Morris Shatzkes
Moshe Av-Ron

Statistics of Breakdown

A derivation is given for the statistics of breakdown, including the effect of defects. The particular case where breakdowns are random in time and defects are randomly distributed is considered in detail. Results are obtained that are applicable to any kind of breakdown test. It is shown how relationships between the breakdown distributions for various breakdown tests and for different conditions in a particular test are generated. These relations are given for life tests (constant field) and for ramp tests (field proportional to time). It is found that an ordering of defect types according to their susceptibility to breakdown is not unique, but generally depends on the test conditions. This implies that the fields and temperatures used in screening procedures must be chosen with care.

Introduction

Dielectric breakdown is an important source of failure in electronic systems. Although the phenomenon has been studied extensively [1, 2], a systematic analysis of the statistics of breakdown has not been previously set forth. However, a statistical model of dielectric breakdown is required for precise reliability projections and assessments of the dielectric quality from data obtained in breakdown tests. Such a model is presented in this paper.

A statistical breakdown model restricted to defect-free devices was derived by Solomon, Klein, and Albert [3]. They assumed breakdowns to be random, independent events. With this assumption, and in the absence of defects, the breakdown statistics are determined by the breakdown probability, which in turn depends on the test conditions and the time on test. We derived a model incorporating the effect of defects for the case of one defect type [4]; *i.e.*, for the case in which the probability of breakdown is the same at each defect. In addition to specification of the probability of breakdown at the defect and in defect-free regions, a knowledge of the defect distribution is required in the determination of the breakdown statistics. In the present paper, we treat the more general case of many defect types, the distinguishing characteristic being the breakdown probability at the defect. For each defect type, the probability of breakdown at that defect is required. The breakdown statistics depend on these probabilities and on the distribution of the defects.

Initially, the problem is formulated and the results are obtained in terms of subsets of devices having equal probabilities for breakdown. These subsets contain devices having identical numbers of defects of each type. The statistical independence of breakdowns at defects and in defect-free regions is then utilized, together with the assumption of independent Poisson distributions, to obtain results in terms of the defect types. Also, by assuming a particular form for the probabilities of breakdown at defects and in defect-free regions as functions of time and field, explicit results are obtained for the mean number of devices broken down and the standard deviation of this number. This is done for the commonly used *ramp* and *life* breakdown tests. Relations among various tests follow directly from these results. Rather than using a single transformation equation to relate the breakdown distribution of different tests, the different defect types are transformed separately.

Statistical model

In a breakdown experiment, or in actual use, an electric field (which may be variable) is applied and the time to breakdown is the quantity of interest. This time generally depends on the field and the temperature, but may fluctuate even at constant field and temperature. In addition, the time to breakdown depends on the structure of the device. Breakdowns may occur with higher probabilities at defects—regions in which deviations of any kind exist that may affect breakdown. For instance, these deviations may be regions of thinner dielectric, impurities, or interface barrier modulations. Given a set of N devices, subsets can be formed where the *a priori* breakdown probability for each device of a subset is the same; *i.e.*, they have identical defects. Let N_α be the number of devices in the αth subset and $p_\alpha(t)$ be the probability that a device in this subset will break down within time t, and let there be κ subsets. The probability that n devices will break down within t, for the given N_α, is

$$p_n(t, \vec{N}) = \sum_{\{\Sigma n_\alpha = n\}} \prod_{\alpha=1}^{\kappa} \begin{bmatrix} N_\alpha \\ n_\alpha \end{bmatrix} p_\alpha^{n_\alpha}(t)[1 - p_\alpha(t)]^{N_\alpha - n_\alpha}. \quad (1)$$

Here, the notation $\vec{N}$ is used to indicate the dependence of p_n on the N_α ($\alpha = 1, \cdots, \kappa$) and the sum is over the non-negative integer solutions of $\Sigma n_\alpha = n$.

If many sets of N devices are extracted from the same population, $p_n(t, \vec{N})$ will vary if the N_α are not all identical. Let $\rho(\vec{N})$ be the probability that N_α out of N devices will be of the subset α for $\alpha = 1, \cdots, \kappa$. Then the probability that n devices will break down within t is given by

$$p(t, n) = \sum_{\{\Sigma N_\alpha = N\}} \rho(\vec{N}) p_n(t, \vec{N}). \quad (2)$$

The expectation value of n, $\langle n \rangle$, and the standard deviation σ are given by

$$\langle n \rangle = \sum_\alpha \langle N_\alpha \rangle p_\alpha(t), \quad (3)$$

$$\sigma^2 = \sum_\alpha \langle N_\alpha \rangle p_\alpha(t)[1 - p_\alpha(t)] + \sum_{\alpha,\beta} p_\alpha p_\beta (\langle N_\alpha N_\beta \rangle - \langle N_\alpha \rangle \langle N_\beta \rangle), \quad (4)$$

where

$$\langle N_\alpha \rangle = \Sigma N_\alpha \rho(\vec{N}), \quad (5)$$

$$\langle N_\alpha N_\beta \rangle = \Sigma N_\alpha N_\beta \rho(\vec{N}), \quad (6)$$

and the sums in Eqs. (5) and (6) are over the non-negative integer solutions of $\Sigma N_\alpha = N$. These results are easily interpreted. The results [Eq. (3)] for $\langle n \rangle$ and the first sum on the right of Eq. (4) follow because for each α, n_α is a binomial distribution and $\langle N_\alpha \rangle$ is the expected number of devices in the αth subset. The second sum on the right of Eq. (4) is due to deviations in the N_α.

If the probabilities of choosing devices of different subsets are independent,

$$\rho(\vec{N}) = \frac{N!}{N_1! \cdots N_\kappa!} P_1^{N_1} \cdots P_\kappa^{N_\kappa} \qquad \Sigma N_\alpha = N, \quad (7)$$

where P_α is the probability of choosing a device of the αth subset. In this case, we have

$$\langle N_\alpha \rangle = NP_\alpha, \quad (8)$$

$$\langle N_\alpha N_\beta \rangle = \begin{cases} N(N-1)P_\alpha P_\beta & \alpha \neq \beta, \\ NP_\alpha + N(N-1)P_\alpha^2 & \alpha = \beta, \end{cases} \quad (9)$$

and therefore,

$$\langle n \rangle = N \sum_\alpha P_\alpha p_\alpha(t), \quad (10)$$

$$\sigma^2 = N\left[\sum_\alpha P_\alpha p_\alpha(t)\right]\left[1 - \sum_\alpha P_\alpha p_\alpha(t)\right] = \langle n \rangle (1 - \langle n \rangle / N). \quad (11)$$

Independent Poisson distributions of defects

The device subsets are characterized by the breakdown probability and this implies that devices belonging to the same subset have equal numbers of defects of the various types. Let there be r different types of defects. Then, the αth subset can be characterized by the r numbers, $n_{\alpha,1}, \cdots, n_{\alpha,r}$, which specify the number of defects of each type for that subset. Now it is assumed that breakdowns at the defect and in the defect-free region are independent events. The probability $p_\alpha(t)$ of a breakdown within time t for devices in the subset α can be formulated as

$$p_\alpha(t) = 1 - q_0 q_1^{n_{\alpha,1}} \cdots q_r^{n_{\alpha,r}}, \quad (12)$$

where q_i is the probability that a breakdown does not occur at a defect of type i within t for $i = 1, \cdots, r$ and q_0 is the corresponding probability for the defect-free region.

If the defects of each type are distributed independently with Poisson distributions,

$$P_\alpha = \prod_{i=1}^{r} \frac{\lambda_i^{n_{\alpha,i}} \exp(-\lambda_i)}{n_{\alpha,i}!}, \quad (13)$$

where λ_i is the mean number of defects of type i per sample. In this case, there are an infinite number of subsets corresponding to different $(n_{\alpha,1}, \cdots, n_{\alpha,r})$, where each $n_{\alpha,1}$ ranges from zero to infinity. The expectation value for the fraction of samples breaking down within t is obtained for this case by using Eqs. (12) and (13) in Eq. (10). The result is

$$f(t) = 1 - q_0 \exp\left[-\sum_i \lambda_i(1 - q_i)\right]. \quad (14)$$

The standard deviation is obtained directly from Eq. (11):

$$(\sigma/N)^2 = f(1 - f)/N.$$

The interpretation of Eq. (14) is straightforward. In a breakdown test, a defect is only counted if a breakdown occurs at the defect. Therefore, the effective mean number of defects of type i is the product of the mean number λ_i and $(1 - q_i)$, the probability that a breakdown occurs at a defect of type i. The factor q_0 is the fraction of samples not having any defects and not broken down at time t. Thus, $(1 - f)/q_0$ is the usual expression for the fraction of samples not having "defects"; *i.e.*, if defects are considered to be only those which would break down at time t on the average, for which the effective mean numbers are $\lambda_i(1 - q_i)$.

As a function of time, the shape of the fraction broken down, $f(t)$, is apparent from Eq. (14). The q_i, being probabilities that a breakdown did not occur up to time t, decrease monotonically from one to zero as t increases from zero to infinity. This will always be the case; the details of this excursion will, however, depend on the relationship between the q_i and the time, field, and temperature. Suppose that an interval exists such that, during this interval, all q_i with $i < j$ and $i \neq 0$ can be considered to vanish, and all q_i with $i > j$ and $i = 0$ are approximately unity. Also in this time interval, q_j goes from zero to one. Then $f(t)$ increases from

$$\left[1 - \exp\left(-\sum_{i=1}^{j-1} \lambda_i\right)\right] \quad \text{to} \quad \left[1 - \exp\left(-\sum_{i=1}^{j} \lambda_i\right)\right];$$

i.e., a step in $f(t)$ will occur of size

$$\left[\exp\left(-\sum_{i=1}^{j-1} \lambda_i\right)\right][1 - \exp(-\lambda_j)],$$

which is just the fraction of devices which do not have the defects of types 1 to $(j - 1)$ but do have the defect of type j. This result is physically intuitive and could have been anticipated without calculation. In general, breakdowns at different defects will overlap in time and $f(t)$ will deviate from a stepped structure. Nevertheless, the physical picture will still prevail; those samples having defects more susceptible to breakdown will most likely break down first and these are defects for which the excursion of q_i from 1 to 0 will occur at earlier times. A screening procedure, the application of a field for some preset time, is thus possible. This would eliminate devices containing those defects for which $q_i \rightarrow 0$ in that time.

Breakdowns random in time

To proceed further, a form must be chosen for the breakdown probabilities in the various regions. Klein [5-7] has advocated that intrinsic breakdown mechanisms (defect-free) are inherently random, and experimental evidence exists supporting this view [3, 6, 8]. Evidence also exists for the random nature of breakdowns at defects [4]. Assuming that breakdowns are random events, the q_i are given by

$$q_i = \exp\left[-\int_0^t (1/\tau_i)dt\right]. \tag{15}$$

Solomon, Klein, and Albert [3] prescribe the field dependence of τ_0, the mean time to breakdown for the defect-free region, to be

$$\tau_0 = \tau_{00} \exp(-F/F_{00}), \tag{16}$$

where F is the electric field, indicating evidence for such an exponential field dependence. Adopting the point of view that the processes involved in breakdown at defects are the same as for defect-free regions, with possibly altered rates, the form for τ_i is

$$\tau_i = \tau_{i0} \exp(-F/F_{i0}) \tag{17}$$

for each of the defect types. In general, we would expect the τ_{i0} and F_{i0} to be functions of temperature. The temperature dependence arises from that of the processes involved in breakdown, which for electronic breakdown are charge-carrier injection and motion, impact ionization, and recombination.

The expectation value for the fraction of devices broken down within time t, $f(t)$, is now specified by using Eqs. (15)-(17) in Eq. (14). It is convenient to recast Eq. (14) as

$$f(t) = 1 - q_0 y_1 y_2 \cdots y_r,$$
$$y_i = \exp[-\lambda_i(1 - q_i)]. \tag{18}$$

There are $(3r + 2)$ parameters which enter $f(t)$, three for each of the defects, λ_i, F_{i0}, and τ_{i0}, and two for the defect-free region, F_{00} and τ_{00}. We discuss now the role of these parameters for several breakdown tests. Independent of the nature of the test, $f(t)$ is given by Eq. (18) and if the parameters are known, the breakdown distribution can be predicted.

Life test

In a life test, the field and temperature are constant, and therefore

$$q_0 = \exp(-t/\tau_0), \qquad \tau_0 = \tau_{00}\exp(-F_L/F_{00}),$$
$$q_i = \exp(-t/\tau_i), \qquad \tau_i = \tau_{i0}\exp(-F_L/F_{i0}),$$
$$y_i = \exp\{-\lambda_i[1 - \exp(-t/\tau_i)]\}, \qquad i = 1, 2, \cdots, r, \tag{19}$$

where the life test field is designated as F_L. Since the range of times t covered in a life test is often many orders of magnitude, it is more convenient to use $s \equiv \ln t$ as the variable. Another advantage is that in terms of s the analysis parallels that for the ramp test to be discussed next.

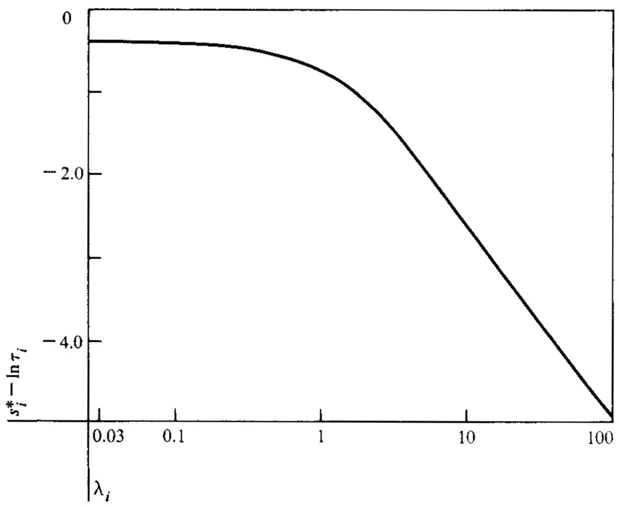

Figure 1 $(s_i^* - \ln \tau_i)$ and $[(F_i^*/F_{i0}) - \ln (R\tau_{i0}/F_{i0})]$ as functions of λ_i [Eqs. (22) and (31)]. Note the sensitivity to λ_i for $\lambda_i \to 1$.

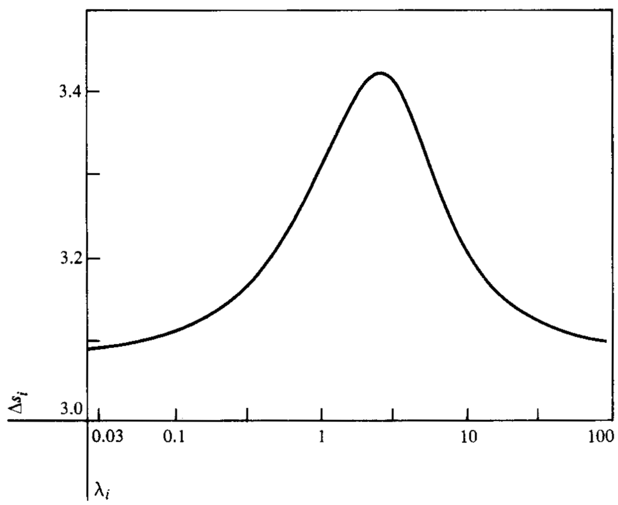

Figure 2 Δs_i and F_i/F_{i0} [Eqs. (23) and (31)] as functions of λ_i, indicating insensitivity to λ_i over the entire range as inferred from the scale of the ordinate.

Consider the case of two life tests, one at a field F_{L1} and one at F_{L2}. Equal values of y_i are obtained in the two tests at values of s_1 and s_2 such that the value of q_i is the same in the two tests. This occurs when

$$\frac{\exp s_1}{\tau_{i1}} = \frac{\exp s_2}{\tau_{i2}}, \tag{20}$$

where

$$\tau_{i1} = \tau_{i0} \exp(-F_{L1}/F_{i0}) \text{ and } \tau_{i2} = \tau_{i0} \exp(-F_{L2}/F_{i0}).$$

Thus, equal values of y_i are obtained at s_1 and s_2 for the two tests if

$$s_2 - s_1 = -(F_{L2} - F_{L1})/F_{i0}, \tag{21}$$

and this holds for $i = 0, 1, \cdots, r$. It is seen that as a function of s, the shape of y_i is invariant with respect to a change in field. The entire curve $y_i(s)$ is translated parallel to the s axis by an amount given by Eq. (21) and the magnitude of the translation depends on F_{i0}, thus being different for the different defect types. A pure translation of the entire $f(s)$ curve results only if all the F_{i0} are the same, $i = 0, 1, \cdots, r$. This is only possible if all the τ_{i0} differ. In fact, since the defect types are distinguished by the two parameters τ_{i0} and F_{i0}, at least one of these must be different for any pair of defects and for τ_{00} and F_{00}.

As a consequence of Eq. (21), defects cannot generally be ordered according to their susceptibility to breakdown in a life test, since the ordering depends on the field. This stems from the dependence of the τ_i on the field. Unless 1) all the F_{i0} are the same, and thus the types are distinguished only by the τ_{i0}, or 2) all the τ_{i0} are the same, and thus the F_{i0} are all different, an ordering of the τ_i according to magnitude will not be preserved as the field changes.

To characterize y_i further, it is desirable to specify its location on the $s \equiv \ln t$ scale and give a measure of its width in s. Its location can be specified by s_i^*, the value of s where $y_i = [1 + \exp(-\lambda_i)]/2$. From Eqs. (19), s_i^* is given by

$$s_i^* = \ln \tau_i + \ln \left[\left(-\ln \{(\lambda_i)^{-1} \ln [(\exp \lambda_i + 1)/2]\}\right)\right]. \tag{22}$$

The second term on the right of Eq. (22) is a slowly varying function of λ_i, which is plotted in Fig. 1. Except for large values of λ ($\lambda > 1$), the variation of s_i^* with λ_i can be neglected, and s_i^* is thus mainly determined by τ_i. A measure of the width of y_i, Δs_i, can be provided by the difference in the values of s for which $1 - y_i$ is $0.1[1 - \exp(-\lambda_i)]$ and $0.9[1 - \exp(-\lambda_i)]$, *i.e.*, the values of s where y_i has decreased by 0.1 and 0.9 of the difference between its maximum and minimum values. Thus,

$$\Delta s_i = \ln \left\{\frac{\ln [(\lambda_i)^{-1} \ln (0.9 + 0.1 \exp \lambda_i)]}{\ln [(\lambda_i)^{-1} \ln (0.1 + 0.9 \exp \lambda_i)]}\right\}. \tag{23}$$

This is a slowly varying function of λ_i and is plotted in Fig. 2, for which $\lambda_i < 1$ is about 3.1. Similarly, we can define s_0^* as the value of s for which $q_0 = 0.5$, and Δs_0 as the difference in the values of s for which $q_0 = 0.1$ and 0.9. The results are

$$s_0^* = \ln \tau_0 + \ln (\ln 2),$$

$$\Delta s_0 = \ln \left(\frac{\ln 0.1}{\ln 0.9}\right) = 3.084. \tag{24}$$

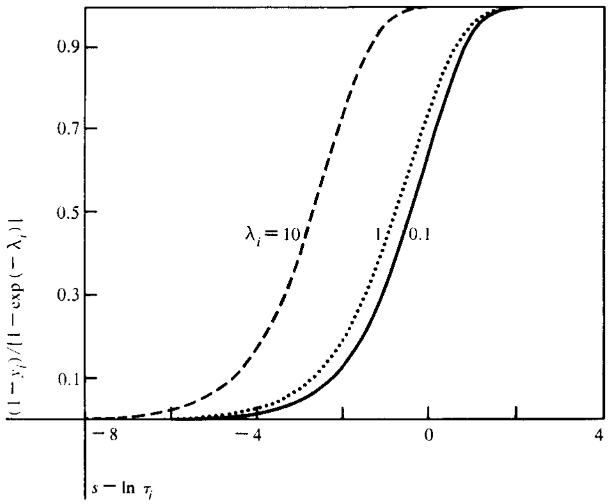

Figure 3 Normalized fraction of breakdown for a single defect type, $(1 - y_i)/[1 - \exp(-\lambda_i)]$, as a function of $(s - \ln \tau_i)$ for $\lambda_i =$ 0.1, 1, and 10.

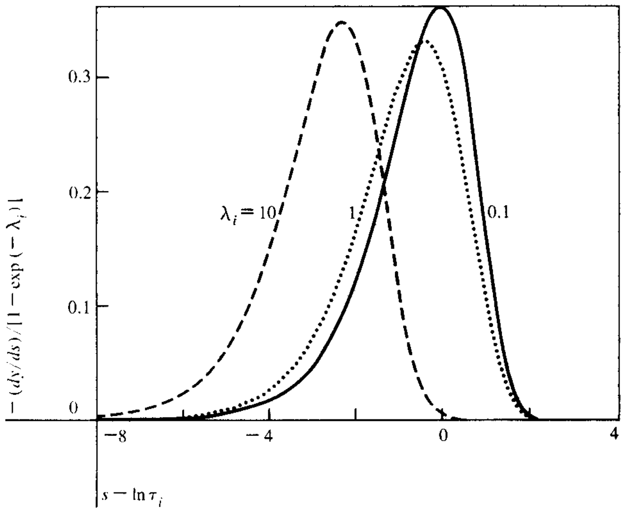

Figure 4 The rate of breakdown (in $s = \ln t$) for a single defect type, $-(dy/ds)/[1 - \exp(-\lambda_i)]$ as a function of $(s - \ln \tau_i)$ for λ_i = 0.1, 1, and 10 (curves are as in Fig. 3).

These equations are essentially the same as Eqs. (22) and (23) for small λ_i.

The parameters entering y_i, λ_i, F_{i0}, and τ_{i0} can be determined in the following way: λ_i from $\lim y_i = \exp(-\lambda_i)$ as $s \to \infty$, F_{i0} from the translation of y_i given by Eq. (21) for life tests at different fields, and τ_{i0} from the measured s_i^* and the already determined values of λ_i and s_i^*. Thus, life tests are required for at least two different fields. Actually, the degree of complexity in extracting the parameters depends on the degree of overlap in s of the y_i. Since different y_i shift by different amounts if the F_{i0} are different, life tests at different fields can resolve the defect types.

In Fig. 3, we plot $(1 - y_i)/[1 - \exp(-\lambda_i)]$, the normalized fraction of breakdown for a single defect type, as a function of $(s - \ln \tau)$ for $\lambda_i = 0.1$, 1, and 10. The insensitivity to λ_i for $\lambda_i \lesssim 1$ is apparent, indicating that λ_i essentially only affects the normalization. The shape of y_i is also not a function of τ_i since y_i can be written as

$$y_i = \exp\left(-\lambda_i\{1 - \exp[-\exp(s - \ln \tau_i)]\}\right). \tag{25}$$

For any life test, breakdowns due to a single defect type or in the defect-free region will occur over an essentially constant interval in $s = \ln t$. Using our measure of width to specify this interval, *i.e.*, Δs_i of Eq. (23) and Δs_0 of Eq. (24), the result is that this interval is about 3.1. This means that a single defect type will cause breakdowns for about a factor of ten in time. The same is true for defect-free breakdowns. However, breakdowns are often observed to cover many factors of ten in time. This can be considered as evidence for the existence of multiple defect types.

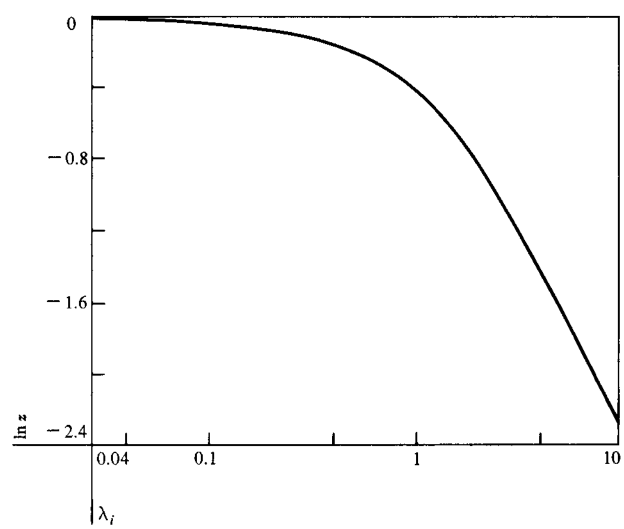

Figure 5 $\ln z$ as a function of λ_i; z is defined by Eq. (26).

The y_i are the expectation values for the fraction of devices not broken down at time t if breakdowns have only occurred at the ith defect type. The rate of device breakdowns due to defects of type i is given by the negative of the time derivative of y_i. In Fig. 4, we plot

$$\frac{d}{ds}\left[\frac{1 - y_i}{1 - \exp(-\lambda_i)}\right]$$

as a function of $(s - \ln \tau_i)$ for $\lambda_i = 0.1$, 1, and 10. A skewness is evident in the rate of device breakdown, and the relative insensitivity to λ_i is apparent. The peak occurs at a value of $(s - \ln \tau_i) = \ln z$, where z is the solution of

$$\lambda_i = (1 - z)/z \exp(-z). \tag{26}$$

Figure 5 is a plot of z *vs.* λ_i. It is interesting to compare $\ln z$, the value of $(s - \ln \tau_i)$ at the peak rate, with $(s_i^* - \ln \tau_i)$ as given by Eq. (24), which is the value of $(s - \ln \tau_i)$ when half the devices have broken down, if breakdowns have only occurred at defects of type i. We find that for $\lambda_i \le 1$, the difference $[\ln z - (s_i^* - \ln \tau_i)]$ is about equal to 0.35 or approximately equal to $-\ln(\ln 2)$.

Similarly, we can consider the rate of device breakdowns if breakdowns have occurred only in the defect-free regions:

$$\frac{d}{ds}(1 - q_0) = [\exp(s - \ln \tau_0)] \exp\{\exp[-(s - \ln \tau_0)]\}. \tag{27}$$

In this case the peak rate occurs at $s = \ln \tau_0$, whereas $(1 - q_0) = 0.5$ occurs at $s_0^* = \ln \tau_0 + \ln(\ln 2)$. The difference between peak rate and s_0^* is exactly $-\ln(\ln 2)$ for defect-free breakdowns.

Ramp test

In a ramp test, the field F varies linearly in time so that $F = Rt$, where R is the ramp rate. The q_i, as given by Eq. (15), are for this case

$$q_i = \exp\{(-F_{i0}/R\tau_{i0})[\exp(F/F_{i0}) - 1]\},$$
$$i = 0, 1, \cdots, r. \tag{28}$$

These q_i are to be substituted into Eq. (18) to obtain y_i and $f(t)$, the expectation value for the fraction broken down up to time t.

The parameter defining a ramp is the ramp rate R, and we consider the effect of changing R from R_1 to R_2. The contribution to breakdown from the defect type will be the same at the two ramp rates for fields F_1 and F_2, such that y_i are the same. Since the q_i will be equal for equal values of $[\exp(F/F_{i0}) - 1]/R$, the relation between F_1 and F_2 is

$$[\exp(F_1/F_{i0}) - 1]/R_1 = [\exp(F_2/F_{i0}) - 1]/R_2, \tag{29}$$

or, since usually $\exp(F/F_{i0}) >> 1$,

$$F_2 - F_1 = F_{i0} \ln(R_2/R_1). \tag{30}$$

Therefore, as was the case for life tests at different fields, y_i in ramp tests at different ramp rates is translated parallel to the field axis with no alteration of its shape. This translation differs for defect types with different F_{i0}, so that an ordering of defect types according to their susceptibility is also not generally possible in ramp tests. This ordering will depend on the ramp rate.

Since, as a function of field, the shape of y_i is invariant with ramp rate, we can characterize its location by F^*, the value of the field where $(1 - y_i)/[1 - \exp(-\lambda_i)] = 0.5$, as for a life test, and its width ΔF_i by the difference in the fields where $(1 - y_i)/[1 - \exp(-\lambda_i)] = 0.9$ and 0.1. These are then given by

$$F_i^* = F_{i0}[\ln(R\tau_{i0}/F_{i0})] + F_{i0} \ln\left(-\ln\{\lambda_i^{-1} \ln[(\exp \lambda_i + 1)/2]\}\right),$$

$$\Delta F_i = F_{i0} \ln\left\{\frac{\ln[\lambda_i^{-1} \ln(0.9 + 0.1 \exp \lambda_i)]}{\ln[\lambda_i^{-1} \ln(0.1 + 0.9 \exp \lambda_i)]}\right\}, \tag{31}$$

where we approximated $[\exp(F/F_{i0}) - 1] \simeq \exp(F/F_{i0})$. The logarithmic functions of λ_i occurring here and plotted in Figs. 1 and 2 are the same as those in Eqs. (22) and (23) for the life test. In fact, $[(F_i^*/F_{i0}) - \ln(R\tau_{i0}/F_{i0})]$ is the same function of λ_i as $(s_i^* - \ln \tau_i)$ and $\Delta F_i/F_{i0}$ is the same function of λ_i as Δs_i. This is to be expected since with the above approximation, in which F/F_{i0} is replaced with s and $R\tau_{i0}/F_{i0}$ with $\tau_i(F_L)$, the q_i for a ramp test [Eq. (28)] become the q_i for a life test [Eq. (19)]. Thus, there is no fundamental difference between these two types of tests.

In fact, any test gives essentially the same information since it is described by Eq. (17). The only difference between tests is the particular forms of the q_i, which are given by Eq. (15). Between these tests there exist transformations which relate the test variables. The requirement that y_i in one test be equal to that in the other is the condition which gives rise to these relations. This condition was used to obtain Eqs. (21) and (22), which related the times for equal y_i in life tests at different fields and fields for equal y_i in ramp tests at different ramp rates. The conditions are, more explicitly, that for the two tests the integrals in the exponent of Eq. (15) be equal. As another application of this transformation we derive the time t in a life test at F_L and field F in a ramp test at ramp rate R such that y_i in the two tests are equal. The requirement of equal values for the integral in Eq. (15) for the two tests is that

$$(t/\tau_{i0}) \exp(F_L/F_{i0}) = (F_{i0}/R\tau_{i0})[\exp(F/F_{i0}) - 1] \simeq (F_{i0}/R\tau_{i0}) \exp(F/F_{i0})$$

so that the relation is [9]

$$\ln t = \ln(F_{i0}/R) + (F - F_L)/F_{i0} \qquad i = 0, 1, \cdots, r. \tag{32}$$

This relation will be different for those defect types having different F_{i0}.

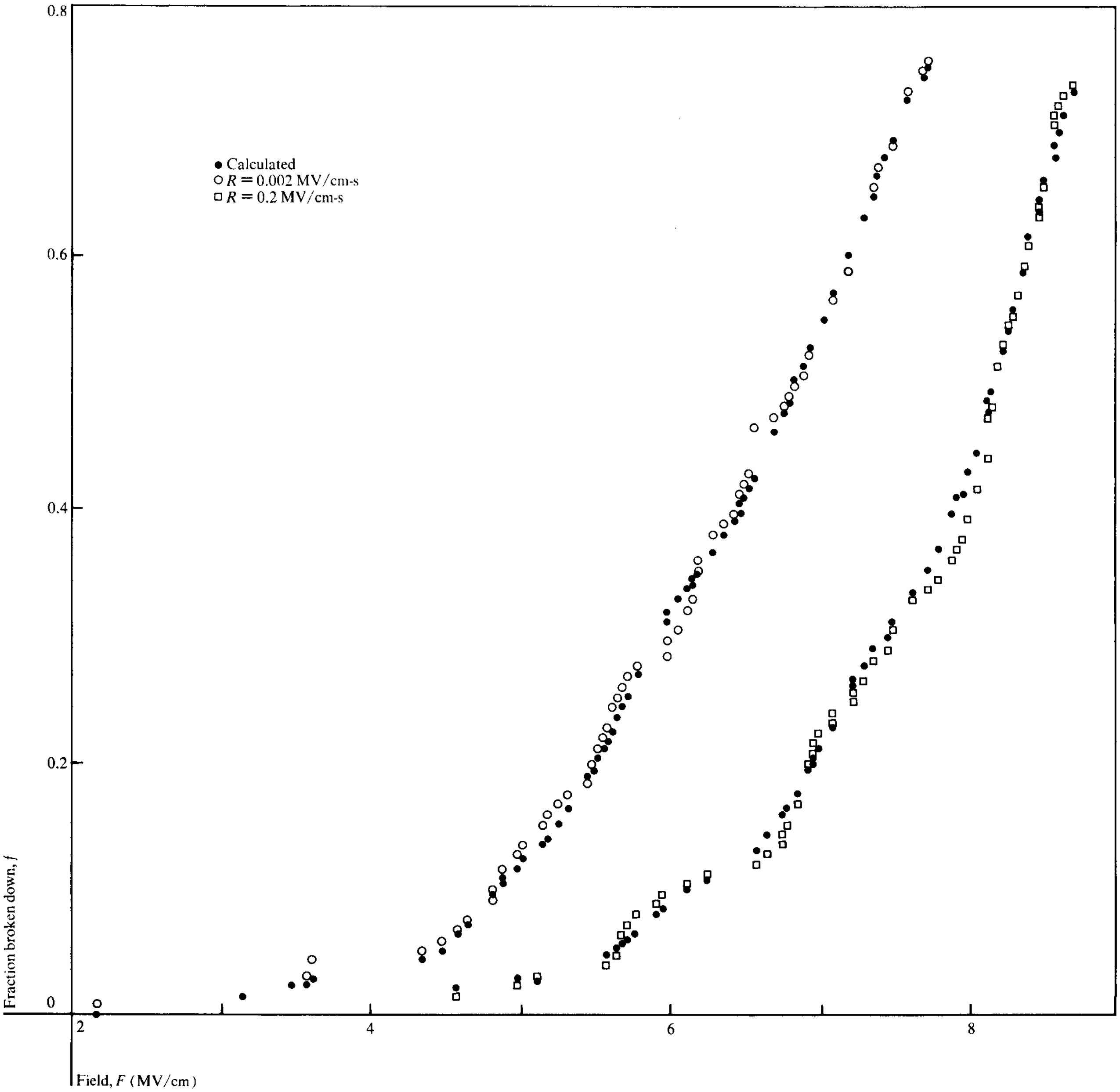

Figure 6 Experimental results and calculation of cumulative fraction of breakdown *vs.* the electric field for Al-SiO_2-Si MOS capacitors at two different ramp rates.

A distinction between ramp test and life test is that while for life tests the width Δs_i is only a weak function of λ_i [Eq. (23)], in a ramp test the width ΔF_i is proportional to F_{i0} [Eq. (31)]. Thus, whereas in a life test device breakdowns caused by any of the defect types occur in about a factor of ten in time, in a ramp test the range of fields for which a defect type will contribute breakdowns will vary with F_{i0}.

Also, as for life tests, the determination of parameters requires at least two tests at different ramp rates. Thus the F_{i0} can be determined, using Eq. (30), from the translations of the $f(F)$ curve caused by the different defect types. The $f(F)$ curve will not generally translate as a whole, since the translations, depending on F_{i0}, will be different for different types. The same is true for $f(s)$ in life tests at different fields. Again, λ_i mainly affects the total excursion of y_i, which is $1 - \exp(-\lambda_i)$, and τ_{i0} can be found from the field F_i^* [see Eq. (31)].

The form of y_i, $(1 - y_i)/[1 - \exp(-\lambda_i)]$, as a function of $[F/F_{i0} + \ln(F_{i0}/R\tau_{i0})]$ is the same as for a life test as a

Table 1 Parameters required to fit data of Fig. 6 according to Eq. (17) for f; the q_i are given by Eq. (29).

	Defect types							
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*
λ_i	0.0253	0.0725	0.1068	0.1278	0.1143	0.2333	0.4902	0.4316
τ_{i0}	1.75×10^7	3.32×10^{10}	6.328×10^8	1.92×10^{11}	1.16×10^{10}	5.235×10^{10}	1.807×10^{15}	3.037×10^{17}
F_{i0}	0.271	0.244	0.37	0.2714	0.3693	0.338	0.242	0.219

function of $s - \ln \tau_i$, which is depicted in Fig. 3. Similarly, if the abscissa of Fig. 4 is interpreted as $F/F_{i0} + \ln (F_{i0}/R\tau_{i0})$, Fig. 5 gives $(d/ds)\{(1 \times y_i)/[1 - \exp(-\lambda_i)]\}$ for the ramp test as a function of this variable.

As an example, we present in Fig. 6 experimental data obtained using MOS capacitors (Al-SiO_2-Si, p-type) at two ramp rates, $R = 0.2$ MV/cm-s (□) and 0.002 MV/cm-s (○). This field was such that the silicon at the interface was inverted and a strong light source was used to ensure this condition. However, at fields above ≈ 8.5 MV/cm and at the faster rate, a significant voltage drop was present in the Si and not all the capacitors were broken down. The shift of f with ramp rate is clearly evident in this figure. Also, the shift is not constant over the entire distribution. Parameters required to fit the data according to Eq. (17) for f, with the q_i given by Eq. (28), are listed in Table 1. Eight defect types were required to achieve the fit; $q_0 = 1$ for the range of fields covered.

Effect of temperature

As we stated earlier, dependence of f on temperature enters via the temperature dependence of the processes involved in breakdown. For electronic breakdowns, these processes are charge injection and motion, impact ionization, and recombination [1, 7]. The field and temperature dependences of these processes are, in general, separable so that the τ_{i0} and F_{i0} are expected to depend on the temperature. For different defect types the τ_{i0} and F_{i0} are functions of temperature such that for a given test condition (say, F in a life test or R in a ramp test), the ordering of the τ_i could be altered with temperature. Thus, a change in temperature can change the order in time (field) in which the different defects contribute to breakdown in a life (ramp) test.

Screening

A screening procedure is sometimes used which eliminates devices that would otherwise fail at early times when used in an electronic system; for dielectric devices, this is a breakdown test. Let q_{is} be the probability that a breakdown occurs at a defect of the ith type during the screen, and q_{ia} be the same quantity under conditions of actual use. The expectation value for the fraction failed during use $f_a(t)$, not counting devices broken down during the screen, is then

$$f_a(t) = 1 - q_{0s}q_{0a} \exp[-\Sigma\lambda_i q_{is}(1 - q_{ia})], \quad (33)$$

where t is the time under use.

For the screen to be effective, q_{is} should be small for defect types for which q_{ia} is small at use times less than the desired lifetime of the system. The order in time at which defects cause breakdowns depends on the test conditions, *i.e.*, the field in a life test, ramp rate for a ramp test, and in both tests the temperature. Thus, the conditions must be chosen to break down the right defects.

Discussion

A statistical model of breakdown has been presented relating the breakdown distribution to distributions of defects and probabilities of breakdown of defects and defect-free regions. A specific form was chosen for these probabilities [Eqs. (15), (16), and (17)] which assumed that the mean times to breakdown are exponential functions of the field. This form was chosen, basically, because it is a reasonable fit to experimental evidence [3] and lends itself to simplified mathematical manipulations.

Processes involved in electronic breakdown have previously been considered in detail [1, 10-12]. The essential point is that breakdown results from a positive feedback mechanism which consists of the following processes:

1. Charge injection (assumed to be electrons) from the cathode into the insulator.
2. The dynamics of the motion of the injected electrons, which determines their energy distribution and the rate of impact ionization.
3. A resulting positive charge distribution, which can be diminished by recombination with electrons, moving towards the cathode and enhancing the field between the positive charge and the cathode.
4. This field enhancement increases the rate of electron injection and impact ionization—giving rise to further field enhancement, etc.

Computations based on the average rates of these processes indicate that a limited current is possible only below a certain field, which depends on the temperature and properties of the insulator and its interface affecting the enumerated processes. Thus, based on the average rates, breakdown occurs only when this field is exceeded. However, allowing for fluctuations in these processes, breakdowns can occur below this deterministic field. Therefore, in order to derive a form for the mean time to breakdown, a solution which includes the effects of fluctuations is required. Such a solution is as yet, to the best of our knowledge, unavailable. Although an exponential dependence of the mean time to breakdown on the field is unlikely to be an exact solution, it should nevertheless provide a reasonable approximation, since most of the enumerated processes are very strong functions of the field. Furthermore, as already stated, an exponential form is supported experimentally.

Acknowledgments

W. Rausch of this laboratory provided the MOS samples, and D. Nahum, a summer student in our laboratory, aided in data reduction of the experimental results, for which we are thankful.

References and note

1. J. J. O'Dwyer, *The Theory of Electrical Conduction and Breakdown in Solid Dielectrics*, Clarendon Press, Oxford, 1973.
2. P. Solomon, "Breakdown in Silicon Oxide—A Review," *J. Vac. Sci. Technol.* **14,** 1122 (1977).
3. P. Solomon, N. Klein, and M. Albert, "A Statistical Model for STEP and RAMP Voltage Breakdown Tests in Thin Insulators," *Thin Solid Films* **35,** 321 (1976).
4. Morris Shatzkes, Moshe Av-Ron, and Robert A. Gdula, "Defect-Related Breakdown and Conduction in SiO_2," *IBM J. Res. Develop.* **24,** 469 (1980).
5. N. Klein, "Theory of Localized Electron Breakdown in Insulating Films," *Adv. Phys.* **21,** 605 (1972).
6. N. Klein, "Switching and Breakdown in Films," *Thin Solid Films* **7,** 151 (1971).
7. N. Klein, "Electrical Breakdown Mechanisms in Thin Insulators," *Thin Solid Films* **50,** 223 (1978).
8. H. H. DeWit, Ch. Wijenbergand, and C. Crevecoeur, *J. Electrochem. Soc.* **123,** 1479 (1976).
9. A. Berman of IBM GTD, East Fishkill, first published such a relation in a memo to file, "Time-zero dielectric reliability test by the ramp method," 3/21/80. Our relations differ in that he requires a single value for all the F_{i0}. See also E. S. Anolich and G. R. Nelson, "Low-field Time-Dependent Dielectric Integrity," *IEEE Trans. Reliability* **R-29,** 21 (1980). These authors have a similar equation in which they do not specify the ln (F_{i0}/R) term and do not distinguish defect types.
10. T. H. DiStefano and M. Shatzkes, "Impact Ionization Model for Dielectric Instability and Breakdown," *Appl. Phys. Lett.* **25,** 685 (1974).
11. T. H. DiStefano and M. Shatzkes, "Dielectric Instability and Breakdown in SiO_2 Thin Films," *J. Vac. Sci. Technol.* **13,** 50 (1976).
12. Morris Shatzkes and Moshe Av-Ron, "Impact Ionization and Positive Charging in Thin SiO_2 Films," *J. Appl. Phys.* **47,** 3192 (1976).
13. N. Klein and T. Solomon, "Current Runaway in Insulators Affected by Impact Ionization and Recombination," *J. Appl. Phys.* **47,** 4364 (1976).

Received November 15, 1980

The authors are located at the IBM General Technology Division laboratory, East Fishkill Facility, Hopewell Junction, New York 12533.