JOURNAL OF MATERIALS SCIENCE 20 (1985) 3487–3500

A grain-boundary defect model for instability/stability of a ZnO varistor

T. K. GUPTA*, W. G. CARLSON
Westinghouse Research and Development Center, Pittsburgh, Pennsylvania 15235, USA

A defect model for the grain-boundary barrier has been proposed to explain the phenomena of voltage instability/stability of the ZnO varistor. The key element of the proposed model is the zinc interstitials which are present in the depletion layer as excess zinc, arising from the non-stoichiometric nature of ZnO. Both instability and stability have been described in terms of diffusion of these interstitials in the depletion layer, followed by chemical interactions with defects at the grain-boundary interface. Finally, a large body of experimental data is presented to indirectly validate the proposed defect model.

1. Introduction

Zinc oxide varistors [1–4] are known to exhibit a degradation in electrical characteristic when stressed continuously by an external field. This degradation phenomena [5–12] has been studied under a.c., d.c. and pulsed electric fields and several mechanisms [5, 6, 8, 11, 12] have been proposed to explain the observed degradation, e.g. electron trapping, dipole orientation, ion migration and oxygen desorption. Among these, ion migration has found strong support on the basis of experimental evidence. There is also a strong indication that the predominant migrating ions are the zinc interstitials [11] whose presence in the depletion layer gives rise to a metastable component [13] in the Schottky barrier [3]. In previous studies [13, 14], it has been shown that this metastable component is responsible for the instability of the varistor. It now appears that a large body of additional experimental results can be explained by invoking this concept of metastable barrier consisting of zinc interstitials. This has led to the gradual emergence of an atomic defect model for the grain-boundary barrier. The purpose of this paper is to develop this defect model as it emerges from past studies, and to present experimental evidence which is consistent with this model. The key element of this model is the zinc interstitials, and it will be shown that both instability and stability can be explained by the presence or absence of these interstitials.

2. Defect model

2.1. Defect distribution in the barrier

The basic concept of our model is that the depletion layer comprising the barrier consists of two components: a stable component consisting of *spatially fixed* positively charged ions and a metastable component consisting of *mobile* positively charged zinc interstitials. The former ions are the trivalent substitutional ions, the so-called donor ions, $D_{Zn}^{\cdot}$ (D = Bi, Sb, etc) and the native oxygen vacancies $V_0^{\cdot}$, $V_O^{2\cdot}$, whereas the latter ions are the singly and doubly charged native zinc interstitials $Zn_i^{\cdot}$, $Zn_i^{2\cdot}$. Recent studies [15] have assigned $V_O^{\cdot}$ and $V_{O_2}^{2\cdot}$ donor levels at 0.5 and ~ 2 eV below the conduction band and $Zn_i^{\cdot}$ and $Zn_i^{2\cdot}$ donor levels at 0.05 and 0.2 eV below the conduction band edge. These positively charged ions extend from both sides of the grain boundary into the adjacent grains and are compensated by a layer of negatively charged ions at the grain-boundary interface, comprised

Present address: Aluminium Company of America, Alcoa Technical Center, Alcoa Center, PA 15069, USA

0022–2461/85 $03.00 + .12

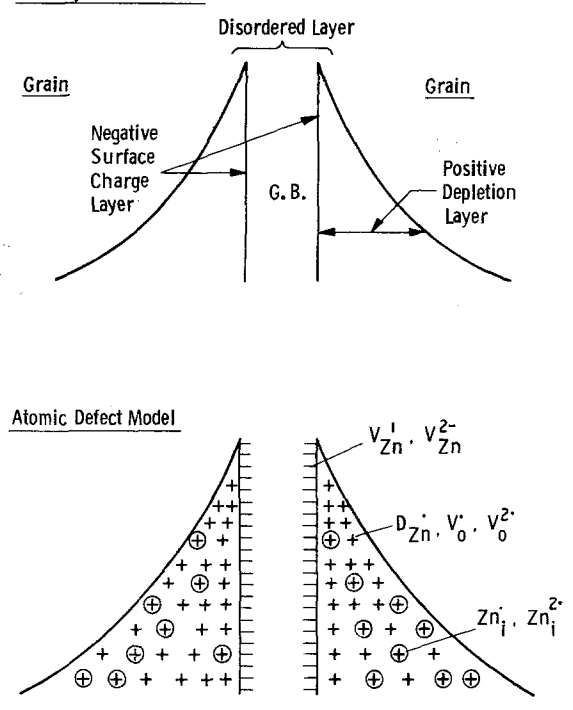

Figure 1 The proposed atomic defect model for Schottky barrier at the grain boundary. Also shown is the analogy with band model.

primarily of native zinc vacancies, V'_{Zn} and $V^{2'}_{Zn}$ whose acceptor levels have been assigned [15] at 0.7 and ~ 2.8 eV, respectively, above the valence band edge. The oxygen interstitials, O'_i and O''_i, are not considered as major defect types in ZnO. For simplicity, we assume that these charges arise [16] because the energy of formation for cation and anion vacancies or interstitials at the grain boundary are different, especially in the presence of donor atoms. The existence of grain-boundary charge has been confirmed in other polycrystalline ceramics, e.g. NaCl, Al_2O_3 and MgO wherein the sign and magnitude of the charge depended on the concentration of solute atoms. An electrical neutrality in the bulk of the crystal is, however, maintained at equilibrium, with the negative charge on the grain boundary being balanced by an equal and opposite space-charge pentrating some distance (depletion layer thickness) into the crystals as shown schematically in Fig. 1.

Fig. 1a represents the band model described previously [3]; Fig. 1b illustrates the corresponding defect model for the barrier. The negative charges on the grain boundary are balanced on both sides by positive charges in the depletion layers in the adjacent grains. The important feature about the charges in the depletion layer is that the spatial locations of these positive ions are different: the substituted ions and vacancies are located on the lattice sites whereas the zinc interstitials are located on the interstitial sites in the ZnO (Wurtzite) structure [18]. This situation results in an important difference in the relative mobility of ions. Since all the octahedral and half the tetrahedral interstitial sites are empty in the ZnO structure [18]; zinc interstitials can rapidly migrate within the structure via these interstitial sites [19]. These interstitials are analogous to highly mobile ionic charges that exist in $Si–SiO_2$ system [20] as impurities. On the contrary, the host lattice ions or the ions substituted on the host lattice sites must migrate via vacant lattice sites (vacancy), whose numbers are thermodynamically fixed at a given temperature [21]. As a result, the migration of these ions are severely restricted at the typical operating temperature of the varistor, first, owing to low concentration of vacancies at the temperature of interest, and second, to high activation energies for migration of these ions via vacant lattice sites. By and large, these ions are thus spatially fixed at the typical operating temperature where the interstitials have substantial mobility. The reported activation energy for the migration of interstitials [19] is ~ 0.55 eV, whereas the activation energies for lattice diffusion of zinc [22] and oxygen [23] are 3.2 and 7.2 eV, respectively.

An additional feature of this model which is analogous to the band model proposed earlier is that the grain boundary behaves as if it were a "disordered" layer [3]. The concept of disordered layer, although described previously, has, however, been misinterpreted in a recent article [24]. This disordered layer can be identified with the asymmetric tilt boundary [25], with the dislocations [26] and with the "core" region considered in diffusion models [27] for sintering and grain growth. For the purpose of our model, we assume that this disordered layer has two characteristics. It (1) provides a rapid diffusion path for anions (oxygen), and (2) acts as an infinite source and sink for neutral vacancies, V^x. These assumptions are consistent with the current thinking of ceramic grain boundaries [17]. They also provide necessary mechanisms for maintaining charge neutrality and lattice conservation. The appropriate equations for

electroneutrality and lattice conservation are:

$$n + [V'_{Zn}] + 2[V''_{Zn}] = p + [V_O^{\cdot}] + 2[V_O^{\cdot\cdot}] + [Zn_i^{\cdot}] + 2[Zn_i^{\cdot\cdot}] + [D_{Zn}^{\cdot}] \quad (1)$$

$$[Zn_{Zn}^{x}] + [V_{Zn}^{x}] + [V'_{Zn}] + [V''_{Zn}] + [D_{Zn}^{\cdot}] = [O_O^{x}] + [V_O^{x}] + [V_O^{\cdot}] + [V_O^{\cdot\cdot}] \quad (2)$$

In the above equation, n and p are electrons and holes and, Zn_{Zn}^{x} and O_O^{x} are the neutral zinc ions on zinc lattice sites and neutral oxygen ions on oxygen lattice sites.

2.2. The characteristics of zinc interstitials

As will be shown later, a large body of experimental results related to varistor stability can be attributed to the presence of zinc interstitials in the depletion layer. It is, therefore, necessary, before we describe the experimental results, to take a close look at the origin and the characteristics of the interstitials.

It is the most predominant defect type [18, 20] in ZnO as shown experimentally with varying zinc and oxygen partial pressures [28]. The interstitials are accommodated as excess metal in the large interstices of the ZnO structure [21]. The amount of excess zinc retained at room temperature as zinc interstitials [28–30] has ranged from 48 to 765 ppm, depending on temperature, prior history, and oxygen or zinc pressure. However, the concentration of zinc interstitials is not uniform throughout the ZnO crystals under conditions typically encountered by ZnO varistor compositions during processing. As the varistor is cooled through the fabrication temperature, the interstitials nearer the surface (grain boundary) can diffuse out, but those away from the grain boundary will gradually "freeze-in" in the structure [31] since the diffusivity decreases exponentially with cooling temperature and the diffusion distance becomes increasingly large. The grain-boundary region is, therefore, depleted of zinc interstitials as the temperature is lowered while the grain interior maintains a higher concentration. There is a further depletion of interstitials in the grain-boundary region due to the presence of donor atoms [15] such as bimsuth. However, a complete depletion of the interstitials is never attained in practice due to the exponential decay of diffusion coefficient with cooling temperature. It has been estimated [32] that below a critical temperature of 400 to 500° C, the outward diffusion of interstitials can no longer keep up with the cooling rate and the interstitials are simply "frozen-in" in the structure. As a result of this, a certain fraction of the total "frozen-in" interstitials are "trapped" in the depletion layer in spite of the ability of the grain boundary to better equilibrate with the environment. It is these *trapped interstitials* which have the most deleterious effect on stability as will be shown later.

The second important characteristic of the interstitials lies in their ability to migrate rapidly through the crystal structure. As discussed previously, this arises from the favourable crystal structure [18] and a low migration energy [19]. Finally, the interstitials can readily ionize, the ionization energy being of the order of 0.05 eV for single ionization, and 0.2 eV for double ionization [15].

These three facts taken together make the interstitials the most dominant and mobile defects in the depletion layer. The effect of other atomic defects [33, 34], if any, is believed to be of secondary importance in the determination of stability. A large body of experimental results can be directly attributed to the presence or absence of the interstitials. The discussion that follows presents these experimental results in terms of this defect model. These tests [11, 13, 14] were conducted under a.c. stress.

3. Discussion of experimental results

3.1. Diffusion and chemical interaction during electrical stressing

It has been shown elsewhere [13] that when an unstable varistor is stressed by a.c. voltage, there is a continuous reduction of barrier voltage with time. The barrier voltage has been previously defined [11] as the voltage at 0.5 mA cm^{-2} and designated as $E_{0.5}$. A similar reduction is also expected for barrier height since it is related to barrier voltage [38]. The reduction in barrier voltage is accelerated by an increase in field and/or temperature. The phenomenon is reversed when the varistor is de-energized and can be cycled [13]. This behaviour can be explained by interstitial diffusion and chemical interaction with the grain-boundary defects.

It is proposed that both the electric field and the elevated test temperature provide the

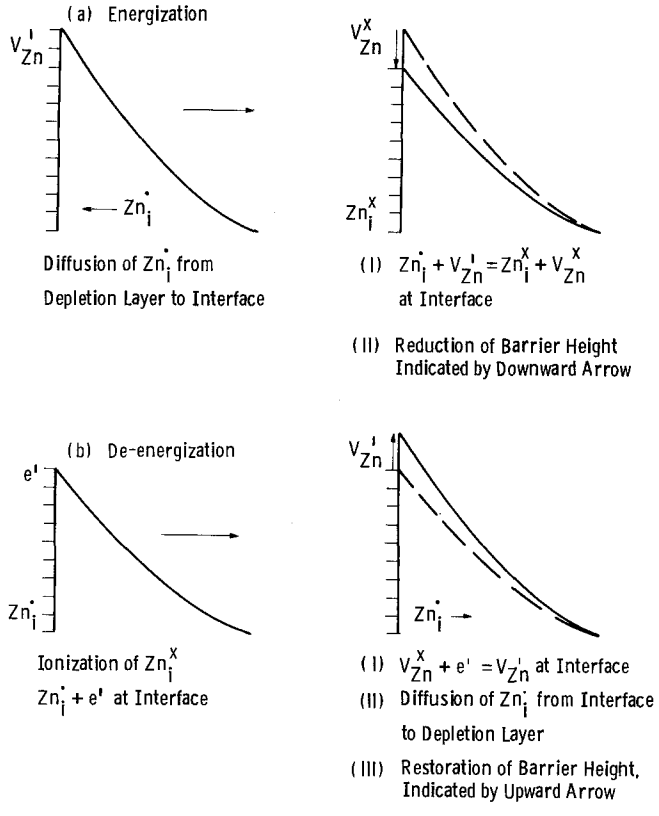

Figure 2 Defect diffusion and chemical interaction at the grain boundary during: (a) energization, (b) de-energization.

necessary driving force for the interstitial migration in the depletion layer. Assuming a typical barrier voltage [3] of ~ 3 volt/grain and depletion layer thickness of ~ 25 nm on either side of the grain boundary, the driving force due to applied voltage alone is of the order of $\sim 10^6\,\mathrm{V\,cm^{-1}}$, with an applied voltage typically set at 80% of the barrier voltage. It is suggested that this electical driving force, in combination with the thermal driving force that prevails at the test temperature ($> 100^\circ$ C), causes the positively charged interstitials to migrate towards the negatively charged grain-boundary interface where charged defects are converted to neutral defects due to chemical reaction between defects:

$$Zn_i^{\cdot} + V'_{Zn} = Zn_i^x + V^x_{Zn} \qquad (3)$$

For simplicity, only the reactions between singly charged ions are considered in this paper. Also the presence of associated defects is neglected. The individual components of the above reaction can be expressed as:

$$V'_{Zn} = V^x_{Zn} + e'$$
$$Zn_i^{\cdot} + e' = Zn_i^x \qquad (4)$$

As a result of these reactions at the interface, a positive charge ($Zn_i^{\cdot}$) in the depletion layer and a negative charge (V'_{Zn}) at the interface are lost, whereas two neutral defects (Zn_i^x and V^x_{Zn}) are formed instead at the grain-boundary interface. Of the latter defects, V^x_{Zn} readily disappears at the grain boundary (sink), leaving only the neutral Zn_i^x at the interface. With the continuous stressing of the varistor, the neutral Zn_i^x keeps accumulating at the interface due to continuous depletion of opposite charges from the adjacent reservoirs of positive and negative charges. The barrier height and barrier voltage are thus continuously reduced by stressing as the concentration of compensating charges are reduced in equal proportion (Equation 3) in the depletion layer as well as in the grain-boundary interface. This is schematically shown in Fig. 2a.

Since there are two depletion regions [3], migration occurs on both sides of the grain-boundary interface with alternation of field polarity [11, 13]. Charge neutralization thus occurs in both depletion layers resulting in a symmetrical lowering of the barrier height [5, 6] under a continuous a.c. stress.

Exactly opposite reactions occur upon removal of the field (driving force) and the

barrier height and barrier voltage are recovered [13]. The reaction starts with the ionization of the accumulated neutral Zn_i^x at the interface

$$Zn_i^x = Zn_i^{\cdot} + e' \quad (5)$$

followed by the capture of e′ by the neutral zinc vacancy V_{Zn}^x which is readily available at the grain boundary (vacancy source):

$$V_{Zn}^x + e' = V_{Zn}' \quad (6)$$

The interface thus regains a negative charge, V_{Zn}' and the corresponding compensating positive charge, $Zn_i^{\cdot}$ diffuses back to the depletion layer which thus regains its positive charge. Again, the prevailing test temperature provides the necessary driving force for the reverse diffusion of the interstitials. The process continues until all the neutral Zn_i^x interstitials are depleted of electrons and exhausted to the adjacent depletion layers. The original prestressed charge configuration is thus restored and the barrier height and barrier voltage are also recovered. This is schematically shown in Fig. 2b.

It should be noted that both diffusion and chemical reactions occur during energization as well as de-energization. However, since diffusion is a slower process, it is reasonable to state that the diffusion is rate controlling in each case. The time-dependent instability phenomena [11] in ZnO varistors can then be directly related to the diffusion of zinc interstitials.

The experimental data related to change in barrier voltage have been presented elsewhere [13, 14]. In the following sections we present the experimental data related to change in barrier height to support the interstitial migration mechanism presented above.

3.2. Effect of interstitial diffusion on barrier height

The model predicts that in an unstable device the barrier height, ϕ_B, should decrease continuously with time upon stressing. The barrier height, ϕ_B, is defined [3] by the conventional current transport equation as

$$J = J_0 \exp(-\phi_B/kT) \quad (7)$$

where J is the current density, J_0 is the pre-exponential factor and kT the usual thermal energy. Assuming a Schottky type of barrier, ϕ_B can be expressed in terms of materials parameters [35]:

$$\phi_B = \frac{N_d e l_D^2}{2\varepsilon\varepsilon_0} \quad (8)$$

with [36]

$$l_D = \left(\frac{A}{C}\right)\varepsilon\varepsilon_0 \quad (9)$$

where, ε is the relative dielectric constant $\simeq 16$ [2], $\varepsilon_0 = 8.85 \times 10^{-14}\,\mathrm{F\,cm^{-1}}$, e the electron charge $= 1.602 \times 10^{-16}$ coulomb, N_d the donor density $\simeq 3 \times 10^{17}\,\mathrm{cm^{-3}}$ [35], C the capacitance (F), and A the area associated with capacitance. Combining Equations 8 and 9,

$$\phi_B = \left(\frac{N_d e}{2}\right)\left(\frac{A}{C}\right)^2 \varepsilon\varepsilon_0 \quad (10)$$

which upon substitution of the above value and rearrangement yields,

$$\phi_B/A^2 = \frac{3.4 \times 10^{-14}}{C^2} \quad (11)$$

wherein the quantity A, being unknown but constant for a given device, is lumped with ϕ_B.

The change in barrier height with time upon stressing can thus be estimated in terms of the left-hand quantity of Equation 11 from a knowledge of the capacitance with time. However, it is convenient instead to measure capacitive current, I_C, as a function of time. The sample reactance X_C yields a relation between I_C and C:

$$X_C = \frac{V_{pk}}{I_C} = \frac{1}{2\pi f C} \quad (12)$$

where V_{pk} is the applied peak a.c. voltage and f is the frequency. Combining Equations 11 and 12 gives barrier height,

$$\phi_B/A^2 = \frac{3.4 \times 10^{-14}}{I_C^2}(2\pi f V_{pk})^2 \quad (13)$$

in terms of capacitive current which can be readily measured as a function of time.

The typical test data for the capacitive current, I_C, as a function of time when the varistor is stressed by a 60 Hz external voltage is shown in Table I. The test was conducted at 784 V which is equivalent to 80% of the "turn-on" peak voltage, $E_{0.5}$, defined elsewhere [11]. The current was measured with a Tektronix 576 curve tracer using the procedure described previously [11]. The table shows that whereas the capacitive current and capacitance increase continuously with time, the barrier height is

TABLE I Estimate of C and ϕ_B/A^2 from capacitive current measurement. Sample no. 19782005, V_{pk} (applied) = 784 V, f = 60 Hz, test temperature = 160° C

Time (min)	I_C (mA)	$C \times 10^9$ (F)	$\phi_B/A^2 \times 10^{-2}$ (eV cm^{-4})
0	1.75	5.95	9.50
5	1.85	6.29	8.60
10	1.90	6.46	8.15
15	1.95	6.63	7.75
30	2.00	6.80	7.35
60	2.05	6.97	7.00
100	2.10	7.14	6.65
160	2.12	7.21	6.55

Note: V_{pk} is set at $0.8_{0.5}$, $E_{0.5}$ being the peak a.c. voltage at 0.5 mA cm^{-2} at the onset of nonlinearity.

correspondingly decreased consistent with the diffusion model developed earlier. As stated elsewhere [13], the mechanism of varistor instability can be directly attributed to the lowering of this barrier height.

3.3. Effect of interstitial diffusion on interface (surface) state density

Other supporting evidence for the validity of the diffusion model comes from an examination of the density of the negative charge at the grain-boundary interface before and after stressing a varistor. The model predicts that this density must decrease as the concentration of V'_{Zn} is decreased through charge neutralization of Equation 3. The number of charge sites per cm^2, Q_S, is determined by the interface (surface) state density, N_S, and the Fermi function, $f(E)$. For most situations of interest, Q_S, can be approximated by [3]

$$Q_S = \int N_S(E)\, dE \qquad (14)$$

which assumes a steplike $f(E)$. When expressed in terms of ϕ_B, Equation 14 can be approximated by [3]

$$Q_S = \int N_S(E)\, d(\phi_B) \qquad (15)$$

where N_S is in units of cm^{-2} eV^{-1}. The $N_S(E)$ has been deduced previously [3] from I–V and C–V data together with the measurement of grain size. The same technique was used here to deduce N_S as a function of ϕ_B for the unstressed and stressed varistors. The experiment was conducted as follows. The N_S was first determined for the unstressed device which was then stressed at 100° C for 1225 min at a peak voltage of $0.8E_{0.5}$. The device was then quenched in liquid nitrogen to freeze the defects present at the stressed state. Appropriate measurements were then conducted to determine N_S after bringing the device to room temperature. It was assumed that in spite of some defect diffusion during the experiment, the N_S profile would nevertheless show a difference between the unstressed and stressed states. The results are shown in Fig. 3. It is seen that the peak value of N_S for the stressed state occurs at a lower value of ϕ_B and, as a result, the integrated value (the area under the curve) which is a measure of Q_S is lower for the stressed varistor. Although not shown here, it was found that upon removal of the field, the N_S–ϕ_B curve shifts to the right with time, indicating the onset of recovery on reverse diffusion.

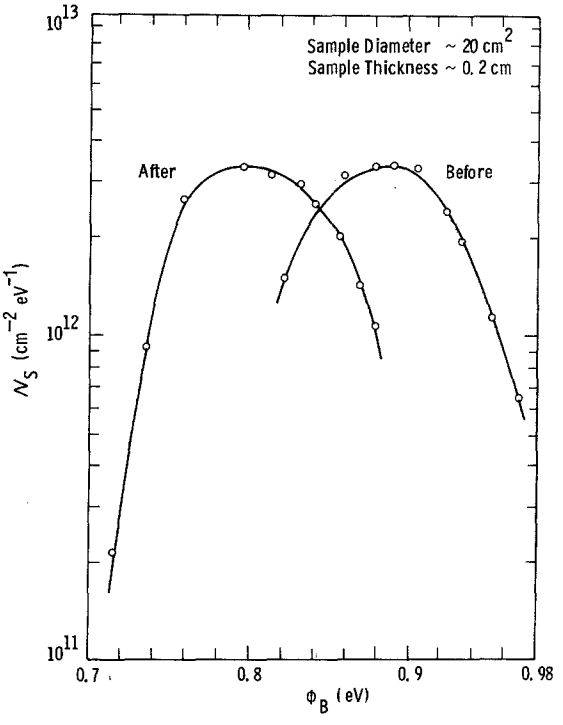

Figure 3 Varistor surface state density before and after stressing at $0.8E_{0.5}$, 100° C, and 1225 min. The varistar was quenched in liquid nitrogen after stressing and then allowed to come to room temperature before measurement.

In this connection, it is interesting to examine the I–V characteristic of the same device before and after stressing. These data have appeared elsewhere [13] but for the sake of completeness we also show it here in Fig. 4. Notice that after stressing, the I–V curve shifts to the right with a decrease in $E_{0.5}$ (barrier voltage) and a simultaneous increase in leakage current. The varistor

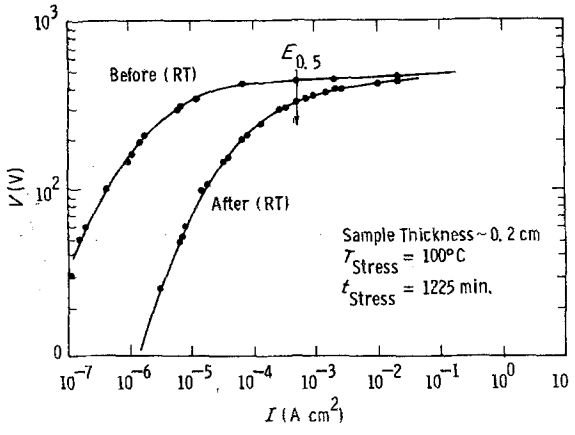

Figure 4 Varistor I–V curve at room temperature before and after stressing at $0.8E_{0.5}$, 100° C, and 1225 min. The varistor was quenched in liquid nitrogen after stressing and then allowd to come to room temperature before measurement. Note that Figs. 3 and 4 are for the same device.

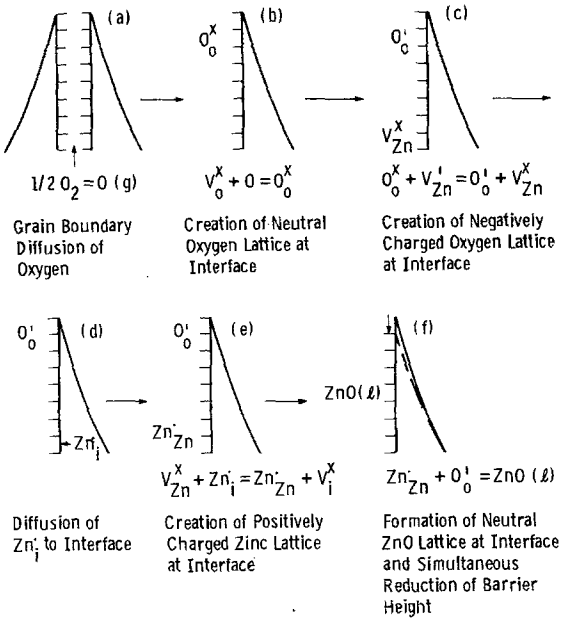

Figure 5 Defect diffusion and chemical interaction at the grain boundary during heat treatment in an oxidized atmosphere. The sequence of quasi-chemical reactions are described in steps from (a) to (f) with explanation in the text.

instability is thus, as previously concluded [13], a direct result of the reduction of barrier voltage and barrier height. As also noted previously [13, 14], this reduction is caused by the metastable component of the barrier.

3.4. Diffusion and chemical interaction during heat treatment

If the zinc interstitials in the depletion layer comprising the metastable component of the barrier are responsible for the instability of the device, it follows that the removal of the metastable component should restore stability. In theory, the metastable component can be removed chemically or thermally. There are numerous evidences in the literature [37–41] to conclude that the stability can be achieved in an otherwise unstable varistor when it is heated for several hours in an oxidizing atmosphere, preferably at 600 to 800° C, although higher and lower temperatures were also used. Such devices when stressed show little or no change in barrier voltage or leakage current with time [13, 14]. In the literature [40, 41], this was presumed to be due to a change in the Bi_2O_3 crystal phase. However, for our model to be consistent, we must conclude that the zinc interstitials are removed by heat treatment. Although this conclusion was inferred previously [13], no defect interaction mechanism was presented. We propose here a sequence of quasichemical reactions for the removal of interstitials from the depletion layer. The mechanism is schematically illustrated in Fig. 5.

We start with the concept that during heat treatment in air, there is a rapid grain-boundary diffusion of oxygen, O(g), through the disorderd layer at the grain boundary (Fig. 5a). The occurrence of a rapid grain-boundary diffusion of oxygen in polycrystalline oxides such as ZnO [42], Al_2O_3 [43] and MgO [44] has already been established in the literature from sintering studies and, therefore, this assumption is not unrealistic. At the grain boundary (the vacancy source), the oxygen atom reacts with the neutral oxygen vacancy, V_O^x, to produce a sub-lattice of neutral oxygen (Fig. 5b):

$$V_O^x + \tfrac{1}{2}O_2(g) = O_O^x \qquad (16)$$

Owing to high electron affinity for oxygen, a charge transfer occurs between the neutral oxygen atom and the negatively charged zinc vacancy at the grain-boundary interface (Fig. 5c)

$$O_O^x + V'_{Zn} = O'_O + V_{Zn}^x \qquad (17)$$

with the formation of a neutral zinc vacancy V_{Zn}^x which is readily annihilated at the grain boundary (vacancy sink), and the creation of a negatively charged oxygen sub-lattice, O'_O, at the

interface. However, since there exists a mass balance between the concentration of V'_{Zn} and $Zn_i^{\cdot}$ due to Frenkel equilibrium,

$$[V'_{Zn}] = [Zn_i^{\cdot}] = K_F^{1/2} \quad (18)$$

a reduction in the concentration of $[V'_{Zn}]$ (Equation 17) requires that the concentration of $[Zn_i^{\cdot}]$ be increased at the interface. A chemical potential gradient of $[Zn_i^{\cdot}]$ is thus established between the depletion layer and the interface, which thus provides the necessary driving force for the diffusion of interstitials from the depletion layer to the grain-boundary interface (Fig. 5d). This diffusion is enhanced at the temperature of 600 to 800° C which is typically employed for heat treatment [37–41]. The interstitial diffusion coefficient at this temperature range is of the order of $\sim 10^{-6}$ to $10^{-7}\,cm^2\,sec^{-1}$ [19].

As the positively charged $Zn_i^{\cdot}$ reaches the grain-boundary interface, it reacts with the neutral V^x_{Zn} (Fig. 5d) which is readily available at the grain boundary (Fig. 5e)

$$V^x_{Zn} + Zn_i^{\cdot} = Zn^{\cdot}_{Zn} + V_i^x \quad (19)$$

with the formation of a positively charged zinc ion at the zinc sub-lattice, $Zn^{\cdot}_{Zn}$, and a neutral interstitial vacancy V_i^x. Note that in Equation 19 the grain boundary acts both as a source (for V^x_{Zn}) and a sink (for V_i^x) for neutral vacancies.

As a result of the Reactions 17 and 19, we now have two oppositely charged ions at the grain-boundary interface: a positively charged zinc ion at the zinc sub-lattice, $Zn^{\cdot}_{Zn}$, and a negatively charged oxygen ion at the oxygen sub-lattice, O'_O. They react to form a neutral zinc oxide lattice, ZnO(l), at the grain-boundary interface (Fig. 5f):

$$Zn^{\cdot}_{Zn} + O'_O = ZnO(l) \quad (20)$$

In the description given above, the ions are considered as singly charged. A similar set of quasi-chemical reactions can be written for the doubly charged ions as well. The net result in either case is the creation of a thermodynamically *stable ZnO lattice* in the grain boundary at the expense of an *unstable zinc interstitial* in the depletion layer. As long as oxygen is available through grain-boundary diffusion, the Reactions 16 to 20 are likely to proceed until all the interstitials are consumed in the creation of new lattice sites at the grain-boundary interface. This is believed to be the reason why an oxidizing atmosphere [12] has been found to be essential to obtain a stable varistor. Also, the reported heat treatment temperature (600 to 800° C) is consistent with the fact that it is higher than the temperature [32] at which the interstitials were estimated to 'freeze-in" [400 to 500° C]. Finally, since several hours of heat treatment are required to obtain stable varistors, the diffusion of interstitials in the depletion layer is inferred to be the rate controlling step, the diffusion of oxygen is presumed very rapid at the grain boundary.

3.5. Effect of diffusional heat treatment on varistor characteristics

It is well known that the heat treatment makes the device stable [37–41]. We remark here that this stability arises due to the elimination of the metastable component of the barrier by diffusional transport of interstitials from the depletion layer to the grain-boundary interface, followed by the interaction of these interstitials with oxygen atoms to form ZnO lattice sites. If this interpretation is correct, one would then expect various changes in varistor characteristics which are related to grain-boundary properties.

The model predicts that there should be a *permanent* reduction in the barrier height and barrier voltage upon diffusional elimination of interstitials. Since it is relatively easy to measure barrier voltage, $E_{0.5}$, we have routinely monitored $E_{0.5}$ values of innumerable devices before and after heat treatment, and in every case, irrespective of composition and processing, there is a reduction in the value of $E_{0.5}$. This is illustrated in Table II which has been reproduced from a previous investigation [13]. The table indicates that the amount of metastable component is different for different compositions, and cannot be predicted *a priori.*

Since the metastable component is absent in a stable device, it further follows that the barrier voltage will remain constant with time when externally stressed, i.e.

$$E_{0.5}(t) = E_{0.5}(0) \quad (21)$$

in contrast to [13]

$$E_{0.5}(t) = E_{0.5}(0) \exp(-t/\tau)^{1/n} \quad (22)$$

observed for devices containing a metastable component in the barrier. In the above

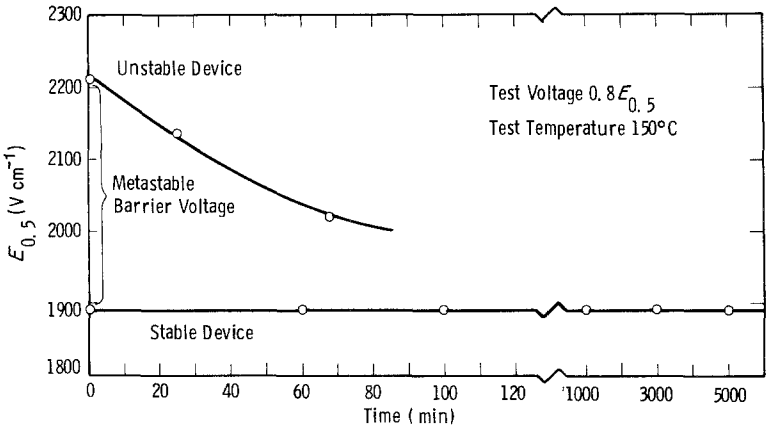

Figure 6 The comparison of the barrier voltage ($E_{0.5}$) with time for a stable and an unstable device. Note the degradation of barrier voltage for the unstable device.

equations, $E_{0.5}(t)$ is the value of $E_{0.5}$ at time t and $E_{0.5}(0)$ at time zero, τ is the time constant, and $n \simeq 2$ to 4. Fig. 6 compares the relative change in $E_{0.5}$ of a stable and an unstable device, respectively. The two devices were initially from the same population with nearly identical values of $E_{0.5}$. The metastable component was then removed from one device by heat treatment whereas the other device remained untreated. The magnitude of the metastable component removed was $\sim 320\,\mathrm{V\,cm^{-1}}$ as shown in the figure. Now, when both devices were stressed, the untreated device showed the time-dependent reduction of $E_{0.5}$ according to Equation 22, whereas the stable device maintained a nearly constant value of $E_{0.5}$ as predicted in Equation 21.

The model predicts further that as a result of the decrease in the concentration of $[V'_{Zn}]$ by heat treatment, there should be a *permanent* reduction in the value of Q_S. This is shown in Fig. 7 where the effective surface state density N_S is plotted against barrier height ϕ_B. It is seen that the effect of heat treatment was to shift the surface state density profile to the left resulting in a permanent decrease in the total number of charge sites, Q_S (the area under the $N_S - \phi_B$ curve) for the heat treated device. It is possible that the donor concentration profile, N_d, will also change upon heat treatment, but it has not been determined in this investigation.

The heat treatment thus causes a significant reduction in the concentration of oppositely charged ions that constitute the barrier at the grain boundary. Since the low curent pre-breakdown region of the I–V curve of the varistor is grain-boundary controlled whereas the high current nonlinear region is grain controlled, it follows that the I–V characteristic of a heat treated varistor will show a significant change in

TABLE II Effect of heat treatment on $E_{0.5}$ and estimate of metastable barrier voltage

Sample number	Composition	Temp. (°C)	Time (h)	$E_{0.5}$ (V cm^{-1}) at room temperature (24° C)			
				Before heating	After heating	Metastable barrier voltage, $\Delta E_{0.5}$	Standard deviation, σ
9982005	A	500	8	2062	1962	100	25
9982011	A	600	8	2050	1918	132	12
9981008	A	700	8	2063	1807	256	1
10022007	B	600	4	2195	2058	137	15
10022010	B	600	8	2179	2042	137	27
9722001	C	600	8	2084	2008	76	26
9732001	D	600	8	1898	1619	279	82
9742001	E	600	8	1913	1648	265	73
9752001	F	600	8	1713	1554	159	5
9762001	G	600	8	1676	1527	151	16
9772001	H	600	8	1759	1601	158	16

Note: Measured values of $E_{0.5}$ are average of three or more samples.

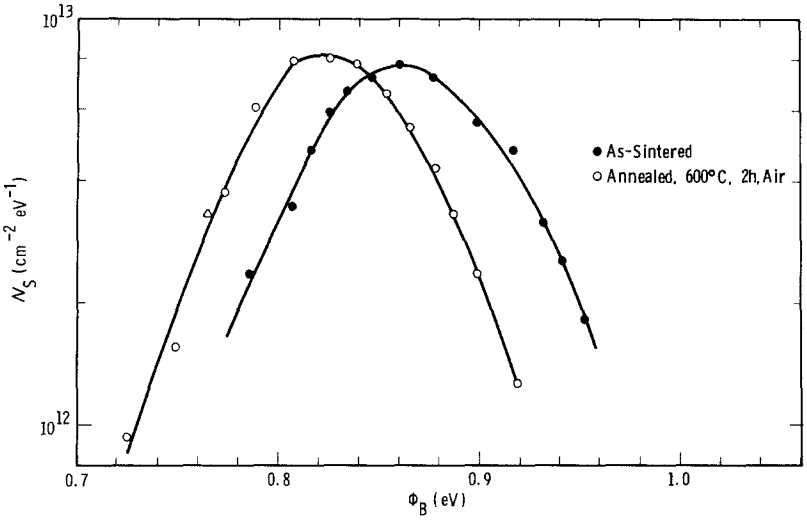

Figure 7 Varistor surface state density before and after annealing the device at 600° C for 2 h in air.

the low current region and little change in the high current region. This is shown for an actual device in Fig. 8 where the I–V data are plotted over a wide range of current using the procedure described previously [11]. The I–V curve of the heat treated device (the dotted curve) is slightly depressed in the low current ($\approx 10^{-5}$A cm^{-2}) region which is consistent with the permanent reduction of $E_{0.5}$ mentioned previously. In contrast, there is a negligible change in I–V characteristic in the high current ($\approx 10^{-1}$A cm^{-2}) region which is consistent with the fact that the grain characteristics are hardly altered by heat treatment. Additionally, the I–V curve is shifted slightly to the right in the low current region.

Since the I–V characteristic in the low current region is depressed by heat treatment due to reduction in $E_{0.5}$, it further follows that there will be a reduction in the nonlinear coefficient, α, which was conveniently measured [45] between two values of currents and voltages by using the equation:

$$\alpha = \frac{\log I_2/I_1}{\log V_2/V_1} \tag{23}$$

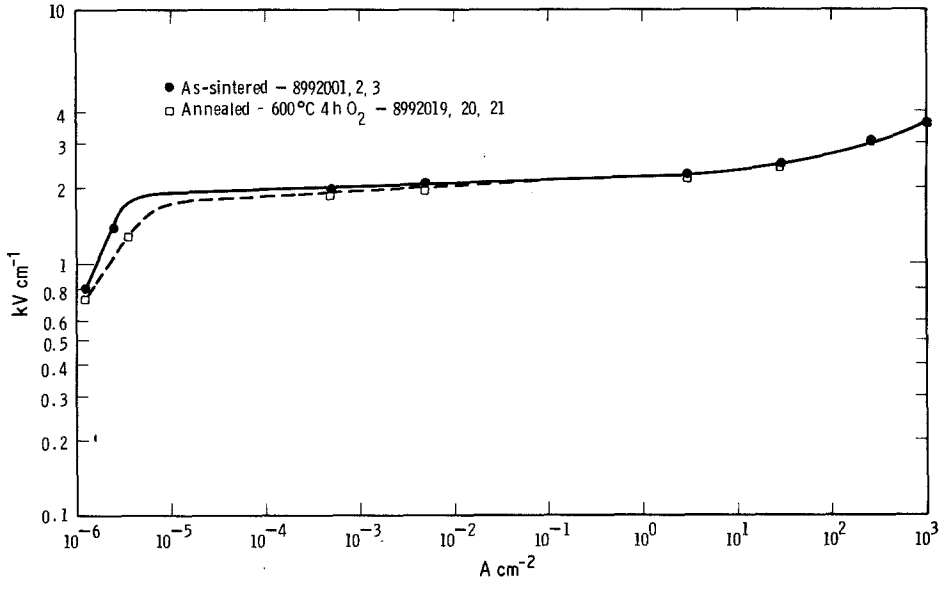

Figure 8 Varistor I–V characteristics from low to high current density (from 10^{-6} to 10^{3}A cm^{-2}) before and after annealing at 600° C for 2 h in air.

TABLE III Effect of heat treatment on the nonlinear coefficient α for ZnO varistors from same population

Heat treatment		α
Temp. (° C)	Time (h)	
As-sintered		35
400	2	33
600	2	30
600	4	27
800	2	29

Note: The value of α was measured between 5×10^{-4} and 250 A cm^{-2}.

The effect of heating on the nonlinear coefficient α between the current ranges of 5×10^{-4} and 250 A cm^{-2} is shown for several devices from the same population in Table III. The α values have been lowered in each case after heat treatment. This lowering of α value due to heat treatment is a common feature of all varistors regardless of composition and fabrication procedure and is consistent with our model for the diffusional annihilation of the interstitials at the grain boundary. Secondly, due to the shift of the I–V curve (Fig. 8) to the right, the heat treated device will produce a greater leakage current than the untreated device for a given applied voltage. This is found to be true for all devices irrespective of the composition and processing. Examples of this increase in leakage current (in terms of the resistive current I_R) as a function of heat treating conditions for two different compositions are shown in Table IV. The I_R was measured at room temperature with an applied voltage set at 80% of $E_{0.5}$, the area being the same for all devices.

In contrast to the above, the durabilities of the device to high current surges and low current, long duration square wave pulses were found to improve upon heat treatment. Both surge and square wave durabilities can be conveniently described by a change in I–V characteristics, or more precisely $E_{0.5}$, upon application of a prescribed number of high current surges of $8 \times 20\,\mu$sec wave shape, and low current, long duration pulses (200 A $\times$ 2500 μsec) of square wave shape. The details of these measurements are beyond the scope of this paper but some results will be briefly mentioned. In one typical test, the varistors were subjected to an increasing current surge of 2.5, 25, 250 and 1000 A cm^{-2}, respectively, and $E_{0.5}$ was monitored before and after surges to determine the percentage of $E_{0.5}$ retained after surges. Similar measurements were also conducted before and after the varistors were subjected to 20 square wave pulses in 20 min. Table V shows a set of electrical data which were obtained before and after heat treatment. The data were averaged for three samples. The first two columns give values of $E_{0.5}$ and α; as discussed previously, both values are reduced by heat treatment. The parameters of interest to present discussion are the data in the last three columns. It is seen that for the heat treated device, there is no degradation of $E_{0.5}$ after high current surges, and very little degradation after square wave pulses. Also shown in the last column is the amount of energy that the varistor will absorb before being destroyed by repeated square wave surges. Here again, the energy absorption capability is improved substantially by heat treatment. A summary of the effect of

TABLE IV Change in room temperature value of I_R due to heat treatment

Sample number	Heat treatment		RT I_R at $0.8E_{0.5}$	
	Temp (° C)	Time (h)	I_R (mA)	% increase upon heat treatment
8962010-3	Control		0.05	–
8962010-12	400	2	0.08	60
8962004-6	400	4	0.07	40
8962013-15	600	2	0.08	60
8962019-21	600	4	0.06	20
8972001-3	Control		0.12	–
8972010-12	400	2	0.13	8
8972004-6	400	4	0.15	25
8972013-15	600	2	0.19	58
8972019-21	600	4	0.24	100

Note: The area of all test devices were same, $\sim 20\,\text{cm}^2$.

TABLE V Electrical characteristics of ZnO varistors before and after heat treatment

Treatment number	$E_{0.5}$ (V cm^{-1})	α	% $E_{0.5}$ retained		J cm^{-3}
			Surge	Square wave	
8 992 001, 2, 3 as-sintered (control)	1990	30	94.0	97.8	63
8 992 019, 20, 21 Heat treated (600° C, 4 h, O_2)	1825	27	100.0	98.1	74

Note: The value of α was measured between 5×10^{-4} and 250 A cm^{-2}.

heat treatment on several varistor parameters is shown in Table VI.

Thus, from a practical standpoint, the major changes obtained from heat treatment are the improved voltage and current stability, the energy absorption capability and the surge and square wave durability of the ZnO varistor combined with lower values for nonlinearity and higher values for leakage current. However, a stable varistor will now permit a higher steady-state operating voltage, thus off-setting the effect of lower values for nonlinearity and thereby attaining improved clamping ratio. By the same token, since the leakage current does not increase with time in a stable varistor, a higher initial leakage current may now be permitted to flow through the stable device as long as this current does not exceed the limiting current at which the device runs away thermally.

Finally, some practical aspects of heat treatment (also known as annealing in the literature) are worth mentioning. The purpose of heat treatment, according to the present model, is to facilitate the diffusional anhilation of all the frozen-in interstitials from the structure, and at the same time, prevent the formation of new interstitials from thermodynamically generated Frenkel disorder. This requirement calls for a delicate balance between generation and annihilation of defects. The annealing temperature must, therefore, be optimized depending on composition and processing. If the temperature is lower than the optimal, the diffusion will not be complete. If the temperature is substantially higher, new interstitials will be formed (Frenkel disorder) in the depletion layers as well as in the bulk. According to our model, these added interstitials in the depletion layers will now cause an increase in barrier height and barrier voltage and a consequent reduction in leakage current (Fig. 2b of the present model). However, since these interstitials are mobile, such varistors are also expected to be unstable with time. Varistors with increased values of $E_{0.5}$ and decreased values of leakage current have indeed been reported in the literature [38, 46] when the annealing temperature was raised to 800 to 1200° C. The above explanation appears to be consistent with these observations. Whether such varistors are stable with time or not, however, has not been reported; our model, however, predicts that they are likely to be unstable.

TABLE VI Effect of heat treatment on several parameters

Parameter	Unstable (unheated)	Stable (heated)
1. I–V curve	Permanent change on heating	
2. $E_{0.5}$(RT)	$E_{0.5}$(RT) >	$E_{0.5}$(RT)
3. I_R(RT)	I_R(RT) <	I_R(RT)
4. $E_{0.5}$ against t	$E_{0.5}(t) = E_{0.5}(0)e^{(-t/\tau)^n}$ Recovery upon removing voltage and cycling	$E_{0.5}(t) = E_{0.5}(0)$ None
5. I_R against t	$I_R(t) = I_R(0) + kt^n$ Recovery upon removing voltage and cycling	$I_R(t) = I_R(0)$ None
6. Q_s(RT) $= \int N_s(E)\mathrm{d}(\phi_B)$	Q_s >	Q_s
7. $\alpha = \dfrac{\log I_2/I_1}{\log V_2/V_1}$	α >	α
8. Surge stability	–	Increase
9. Square wave stability	–	Increase
10. Energy absorption	–	Increase

4. Summary

An atomic defect model for the grain-boundary barrier of a ZnO varistor has been presented in this paper to explain instability/stability of the varistor under an externally applied electric field. The key element of this model is the zinc interstitials which are present in the depletion layer as excess zinc, arising from the non-stoichiometric nature of ZnO and are formed during the fabrication of the varistor. The presence of these interstitials in the barrier gives rise to a metastable component in an otherwise stable barrier which results in an instability of the varistor. The instability arises due to field-assisted diffusion of the interstitials in the depletion layer. Quasi-chemical reactions are derived to show how field-assisted diffusional transport of interstitials can cause a reduction in the barrier voltage, barrier height and interface state density with time. The process is reversible with field removal. It is generally recognized that heat treatment in an oxidizing atmosphere makes the varistor stable. Again a series of quasi-chemical reactions are derived to show how the interstitials can be permanently removed from the depletion layer through diffusion and subsequent chemical reaction with oxygen at the grain boundary. The key factor in restoring the stability by heat treatment is the formation of a stable ZnO lattice at the grain boundary at the expense of the zinc interstitials in the depletion layer. Various experimental results are presented to indirectly validate the above model. Notable changes were observed in barrier voltage, barrier height, interface state density, nonlinearity and current–voltage characteristics. The changes in these parameters are consistent with the prediction of the model.

Acknowledgement

The authors wish to thank P. L. Hower for the experimental data on surface density of ZnO varistors.

References

1. M. MATSUOKA, *Jpn. J. Appl. Phys.* **10** (1971) 736.
2. L. M. LEVINSON and H. R. PHILIPP, *J. Appl. Phys.* **46** (1975) 1332.
3. P. L. HOWER and T. K. GUPTA, *ibid.* **50** (1979) 4847.
4. G. E. PIKE, in "Grain Boundaries in Semiconductors", edited by H. J. Leamy, G. E. Pike and C. H. Seager (Elsevier, New York, 1982) pp. 369–79.
5. K. EDA and A. IGA, *Jpn. J. Appl. Phys.* **18** (1979) 997.
6. K. EDA, A. IGA and M. MATSUOKA, *J. Appl. Phys.* **51** (1980) 2678.
7. C. G. SHIRLEY and W. M. PAULSON, *ibid.* **50** (1979) 5782.
8. K. SATO, Y. TAKADA, H. MAEKAWA, M. OTOTAKE and S. TOMINAGA, *Jpn. J. Appl. Phys.* **19** (1980) 909.
9. S. TOMINAGA, Y. SHIBUY, Y. FUJIWARA, M. IMATAKI and T. NITTA, "Stability and Long Term Degradation of Metal Oxide Surge Arrester", presented at 1979 IEEE Summer Power Meeting, Vancouver, BC, IEEE Transactions on Power Apparatus and Systems, Vol. PAS-99, No. 4, 1980) pp. 1548–1556.
10. W. MOLDENHAUER, K. H. BÄTHER, W. BRÜCKNER, D. HIZN and D. BÜHLING, *Phys. Status Solidi (a)* **67** (1981) 533.
11. T. K. GUPTA, W. G. CARLSON and P. L. HOWER, *J. Appl. Phys.* **52** (1981) 4104.
12. K. TAKAHASHI, T. MIYOSHI, K. MEADA, T. YAMAZAKI and S. OHWADA, in "Grain Boundaries in Semiconductors", edited by J. H. Leamy, G. E. Pike and C. H. Seager (Elsevier, New York, 1982) pp. 399-404.
13. T. K. GUPTA and W. G. CARLSON, *J. Appl. Phys.* **53** (1982) 7401.
14. T. K. GUPTA, W. G. CARLSON and B. O. HALL, in "Grain Boundaries in Semiconductors", edited by J. H. Leamy, G. E. Pike and C. H. Seager (Elsevier, New York, 1982) pp. 393-98.
15. M. H. SUKKAR and H. L. TULLER, in "Grain Boundaries and Interfaces in Ceramics", Advances in Ceramics, Vol. 7, edited by M. F. Yan and A. H. Heuer, (American Ceramic Society, Columbus, Ohio, 1983) pp. 71–90.
16. K. LEHOVEC, *J. Chem. Phys.* **21** (1953) 1123.
17. W. D. KINGERY, *J. Amer. Ceram. Soc.* **57** (1974) 1.
18. G. HEILAND and E. MOLLWO, in "Solid State Physics", Vol. 8, edited by F. Seitz and D. Turnbull (Academic Press, New York, 1959) p. 191.
19. D. G. THOMAS, *J. Phys. Chem. Solids* **3** (1957) 229.
20. D. W. HESS, in "Surface and Interfaces in Ceramic and Ceramic-Metal Systems", Materials Science Research, Vol. 14, edited by J. Pask and A. Evans (Plenum Press, New York, 1981) pp. 335–51.
21. W. D. KINGERY, H. K. BOWEN and D. R. UHLMANN, in "Introduction to Ceramics", (Wiley, New York, 1976) Chs. 4 and 6.
22. R. LINDER, *Acta Chem. Scand.* **6** (1952) 457.
23. W. J. MOORE and E. L. WILLIAMS, *Discus. Far. Soc.* **28** (1959) 86.
24. K. SATO and Y. TAKADA, *J. Appl. Phys.* **53** (1982) 8819.
25. R. W. RICE, in "Materials Science Research", Vol. 3, edited by W. W. Kriegel and H. Palmour III (Plenum Press, New York, 1966) p. 387.
26. D. R. CLARKE, *J. Appl. Phys.* **49** (1978) 2407.
27. M. F. YAN, R. M. CANNON, H. K. BOWEN

and R. L. COBLE, *J. Amer. Ceram. Soc.* **60** (1977) 120.
28. J. SHI CHOI and C. H. YO, *J. Phys. Chem. Solids* **37** (1976) 1149.
29. H. J. ALLSOPP and J. P. ROBERTS, *Trans. Faraday Soc.* **55** (1959) 1386.
30. N. DUPONT-PAVLOVSKY, F. CARALP, P. DELHAES and J. AMIELL, *Phys. Status Solidi (a)* **35** (1976) 615.
31. W. ALBERS, C. HASS and H. J. VINK, *Phillips Res. Repts.* **18** (1963) 372.
32. T. K. GUPTA, and W. G. CARLSON, in "Grain Boundaries and Interfaces in Ceramics", Advances in Ceramics, Vol. 7, edited by M. F. Yan and A. H. Heuer (American Ceramic Society, Columbus, Ohio, 1983) pp. 30-40.
32. R. EINZINGE, in "Grain Boundaries in Semiconductors", edited by H. J. Leamy, G. E. Pike and C. H. Seager (Elsevier, New York, 1982) pp. 343-55.
34. Y. M. CHIANG, W. D. KINGERY and L. M. LEVINSON, *J. Appl. Phys.* **53** (1982) 1765.
35. P. EMTAGE, *ibid.* **48** (1977) 4372.
36. W. G. MORRIS, *J. Vac. Sci. Technol.* **13** (1976) 92.
37. H. R. PHILIPP and L. LEVINSON, *J. Appl. Phys.* **50** (1979) 383.
38. J. E. MAY, US Patent 4 165 351, 21 August (1979).
39. J. S. KRESGE, US Patent 4 046 847, 6 September (1977).
40. A. IGA, M. MATSUOKA and T. MASUYAMA, *Jpn. J. Appl. Phys.* **15** (1976) 1161.
41. *Idem, ibid.* **15** (1976) 1847.
42. T. K. GUPTA and R. L. COBLE, *J. Amer. Ceram. Soc.* **51** (1968) 521.
43. R. L. COBLE, *J. Appl. Phys.* **32** (1961) 793.
44. T. K. GUPTA, *J. Mater. Sci.* **6** (1971) 25.
45. W. G. CARLSON and T. K. GUPTA, *J. Appl. Phys.* **53** (1982) 5746.
46. J. E. MAY, US Patent 4 042 535, 16 August (1977).

Received 25 July
and accepted 31 July 1984