K; (b) ERASE transient at 4.2K.
an ERASE pulse at 300K, where the device resistance is changed from $R_{ON} = 3 \times 10^9 \Omega$ to $R_{OFF} = 6 \times 10^9 \Omega$. Figure 14.27(b) shows a similar ERASE transient at much lower temperature (4.2K) where the device resistance has also changed from $R_{ON} = 3 \times 10^9 \Omega$ to $R_{OFF} = 6 \times 10^9 \Omega$. It can be seen that despite the large temperature difference, the current level and the threshold voltage required to achieve the same ERASE process as at room temperature increase by less than a factor of two. It should also be emphasized that the device continues to operate even at liquid helium temperatures. The analog switching effect is still observed at 4.2K without significant changes either in the threshold voltage or the current level. This suggests that the electrical conduction and memory phenomena are possibly connected to a temperature-independent physical process, which we propose may be associated with tunneling, for example, or with conduction through a channel of very small dimensions in which metallic-like inhomogeneities are distributed.

14.2.2.1 Low-Temperature Quantization Effects in a-Si:H Structures

The behavior proposed in the previous section is supported by recent results obtained during investigations of the low temperature conductivity of the analog memory states [61], [62]. Typical $I-V$ characteristics of a formed memory-ON state at 4.2K are shown in Figure 14.28(a). In the voltage region from 0 to 0.36V, the current around zero bias is of the order of $\sim 10^{-8} \text{A}$ but increases to $\sim 10^{-6} \text{A}$ at voltages approaching 0.3V (i.e., a strong nonlinear behavior is found). It is important to emphasize that the room-temperature $I-V$ characteristics of the memory-ON state are linear. The observed large increase in the resistance around zero bias is shown in the inset to Figure 14.28(a) and is consistent with tunneling conduction between metallic particles embedded in an insulating matrix [63]. Further experimental evidence for tunneling conduction comes from the temperature dependence of the device current at low bias voltages. In this regime, $V$ is smaller than the tunneling barrier $\phi$ and the conduction is associated with field-assisted tunneling. The approximate form of the current density/voltage characteristics is [63]:

\[
J(V, T) = J(V, 0) \left(1 + \frac{3 \times 10^{-9}d^2T^2}{\phi}\right) \tag{14.12}
\]

where $d$ is in Angstrom units, $\phi$ is in electronvolts and $T$ is in Kelvin. The experimentally observed $T^2$ dependence is consistent with $d = 50\text{Å}$ and $\phi = 1 \text{eV}$.

At 0.36V, a current jump occurs and the resistance of the sample is lowered to the order of a few k\Omega. After the first current jump at 0.36V, further current steps can be observed at 0.47, 0.53, and 0.70V (curve A, Figure 14.28(a)). At 0.36V (henceforth called the critical voltage $V_*$) there is thus a dramatic change
in the behavior of the sample. At voltages lower than $V_{\sigma}$, no discontinuities are observed but at voltages higher than $V_{\sigma}$ the resistance is lowered and the current increases in discrete steps. The current-voltage characteristics are symmetrical (i.e., the same behavior is observed for the opposite polarity. Curve B in Figure 14.28(a) depicts the $I$-$V$ characteristics of the same sample under the influence of a $B = 0.2$ $T$ magnetic field. The curve has been displaced by 50 $\mu$A in the current scale for clarity. The direction of the magnetic field is 30° with respect to the conducting channel (the filament). Additional steps can be observed at 0.34, 0.42, 0.5, 0.59,
and 0.79V together with the steps observed in the zero magnetic field case (curve A). The effect of the magnetic field is reversible, and if the field is removed the I-V characteristics revert to the zero magnetic field case. A number of I-V characteristics have been obtained in which sharp and well-defined steps can be observed at 4.2K. However, it should be pointed out that the position of the observed steps is dependent on the direction of the voltage sweep (i.e., some hysteresis is observed as illustrated in Figure 14.28(b)).

The critical voltage $V_c$, at which the first current jump is observed and the magnitude of the first current jump are dependent on the resistance of the memory-ON-state investigated. Figure 14.29 shows the effect of changing the memory-ON-state resistance on the observed resistance steps. Resistance (in kΩ) is plotted on the right hand vertical axis; on the left, resistance is plotted in the dimensionless quantized units of $(h/2e^2)$. The critical voltage and the magnitude of the first jump increase with increasing resistance ($R_a > R_c > R_s > R_d$), but after the first jump the characteristics are rather similar, suggesting a similar conduction mechanism at higher voltages. The first jump appears to be associated with the formation of a highly conducting path within the structure whose characteristics are independent of the low-bias behavior (i.e., of the different memory states). Therefore, we propose that the first large current step (at 0.36V) is associated with the formation of a narrow, highly conducting channel that significantly lowers the resistance of the sample. With further increase in the applied voltage the resistance is lowered.

![Figure 14.29 I-V characteristics of different memory states at 4.2K ($R_c > R_s > R_d$).](image-url)
in steps, corresponding to quantized resistance values \( R = \frac{h}{2ie^2} \) where \( i \) is an integer. In the voltage range from 0.36 to 0.8V, there are four steps, with \( i \) being 2, 3, 4, and 5. These correspond to the observed current rises in curve A, Figure 14.28(a). Higher voltages have not been applied to the sample because this could change the resistance of the particular memory state. If a magnetic field is now applied to the sample, further quantization of resistance is observed at values \( R = \frac{h}{2(i + 1/2)e^2} \). This is illustrated in Figure 14.30(b), in which the data correspond to curve B in Figure 14.28(a). It can be seen that extra steps in the resistance occur at \( i = 2.5, 3.5, 4.5, \) and 5.5. The effect of temperature is illustrated in Figure 14.31. The curves have been shifted along the current axis for clarity. The observed current steps gradually decrease with increasing temperature, finally disappearing at the remarkably high temperature of 190K. In our more recent work [64], using samples that have undergone a further conditioning step, similar quantization effects have been observed at temperatures up to 400K.

![Figure 14.30](image)

In analyzing the main results of this work, two important facts should be emphasized. First, the observation of quantized electron transport provides a vital clue to the structure of the analog memory element. Second, the programmability of the analog memory provides information about the possible mechanism of the switching process itself. The analog memory effect in amorphous silicon metal-\( p^* \)-metal structures can only be observed if the sample is subjected to an initial forming.
process. The forming process is characterized not only by the breakdown of the high-resistance state of the structure but, more importantly, by the presence of a positive feedback mechanism that provides a low-resistance ON state so that the breakdown is nondestructive and repetitive switching is possible. Furthermore, the experimental results suggest a strong influence of the choice of the top metal contacts (summarized in Table 14.2) on the type of memory switching observed (digital or analog switching) and on the success of obtaining stable and reproducible switching. These results are in good accordance with the previous observations [65] that the first switching event (forming) causes a local structural modification of the $p^+$ amorphous silicon layer, producing a highly conducting filament that does not revert to the original amorphous material when the device is switched OFF. After the forming process, the $p^+$ devices usually exhibit a lower OFF resistance than the unformed device resistance, in contrast to our original data for $p^+\cdot n\cdot i$ devices [55]. The temperature dependence of the conductivity is also greatly reduced by the forming process. The area independence of $R_{ON}$ (Figure 14.15) suggests localized electrical conduction after forming. These results, together with the newly observed quantized electron transport phenomena (Figures 14.28–14.31), provide strong experimental evidence that the forming process creates a filamentary region consisting of a new material whose properties have changed significantly compared to those of the unformed original material. It is feasible that the high fields and

![Figure 14.31 I-V characteristics as function of temperature (curves displaced vertically for clarity).](image-url)
current densities present during forming result in high temperatures developing locally, which could lead to enhanced diffusion of metallic particles from the electrode into the thin amorphous film. Such a region would become the preferred current path carrying the electron current in the ON state. The I-V characteristics in the ON state (Figure 14.10) suggest that any material rearrangement within the filament occurs so as to destroy the rectifying properties of the original metal-p+ Schottky contact at room temperature; it should also be noted that this is the case even in a typical OFF state.

Lowering the temperature of the memory device reveals further information about the possible structure. The low-temperature I-V characteristics (Figure 14.28(a)) show that we have observed a zero-bias anomaly and quantized resistance in the ON state of amorphous silicon Cr-p+-V structures. The phenomenon of quantized electron transport is usually observed in the case where the mean free path of the electron $\lambda$ is larger than the length of the conducting channel, resulting in ballistic conduction [66, 67]. The usual approach to the fabrication of devices based on ballistic conduction is to use a very high mobility material (usually high-purity GaAs), where the mobility can reach values in the order of $\sim 10^5 \text{cm}^2\text{V}^{-1}\text{s}^{-1}$ leading to the values of electron mean free path:

$$\lambda \sim \tau v_m \sim \frac{\mu}{q} (3kTm^*)^{1/2} \sim 1 \mu\text{m} \tag{14.13}$$

where $\tau$ is the momentum relaxation time, $v_m$ is the thermal velocity, $T$ is the lattice temperature, $m^*$ is the effective mass of the electron, $q$ is the electronic charge and $k$ is the Boltzmann's constant. For the above estimate, $m^* \sim 0.068 m_e$ ($m_e$ is the free electron mass) has been used in the case of GaAs. The value of $\lambda \sim 1 \mu\text{m}$ is certainly longer than the device dimensions that can be achieved by modern submicron technology. The structure in which ballistic transport is most widely investigated is a GaAs-AlGaAs heterojunction with a split gate field-effect transistor (FET) configuration, usually less than 0.5 $\mu\text{m}$ in length, with a gap of about 0.7 $\mu\text{m}$ [65, 66]. As the voltage on the gate is made increasingly negative, the width of the conducting channel decreases and becomes sufficiently small that one-dimensional quantization occurs as the channel width becomes comparable to the electron wavelength. This is sufficiently short that electrons may pass through the conducting channel without appreciable scattering. Our structure is quite different and one possible mechanism that would explain the quantized transport in amorphous silicon structures is outlined below. The observed I-V characteristics at low temperatures show two distinct regimes, as shown in Figures 14.28-14.31. Below the critical voltage $V_c$, the curves are nonlinear but there are no discontinuities; however, above $V_c$ sudden jumps in current are observed associated with quantized resistance. These observations suggest that the current flow below and above $V_c$ is determined by different mechanisms.
The $I-V$ characteristics below $V_o$ (see inset, Figure 14.28(a)) suggest that the filament has a relatively large resistance around zero bias (i.e., a barrier for electron flow exists). The observed zero-bias anomaly is consistent with tunneling conduction between metallic particles embedded in an insulating matrix [63]. Below $V_o$, the $I-V$ characteristics show a continuous but strongly nonlinear behavior due to the creation of free carriers by the high electric field across the small tunneling distance. In such a system the creation of free carriers is an activated process related to the increase in the electrostatic potential of a particle when a free electron is added to it. The activation energy can be provided entirely by thermal energy, hence the effect diminishes at higher temperatures; or, in the presence of an applied field, part or all of it can be provided by the field itself. Experimentally, a $T^3$ dependence of the current at a constant voltage (less than $V_o$) is observed in good accordance with the assumption that the electron transport is dominated by field-activated tunneling processes in this region. Figure 14.32 illustrates an idealized model having a single permanent inclusion extending from the top contact, with a narrow channel connecting it to the bottom contact. The evidence [55] is that the overall diameter at the top contact is less than 0.5 $\mu$m. The length of the channel must be consistent with tunneling. With increasing applied voltage, the tunneling current increases exponentially. The sudden increase in current at $V_o$ implies that the tunneling

![Figure 14.32 Schematic description of filament showing proposed metallic inclusion and one-dimensional conducting channel.](image)
barrier effectively breaks down. This is not a destructive effect, however, as the process is completely reversible and no material changes ensue.

At present we have no firm explanation for the steps in the current above $V_{cr}$, which we have associated with quantized resistance. Experiments on other one- and two-dimensional systems show electrical properties that are associated with quantized resistance, but these are only observed at very low temperatures. For example, in a two-dimensional electron gas [66], the quantum Hall effect is related to $e^2/h$, and in a one-dimensional electron gas [67, 68] in the ballistic regime, resistance is quantized in terms of $h/2e^2$. However, it is not expected that quantized resistance states associated with ballistic transport should be observed when, as in the present case, the applied voltage is much greater than $kT$—if, for example, all the applied voltage were to appear across the critical part of the structure, the equivalent temperature would be in excess of 1000K. We have suggested elsewhere [64] that the quantized phenomena might be related to electrical transport through a quantum point contact. Assuming that the inclusion of modified material created during forming has a tapered shape (Figure 14.32), the resulting current flow could be restricted to a very localized area. The device resistance could thus be determined by the contact area where the tip of the modified region and the lower metal electrode are in close proximity. At such distances the resistance might be associated with either a single atom or a small number of contact atoms [69]. It has been shown theoretically [70] that the resistance associated with a single contact atom reaches saturation with a minimum value given by $R = h/2e^2$, provided that no elastic deformation occurs. This is the constriction resistance ($\approx 1290\Omega$) associated with an ideal conduction channel. It should be noted that this quantized resistance value is predicted assuming current flow through a single atomic orbital only. If more orbitals (conduction paths) are involved, these would reduce the quantized resistance value in integer steps according to the number of paths involved. Experimental results from a scanning tunneling microscope (STM) using an Ir tip showed a resistance jump to a value $R \approx 4 \times 10^7\Omega$ at close contact, rather larger than was predicted theoretically [71]. In this experiment the tunneling current is recorded as a function of distance $D$, from the conducting surface. The jump occurs at the transition from the tunneling to the one-atom point contact regime in the STM.

Our measurements show that the resistance of conditioned a-Si:H sandwich structures can be quantized under some circumstances. The voltages at which this occurs correspond to energies that are greatly in excess of the thermal energy $kT$. Furthermore, the quantized resistance values reflected in the $I-V$ characteristics do not involve all possible integer or half integer values. For these reasons we believe that existing theories of quantized resistance associated with ballistic transport are not applicable to our structure. However, the results obtained so far do not appear incompatible with current flow through atomic scale point contacts, and it is possible that this could explain the room temperature quantization observed. The most important feature of the observed quantized behavior is that it can be observed up to $\sim 400K$, much higher than previously reported.
The model outlined above describes one possible mechanism for the quantized electron transport and provides important information about the structure of the filament, but as it stands it does not explain the memory switching. The most important difference between the two phenomena lies in the fact that the quantized electron transport observed in memory devices is seen at applied voltage levels lower than ~1V, that is, at voltage levels where no memory switching would be expected to occur. The observed quantized jumps in the device resistance are threshold-switching type: if the voltage is lowered, the sample resistance reverts to the original low-bias case and there is no permanent change in the I-V characteristics of the device. On the other hand, memory switching occurs at voltage levels from ~1.5 to ~4V, resulting in a permanent change in the I-V characteristics of the device (i.e., a different resistance state of the memory).

14.2.2.2 Models for Memory Switching

Some of the possible mechanisms that might explain the memory-switching behavior are discussed below.

**Thermal Models**

The thermal models [26, 29] used to explain the behavior of the amorphous chalcogenide memories might appear to offer a basis for explaining the a-Si:H switching process. In the chalcogenide memories the ON state is associated with a filament of crystalline material that is formed after sufficient power has been applied to the layer to melt a small area of the material. Switching OFF is achieved by burning out this filament using a number of relatively short high-power pulses and allowing rapid quenching to reform the highly resistive amorphous phase. However, there are a number of important differences between the amorphous silicon and the chalcogenide memories:

1. It has been established that the a-Si:H memories do not form or WRITE at constant power; in general, forming occurs at much lower energy (<10^{-4} J) than in the chalcogenides (10^{-2}-10^{-4} J).
2. The forming, WRITE, and ERASE operations for the a-Si:H memories are generally polarity dependent.
3. No rise in the temperature of the a-Si:H specimens can be observed prior to switching.
4. The a-Si:H WRITE and ERASE times are many orders of magnitude shorter than those for the chalcogenides (e.g., 10^{-8} s for the WRITE operation compared with 10^{-3} s).

It is thus unlikely that the crystalline/amorphous thermal model is applicable to a-Si:H memories.
Models Based on Trapped Space Charge

In many respects the behavior of the a-Si:H layers appears to be closely related to that of crystalline-silicon MISS structures in that both show fast polarity-dependent switching, both show current instabilities, and both have high conductance states associated with current filaments. However, the crystalline MISS structures are threshold switches that always revert to the OFF state when the power is removed, whereas the a-Si:H structures have the additional feature (and complexity) of nonvolatile memory behavior. It is nonetheless possible that the initiation of memory switching in the amorphous silicon devices is closely related to the mechanism proposed to explain threshold switching in crystalline MISS structures. Essentially two models have been used to explain MISS behavior [10]. These are generally referred to as the “punch-through” and “avalanche” modes, and both mechanisms require that high fields be developed across space-charge barriers in the films. In addition, in both models the low-impedance (ON) state is produced by injected charge, causing inversion of the Si at the Si/insulator interface. It is tempting to suggest that the “permanent” memory of the a-Si:H layers may be produced by a similar mechanism in which the charge is trapped in deep gap states at the insulator semiconductor interface for which the probability of release is very small. However, the a-Si:H devices retain their ON-state memory conductance without any observable change for at least 1 year at room temperature and at least 24 hours at 95°C.

Using thermal release rates from deep midgap states of energy $E_r$ as a measure of the persistance of the trapped space charge, the average thermal release time ($\tau \exp[-(E_r - E_f)/kT]$) indicates that the capture cross section of these centers would have to be less than $10^{-16}$ cm$^2$ in order to agree with the experimental data. Although extremely small, such values would be consistent with Coulomb repulsive centers identified in crystalline materials. However, the problem is that recombination of the trapped charge distribution through tunneling or diffusion may well invalidate the above estimate by leading to a much faster decay of any trapped space charge distribution. All that can be said at present is that a model in which the observed memory is associated with a trapped space charge cannot be excluded, but in view of the remarkable nonvolatility of the memory states, it is unlikely to be the basis of anything more than a partial explanation.

Hydrogen Motion in a-Si:H

It is known that significant amounts of hydrogen are incorporated in the random network of the a-Si:H layer. The possibility exists, therefore, that memory switching may be associated with some atomic motion of hydrogen in the material. For instance, it has been reported that in $n$-type crystalline-silicon-Cr Schottky barrier
structures, hydrogen plays an important role in lowering the contact barrier [74]. Also, the polarity dependence of the threshold voltages for the a-Si:H memories could be understood on the basis of field-assisted diffusion.

**Tunneling Method**

One possible explanation for the quantized electron transport and the zero-bias anomaly observed at lower temperatures is that tunneling processes are involved and hence must be taken into account when suggesting a model for memory switching. In order to explain the different and permanent resistance states, the model should embody the following features:

1. Changes in the conductivity of the whole filament;
2. Local changes in conductivity at a certain point(s) in the filament;
3. Changes in the filament geometry.

Thermal models would presumably require an average energy for switching of an order similar to that established for forming, but it has been found that ON $\rightarrow$ OFF $\rightarrow$ ON switching can be achieved using single-pulse energies of $\sim 10^{47}$ J (Figure 14.26). In addition, if it is assumed that the filament is homogeneous, there is no obvious way of introducing a polarity dependence. However, if there are inhomogeneities in the filament it may be possible to substantially increase the local energy density and field strength. Consider, for example, a simple one-dimensional filament consisting of alternating regions of material A and B. The observed resistance of the filament will depend on the relative contributions to the filament resistance of A and B type material, that is, if $R_A < R_B$, the properties of the A type will be observed, and if $R_A < R_B$, the properties of the type B material will be observed. If highly conducting regions are separated by small gaps, it is also necessary to consider the possibility of inter-island tunneling. In order to establish such a model, a more detailed description of regions A and B would have to be included, for example:

1. Region A is metallic or a degenerately doped material whose contribution to the filament is fixed and $\sigma_A = f(V)$ where $V$ is the applied voltage and $\sigma$ is the conductivity.
2. Region B is a gap between A-type material that can be sufficiently small to allow significant tunneling to occur; $\sigma_B$ is some function of the voltage.

In the ON state, $R_B < R_A$ and the measured conduction properties of the filament are those corresponding to material $A$, namely, ohmic $I$-$V$ characteristics and small thermal activation. In the OFF state, $R_A < R_B$ and the conduction properties are governed by the behavior of the tunneling regions. The current density in such a one-dimensional system is governed by (1) the applied bias $V$, (2) the barrier thickness $d$, (3) the barrier average height $\Phi$ and asymmetry $\Delta \Phi$,
(4) the density of states \( g(E) \) in the metallic regions, and (5) the occupation of \( g(E) \) (i.e., the temperature dependence).

In this case, the functional form of the current density \( J \) versus voltage \( V \) at different temperatures \( J(V, T) \) can be obtained under two limiting conditions. First:

\[
V < \Phi, \text{ and } \Delta \Phi < \Phi, \text{ constant } T
\]

In this case the transmission probability \( P = c \exp(-d\Phi^2) \) may be treated as a constant for electrons near the Fermi energy \( E_F \). Provided that \( g(E) \) variations near \( E_F \) are small, \( J \) will increase linearly with \( V \) as more empty states become accessible in metallic region A (i.e., conduction is approximately ohmic and symmetric). This is consistent with the experimental data shown in Figure 14.12. Second:

\[
V < \Phi, \text{ and the temperature is varied}
\]

In this case \( P \) is essentially independent of \( T \) because usually \( kT \sim 10^{-2} \text{eV} \). Thus only a few electrons are promoted thermally to levels where \( P \) is significantly larger than at \( E_F \). The approximate form of thermal \( J-V \) characteristics in this regime is as given in Equation (14.12).

How might switching be achieved in terms of the tunneling model? The low-bias isothermal resistance of region B-like parts of the filament is determined by geometric factors (included in constant \( c \)) and the transmission probability. Therefore one obtains:

\[
J = cV \exp(-d\Phi^2)
\]

or

\[
R = c \exp(d\Phi^2)
\]

The exponent is almost numerically correct if \( d \) is in \( \text{Å} \) and \( \Phi \) is in eV. Taking \( d = 50 \text{Å} \) and \( \Phi = 1 \text{eV} \), then \( R_1 = c \exp(50) \) and increasing \( d \) by 5\( \text{Å} \) gives \( R_2 = \exp(55) \) (i.e., \( R \) has increased by about a factor of 150). An equivalent increase in resistance is obtained by keeping \( d \) constant and increasing \( \Phi \) from 1\( \text{eV} \) to 1.2\( \text{eV} \). Changes in the tunnel barrier width on an atomic scale, or small (20%) changes in barrier height, could therefore account for the presence of switching. It is not clear, however, how these could occur in a manner that would explain the systematic dependence of device resistance on the magnitude and polarity of the switching voltages as shown, for example, by the data in Figures 14.22(a) and 14.22(b).

Finally we should note that in this model the cross section of the conducting channel is considered to be constant. The possibility exists that memory switching
might be associated with some permanent change to the cross section of the conducting channel (i.e., the high currents involved might cause small-scale material rearrangements).

In summary, we can say that, based on the available experimental evidence, it is possible that memory switching in the amorphous silicon metal-$p^+$-metal structures might be associated with a tunneling barrier within a permanent filament produced by the forming process. A change in the particle size or spatial distribution as a result of localized heating or high-field effects could explain the differences between high- and low-resistance states. However, the high switching speeds (~10 ns) and the insensitivity of switching to temperature suggest that large-scale structural changes are unlikely. It is possible though, that a structure might arise with small variations at a "weak link" in the filament that could produce the observed changes in a short time. As tunneling is a quantum mechanical effect, the nanosecond time scale for the switching is feasible.

Finally we emphasize that the effect has been observed in what is initially a metal-amorphous semiconductor structure. It is not known with any certainty what effect the forming process has on the structure of the conducting channel, but the observed quantized behavior suggests the importance of the very small dimensions of the conducting channel rather than other physical parameters. The significance of using amorphous silicon sandwich structures lies in the forming process, that is, in the process that allows the fabrication of such small structures. The strong influence of the metal contact on the observed memory switching behavior indicates the importance of alloying during the forming process. It has been reported, for example, that solid-phase amorphization or glass formation occurs in V-Si reactions induced by rapid thermal annealing (i.e., at conditions similar to forming) but does not occur in Co-Si and Cr-Si [75]. Therefore, the presence of a very small tunneling conduction path in amorphous silicon $V-p^+-Cr$ structures might be due to a new type of solid-phase reaction, induced by forming, that creates a homogeneous (possibly amorphous) V-Si silicide.

14.3 CONCLUSION

Research into electronic switching in amorphous semiconductor thin films has a long history, with several distinct stages involved. Early work was concentrated on chalcogenide thin-film devices in which the basic switching took place between highly conducting ON and highly resistive OFF states (that is, an essentially digital mode of operation was involved). Memory switching in chalcogenide glasses is based on a crystalline/amorphous phase transition, whereas threshold switching in these glasses can be described by purely electronic mechanisms for films ≤1 μm thick or by electrothermal considerations for film thickness from 1 to 10 μm.

More recent observations on amorphous silicon thin-film structures have shown that, in contrast to chalcogenide thin films, these devices can exhibit an
analog behavior under certain conditions. This analog switching mechanism appears to be associated with interatomic processes such as tunneling, and, as a consequence, is significantly faster than switching in chalcogenide glasses.

Finally, our most recent results have demonstrated that in certain a-Si memory structures, quantized electron transport phenomena occur. These results are most surprising considering the magnitude of the effect and the high temperatures involved and appear to be a consequence of the extremely small (atomic scale) dimensions of the structures involved.

REFERENCES


**SYMBOLS**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Magnetic field strength</td>
</tr>
<tr>
<td>d</td>
<td>Layer thickness</td>
</tr>
<tr>
<td>dₙ</td>
<td>n-layer thickness</td>
</tr>
<tr>
<td>dᵣ</td>
<td>Phase boundary diameter of phase transition</td>
</tr>
<tr>
<td>e</td>
<td>Electronic charge</td>
</tr>
<tr>
<td>E</td>
<td>Electric field (V/cm⁻¹)</td>
</tr>
<tr>
<td>Eᵣ</td>
<td>Critical field for switching</td>
</tr>
<tr>
<td>Eₛ</td>
<td>Fermi energy</td>
</tr>
<tr>
<td>ξ</td>
<td>Energy</td>
</tr>
<tr>
<td>jᵣ</td>
<td>Recombination current density</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
</tr>
<tr>
<td>$J$</td>
<td>Current density</td>
</tr>
<tr>
<td>$k$</td>
<td>Boltzmann's constant</td>
</tr>
<tr>
<td>$g(E)$</td>
<td>Density of states (DOS)</td>
</tr>
<tr>
<td>$G_{OFF}$</td>
<td>OFF-state conductance</td>
</tr>
<tr>
<td>$G_n$</td>
<td>Gradient of phase boundary diameter</td>
</tr>
<tr>
<td>$I_H$</td>
<td>Critical holding current</td>
</tr>
<tr>
<td>$I_i$</td>
<td>Steady current</td>
</tr>
<tr>
<td>$I_{\max}$</td>
<td>Maximum current</td>
</tr>
<tr>
<td>$m^*$</td>
<td>Electron-effective mass</td>
</tr>
<tr>
<td>$m_e$</td>
<td>Free electron mass</td>
</tr>
<tr>
<td>$N_\omega$</td>
<td>Average number of phonons of frequency $\omega$</td>
</tr>
<tr>
<td>$P_{RMS}$</td>
<td>RMS power</td>
</tr>
<tr>
<td>$P$</td>
<td>Transmission probability</td>
</tr>
<tr>
<td>$P_n$</td>
<td>Intercept of phase boundary diameter</td>
</tr>
<tr>
<td>$q_{inj}$</td>
<td>Injected charge</td>
</tr>
<tr>
<td>$Q$</td>
<td>Charge</td>
</tr>
<tr>
<td>$R_{OFF}$</td>
<td>OFF-state resistance</td>
</tr>
<tr>
<td>$R_{ON}$</td>
<td>ON-state resistance</td>
</tr>
<tr>
<td>$t_s$</td>
<td>Switching delay time</td>
</tr>
<tr>
<td>$t_{OW}$</td>
<td>WRITE delay time</td>
</tr>
<tr>
<td>$t_{DE}$</td>
<td>ERASE delay time</td>
</tr>
<tr>
<td>$T$</td>
<td>Temperature</td>
</tr>
<tr>
<td>$T_e$</td>
<td>Electron temperature</td>
</tr>
<tr>
<td>$V_c$</td>
<td>Critical forming voltage</td>
</tr>
<tr>
<td>$V_{cr}$</td>
<td>Critical voltage</td>
</tr>
<tr>
<td>$V_F$</td>
<td>Forming voltage</td>
</tr>
<tr>
<td>$V_H$</td>
<td>Holding voltage</td>
</tr>
<tr>
<td>$V_{th}$</td>
<td>Threshold voltage</td>
</tr>
<tr>
<td>$V_{FTh}$</td>
<td>Forward bias threshold voltage</td>
</tr>
<tr>
<td>$V_{RTh}$</td>
<td>Reverse bias threshold voltage</td>
</tr>
<tr>
<td>$\Delta p/p(B)$</td>
<td>Transverse magnetoresistance</td>
</tr>
<tr>
<td>$\Delta V$</td>
<td>Voltage window for analogue memory programming</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>Conductivity</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Electron mean free path</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Mobility</td>
</tr>
<tr>
<td>$\pi$</td>
<td>Tunneling barrier height</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Relaxation time</td>
</tr>
<tr>
<td>$\tau_D$</td>
<td>Forming delay time</td>
</tr>
<tr>
<td>$\tau_c$</td>
<td>Mean time between electron-lattice collisions</td>
</tr>
<tr>
<td>$\omega$</td>
<td>Frequency</td>
</tr>
<tr>
<td>$\Omega$</td>
<td>Resistance</td>
</tr>
<tr>
<td>ACRONYMS</td>
<td>DESCRIPTION</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
</tr>
<tr>
<td>CCNDR</td>
<td>Current-controlled NDR</td>
</tr>
<tr>
<td>CFO</td>
<td>Cohen-Fritsche-Ovshinsky</td>
</tr>
<tr>
<td>FET</td>
<td>Field-effect transistor</td>
</tr>
<tr>
<td>I-V</td>
<td>Current-voltage</td>
</tr>
<tr>
<td>MISS</td>
<td>Metal/i-n-p*/metal</td>
</tr>
<tr>
<td>NDR</td>
<td>Negative differential resistance</td>
</tr>
<tr>
<td>TONC</td>
<td>Transient ON characteristics</td>
</tr>
<tr>
<td>VCNDR</td>
<td>Voltage-controlled NDR</td>
</tr>
</tbody>
</table>
Appendix - List of Publications August 1994

(Where there are co-authors, the number in parenthesis indicates my position in the list of authors).


37. (2) The Mobility of Photo-induced Carriers in Disordered As$_2$Te$_3$ and As$_{50}$Te$_{48}$Si$_{12}$Ge$_{10}$, with J M Marshall. Phil Mag, Vol. 31, 1341-1356, (1975).

38. (2) Field Effect Measurements in Disordered As$_{30}$Te$_{40}$Si$_{12}$Ge$_{10}$ and As$_2$Te$_3$, with J M Marshall. Phil Mag, Vol. 33, 457-474, (1976).


54. (2) Growth of As$_2$Se$_3$ Single Crystals by an Open-Tube Vapour Transport Technique, with C Hampton and P McA Dryburgh. 2nd European Conf on Crystal Growth, Abstracts Book C132 (September 1979).


* THIS PAPER WAS AWARDED THE MAXWELL PREMIUM, FOR 1982, BY THE INSTITUTION OF ELECTRICAL ENGINEERS.


67. (2) The Electronic Properties of Glasses in the Cu$_4$(As$_{0.4}$Se$_{0.6}$) System with x Between 5 and 30, with M I Fraser. J Non-cryst Sol, Vols 59 and 60, 1031-1034, (1983).


150. (4) Index of Refraction Measurements in Photodoped Ag—As$_{33}$S$_7$ Films, with T I Kosa, T Wagner and P J S Ewen. Phil.Mag. B. In press.

Patents

1. Improvement in and relating to photo-conductive cells: Ultrasonic stirring during lead sulphide photocell depositions. British Patent 29557/56.


Invited Speaker at Numerous Conferences, Summer Schools, Colloquia, etc, eg, since 1982 at: (abbreviated list)

International Centre for Theoretical Physics, Trieste (Italy). Lectures on "The Glass Transition" at Spring College on Amorphous Semiconductors and the Liquid State, May-June 1982.

NATO Summer School on "Coherence and Energy Transfer in Glasses", Clare College, Cambridge University, England, September 1982. Lectures on "Electronic Transport in Chalcogenide Glasses".

SERC Postgraduate School on "Non-Crystalline Materials", University of Leicester, March 1983.

10th Int Conf on Amorphous and Liquid Semiconductors, Tokyo, August 1983. Invited Lecturer.

NATO Advanced Study Institute on "Glass - Current Issues", Tenerife, April 1984.

University of Reading, Department of Physics, November 1984.

University of Nottingham, Department of Physics, Feb 1986.

XIV Int Congress of Glass, Delhi, March 1986. Special invited lecture on "Glass in Electronics".


International Conference on Amorphous Metals and Semiconductors, University of Hyderabad, India, March 1987. Invited speaker.


University of East Anglia, Department of Physics, April 1988.

Texas A&M University, Department of Physics, March 1989.

Jet Propulsion Laboratory (JPL), California Institute of Technology, Pasadena, California, April 1989.

Bellcore Research Laboratories, New Jersey, USA, August 1990.

A.F. Ioffe Physico-Technical Institute, St Petersburg (Leningrad), June 1991.


University of Leeds, Houldsworth School of Applied Science, April 1993.
