Downloaded from https://www.everyspec.com

AD624769

RADC-TR-65-330

UNITED STATES AIR FORCE

CLEARINGHOUSE
FOR FEDERAL SCIENTIFIC AND
TECHNICAL INFORMATION
Hardcopy | Microfiche
$ 6.00 | $ 1.50 | 272 pp
ARCHIVE COPY

RADC

RELIABILITY PHYSICS NOTEBOOK

Rome Air Development Center

EDITOR – J. Vaccaro

Battelle Memorial Institute

EDITOR – H. C. Gorton

CONTRIBUTORS

C. W. HAMILTON

B. J. NICHOLSON

A. B. TIMBERLAKE

H. W. RAY

T. S. SHILLIDAY

PREPARED FOR: ROME AIR DEVELOPMENT CENTER
GRIFFISS AIR FORCE BASE, N. Y.

Downloaded from https://www.everyspec.com

PAGES _____

ARE

MISSING

IN

ORIGINAL

DOCUMENT

Downloaded from https://www.everyspec.com

RADC
RELIABILITY PHYSICS NOTEBOOK

AFLC, GAFB, N.Y., 19 Nov 65-524

Downloaded from https://www.everyspec.com

This book was prepared by the Battelle Memorial Institute, Columbus, Ohio, for the U. S. Government under a contract with Rome Air Development Center, Griffiss Air Force Base, New York.

Certain portions of this book are reproduced with permission of respective copyright owners. Any further reproduction of this material may not be made without the permission of the holders of such rights.

Publication of this document has been made possible through the support and sponsorship extended by the Reliability Branch, Engineering Division, of the Rome Air Development Center under Contract No. AF30(602)-3504.

First Edition
Published October 1965

Downloaded from https://www.everyspec.com

FOREWORD

Continued advances in the state of the art of solid-state technology have led to significant improvement in the reliability of electronic component parts. This improvement has resulted from increased uniformity in material properties and greater control in processing procedures, which have reduced the variability among devices and the incidence of randomly occurring failures. Not enough attention, however, has been devoted to understanding the basic mechanisms in materials which underlie degradation and failure of solid state devices.

All device failures may be said to result from changes in the chemical composition or in the physical configuration of the constituent elements of the device. Discovery of these critical chemical or physical change processes makes possible their reduction or elimination, thus improving device reliability. The development of an understanding of the dependence of such processes on environmental stress makes possible the development of mathematical prediction and testing techniques based on fundamental physical principles. Thus, as the improvement and the prediction of reliability begin to be placed on a scientific basis, the reliability engineer will go more and more to the chemist and physicist for answers to his questions. It is expected that this fundamental approach to reliability will ultimately lead to economic engineering procedures for the accurate assessment of the reliability of electronic devices.

The purpose of this Notebook is to make available to the electronics reliability engineer current state-of-the-art information relating to what may be termed the reliability physics of solid state electronic parts. It explains techniques and procedures for obtaining pertinent data on specific part types and methods of utilizing the data in accelerated testing, screening, and reliability prediction programs. Consideration is largely limited to degradation and failure mechanisms which remain after gross mechanical and quality defects have been screened out.

Because the field of reliability physics is yet relatively new, the Notebook is intended as a preliminary effort, and the format has been designed to facilitate modifications and additions to existing data. Wherever possible, data have been presented in useful tabular form. When topics are still largely conjectural, summaries of the results of various investigators are presented. The interest has been confined primarily to discrete devices, with emphasis on those part types with design or compositional characteristics related to microcircuit structures. A brief section on silicon integrated circuits is included.

Downloaded from https://www.everyspec.com

PUBLICATION REVIEW

This report has been reviewed and is approved.

Approved: Joseph Vaccaro
JOSEPH VACCARO
Chief, Reliability Physics Section
Reliability Branch

Approved: [signature]
WILLIAM P. BETHKE
Chief, Engineering Division

FOR THE COMMANDER: [signature]
IRVING J. GABELMAN
Chief, Advanced Studies Group

Downloaded from https://www.everyspec.com

TABLE OF CONTENTS

Contents		Page
SECTION 1.	Mathematical Models in Reliability Physics	1-1
SECTION 2.	Aging and Failure Mechanisms	2-1
SECTION 3.	Physical Properties of Materials and Processes Pertinent to Reliability Physics	3-1
SECTION 4.	Accelerated Testing	4-1
SECTION 5.	Reliability Screening Procedures	5-1
SECTION 6.	The Use of Statistical Methods in Reliability Physics Experiments	6-1

Downloaded from https://www.everyspec.com

SECTION 1. MATHEMATICAL MODELS IN RELIABILITY PHYSICS

Downloaded from https://www.everyspec.com

TABLE OF CONTENTS

	Page
1. Mathematical Models in Reliability Physics	1-1
1.1 Introduction	1-1
1.2 Evolution of Mathematical Models in Reliability Physics	1-1
1.2.1 Early Models	1-1
1.2.2 Current Models	1-2
1.2.3 Future Models	1-3
1.3 Classification of Mathematical Models	1-3
1.4 Limitations of Mathematical Models	1-4
1.5 Mathematical Models of Rate Processes	1-5
1.5.1 The Arrhenius Reaction-Rate Model	1-5
1.5.1.1 Application of Dimensional Analysis to the Arrhenius Model	1-5
1.5.1.2 A Mathematical Derivation of the Arrhenius Model	1-6
1.5.1.3 Interpretation of the Constants of the Arrhenius Model	1-13
1.5.1.4 Application of the Arrhenius Model in Reliability Physics	1-13
1.5.2 The Eyring Reaction-Rate Equation	1-13
1.5.2.1 Application of Dimensional Analysis to the Eyring Model	1-13
1.5.2.2 A Derivation of the Eyring Model as Applied in the Theory of Component Failure	1-17
1.6 Extensions of the Arrhenius and Eyring Models	1-20
1.6.1 Application of the Arrhenius Model to Failure Data	1-21
1.6.1.1 Application to Hazard Rates	1-21
1.6.1.2 Application to Weibull-Distributed Failure Data	1-22

Downloaded from https://www.everyspec.com

TABLE OF CONTENTS
(Continued)

	Page
1.6.2 Extensions of the Eyring Model	1-23
1.6.2.1 Simultaneous Occurrence of Eyring Processes	1-23
1.7 Summary	1-24
REFERENCES	1-27

LIST OF FIGURES

Figure 1.1. Probability Distribution Over the Energy States for Various Values of θ/T	1-10
Figure 1.2. An Ideal Failure Mode	1-19
Figure 1.3. Superposition of Two Simultaneous Eyring Degradation Processes	1-24

Downloaded from https://www.everyspec.com

1. Mathematical Models in Reliability Physics

1.1 Introduction

The successful application of research results in reliability physics depends in part on the efficacy of the mathematical models that are used to characterize physical processes. The models should accurately and precisely describe the aging effects of environmental stresses, operating time, and their interaction. Moreover, the models should be mathematically tractable. If these desired properties can be attained, reliability-physics models can profitably be used for reliability problems such as materials specifications, component design configurations, component screening, performance prediction, etc.

The objective of this section is to describe the types of mathematical models currently used in reliability physics and the types of models that are expected to evolve in the future. In Sections 1.2, 1.3, and 1.4, the trend in mathematical models in reliability physics is briefly outlined, and the limitations of models are discussed. In Sections 1.5 and 1.6, emphasis is centered on the mathematical derivations of the Arrhenius and Eyring models, interpretation of the model parameters, and extensions of these models in reliability analysis. The section is concluded with a brief discussion of other reliability-physics models.

1.2 Evolution of Mathematical Models in Reliability Physics

1.2.1 Early Models

Most early models in reliability were empirically based. That is, models were derived from experimental data according to statistically defined goodness-of-fit criteria. Moreover, most of these models involved determining whether a device had acceptable input-output characteristics. In general, these models were used:

(1) To compare operating characteristics of competing devices

(2) To serve as a basis for improving the operating characteristics of existing devices

(3) To guide the integration of different devices in synthesizing a system having desired operating characteristics.

An example of a model of the above type is the stress-strength interference model (Section 5.2).[1] This model yields a static reliability estimate based on the "overlap" of the distribution of strength for a device type, C(S), and the distribution of stress levels the device may encounter in operation, E(S). Reliability in this model is defined as the probability, P, that no stress will be encountered that exceeds the device's strength. Symbolically, the model is given by:

$$P = \int_0^\infty \left[\int_0^\infty C(S)\, ds \right] E(S)\, ds \ . \qquad (1.1)$$

Downloaded from https://www.everyspec.com

A second example is the propagation of variance model for estimating the variation of a circuit output parameter, y, based on the distributions of variations in the component parameter values, $f_1(x_i), \ldots, f_n(x_n)$.[2] In the model, the input-output function, $y = H(x_1, \ldots, x_n)$ is expanded into a Taylor series. The variance of y is then approximated by taking only the linear terms of the series, yielding

$$V(y) = \sum_i \left(\frac{\partial H}{\partial x_i}\right)^2 \sigma_i^2 + \sum_i \sum_j \left(\frac{\partial H}{\partial x_i}\right)\left(\frac{\partial H}{\partial x_j}\right) \sigma_i \sigma_j \quad , \tag{1.2}$$

where the partial derivatives are evaluated at the mean values of the x_i. Various reliability inferences can be made using the model in conjunction with some distributional assumptions of the circuit output parameters.

The principal shortcoming of such models is that they are concerned with the estimation of static operating characteristics; the models are generally not used to determine how long a device will operate satisfactorily. Thus, the mathematical forms of these models do not explicitly include operating time as a variable. Hence, the effect of time on operating characteristics can only be determined implicitly by evaluating the performance variables in the model at several fixed points in time.

1.2.2 Current Models

Current reliability-physics models differ from early models in two important respects. First, the current models permit at least a partial theoretical interpretation of physical processes, even though these models are generally derived empirically. In a sense, these models may be considered as representing an intermediate transition state in going from purely statistical to highly theoretical models of physical phenomena. Second, the current models explicitly include operating times or time-rates of device parameter change.

These models are generally used for the same purposes as the early models described in the preceding paragraphs. However, in this case, the problem is changed from a relatively simple one – "Will the device operate satisfactorily?" – to the more cogent problem – "Given an initially good device, how long will it operate satisfactorily?" To answer this question, the form of the mathematical models not only includes operating time but also the environmental stresses as explicit parameters.

An example of such a model is the Arrhenius reaction-rate equation, which expresses the time-rate of degradation, R(T), of some device parameter as a function of the operating temperature, T.[3] Symbolically, the model is given by

$$R(T) = e^{A - B/T} \quad . \tag{1.3}$$

The constants A and B are estimated for a particular device type from experimental test data. This model is described in detail in Sections 1.5 and 4.3 of this Notebook.

Downloaded from https://www.everyspec.com

1.2.3 Future Models

The trend in reliability-physics research suggests that future models will be based on physical theories of rate processes. More specifically, the mathematical form of future models will be derivable from theories of aging processes based on phenomena occurring at the atomistic or molecular level. For these models, statistical estimation procedures will probably have to be developed to "fit" the theoretical model properties; this is contrary to earlier models in which statistical theory dictated to a large extent the form of the models. Continuing improvement and refinement of effective screening techniques will be required, however, before degradation models based upon rate processes can be successfully applied.

The important expected advantage of future models over existing models will be their utility in the initial device-development stages in generating "optimum" designs as a function of reliability, desired input-output characteristics, materials properties, design configuration, manufacturing processes, etc. Further, these models can be expected to yield highly efficient and powerful screening and accelerated testing procedures for existing devices.

The modified Eyring reaction rate equation serves as an excellent example of such a model.[3] Analogous to the Arrhenius model, it relates the time-rate of degradation, R(T,S), to the thermal stress, T, and nonthermal stress, S, acting in a device. Symbolically, the model is given by

$$R(T,S) = \left[AT\, e^{-B/kT}\right]\left[e^{(C+D/kT)S}\right] , \qquad (1.4)$$

where A, B, C, and D denote the constants in the model. However, in this case, it is expected that the constants will be subject to theoretical interpretation. For example, the constant B is interpreted as an activation energy.

1.3 Classification of Mathematical Models

It may be convenient to classify mathematical models for rate processes as aggregate models and singular models. The former is related to the gross behavior of a population of devices, the behavior of the group, etc. Conversely, the singular model is associated with behavior of a single unit. In some cases, the singular model can be applied to a large collection of individual units to deduce the properties of the collection. This is the case, for example, in the kinetic theory of gases in which Boyle's law, $pV = RT$, is deducible from postulates regarding the behavior of individual gas molecules.

In principle, it appears plausible that one could deduce population characteristics from knowledge of the individuals that make it up. A primary obstacle here consists of choosing the description of the individual behavior in such a way that it will be correct for the individual and also will permit the transition to populations of individuals while retaining mathematical tractibility. In addition to characterizing the kinetic theory of gases, this approach characterizes probability theory. Here, the objective consists of specifying in axiomatic terms the detailed mathematical structure for small integral values of n and then studying the limits approached, if any, as n becomes arbitrarily large. In many instances, such a transition yields a normal distribution for the population. This kind of activity is usually associated with "probabilities", in contrast to the statistical approach

Downloaded from https://www.everyspec.com

in which the investigations generally begin (and end) with descriptions of populations rather than individuals. It is clear that knowledge of the population characteristics may not yield knowledge of the characteristics of the individuals that make it up, and conversely knowledge of the behavior of one individual may not permit the population characteristics to be derived. Ideally, however, knowledge in one area should complement that in the other.

In the physics of aging, the Arrhenius model appears to be an example of an aggregate model. It describes the gross behavior of a device based on an empirical determination of the "constants" from performance data on a set of nominally identical devices. In general, these constants are evaluated by graphical or numerical methods based on the assumption that a straight-line plot would result if the logarithm of the rate is plotted versus reciprocal temperature. The approach is completely analogous to conventional statistical data analysis. For example, if it is assumed that the data are log-normally distributed, they are plotted on log-probability paper to obtain a straight-line fit to a normal distribution having a corresponding mean and standard deviation. In neither case does the model explain in physical terms why the straight-line assumption is valid. To do this requires assumptions concerning the underlying theoretical structure. Such is the case with the Eyring model. Here, derivation is given in terms of concepts of quantum mechanics so that the "straight line" of the Arrhenius plot is seen to be an approximation of the Eyring curve; but even more important, the constants associated with the "activation energy" and "frequency factor" are derivable, in principle at least, from the basic physical concepts.

1.4 Limitations of Mathematical Models

As with any mathematical model used to describe a physical process, the mathematical models of reliability physics give approximate descriptions over limited ranges of variables. The validity of the approximation over specified limits must be established by empirical evidence. The structuring of empirical information in a form that permits the formulation of a mathematical model is often a difficult challenge. Frequently, such empirical information is summarized in graphs, tables, photographs, X-rays, flow diagrams, and other aids. The translation of such information into mathematical terms is often a first step in the formulation of a model. The translation may require the introduction of new definitions and concepts, together with methods of statistical analysis and data processing. Thus, the range of validity of a model may be considerably reduced because of the wide variability shown in the kinds of input information.

A second limitation of models in general use in reliability physics is associated with the concepts of "population of devices" and "individual device". In many instances, it is easier to characterize the average behavior and the variability about this average behavior for a population of devices than it is for any particular device within the population. In some cases, the characterization of a population is exactly what is required, as in quality-control operations or lot-acceptance sampling plans. However, in many cases of practical importance, it is the behavior of a particular device of a population that is of interest. It is not clear that information regarding a population of devices is applicable to an individual device from that population. In fact, it is often the case that no "population" even exists. It may be known that no other device exists like the one just made, and it may still be meaningful to want to know the characteristics of that particular device. It appears that models giving population characteristics would necessarily be different from those that

Downloaded from https://www.everyspec.com

hold for an individual device. One of the major objectives of current model building for reliability-physics problems consists of finding models that are valid for individual devices. The possibility of such models is evident, even with the Arrhenius or Eyring models where, for example, the activation energies and frequency factors could, in principle, be determined for a single device. After these determinations are made and the numerical values are substituted into the equations, the models become predictors of the behavior of the specific device. To the extent that this can be done, it is not necessary to determine population characteristics in order to describe approximately the characteristics of a specific device. Alternatively, in practice it is highly desirable to apply constants determined with one device to others of the same set.

In attempting to describe performance characteristics over time for a specific device, it is convenient to suppose that the device will be subjected to various levels of stress during its life. In order to formulate a model, the general term "stress" must be made specific, such as thermal stress, voltage stress, stress due to vibration, shock, radiation, etc. In most practical situations, it is very difficult to isolate these stresses with sufficient clarity to permit a mathematical structuring of the composite stress as a function of time. Some of the stresses may interact so that the effect of a given level of vibration at high temperatures, for example, may be quite different from the effect of the same level of vibration at low temperatures. Other stresses may operate independently and simultaneously in time; still others may operate sequentially in time. The latter may occur, for instance, when an impurity must first penetrate a canister before it can attack the surface of the encapsulated device.

1.5 Mathematical Models of Rate Processes

1.5.1 The Arrhenius Reaction-Rate Model

The Arrhenius equation was determined empirically in 1889 to account for the influence of temperature on rate of inversion of sucrose. In its original form, the equation was given by

$$k = Ae^{-E/RT} \quad , \tag{1.5}$$

where A and E denote constants in the model. The constant A has been labeled the "frequency factor" and is invariant over relatively small temperature ranges. The constant E has been labeled the "experimental activation energy".

Subsequent to the original postulation of the model, it has become generally accepted that a relationship of this type represents the temperature dependence of the rates of most chemical reactions and of some physical processes. Extension of this equation to describe temperature-dependent degradation rates of electronic components has resulted from empirical investigations analogous to those associated with earlier applications to chemical processes. The validity of the Arrhenius model as applied to electronic components is based primarily on the fact that it is an excellent approximation to the more theoretically sound Eyring model over the range of temperature conditions normally encountered.

1.5.1.1 Application of Dimensional Analysis to the Arrhenius Model

Dimensional analysis is usually not used where theoretical results can be derived. It is used in those areas where the relevant variables are believed to be known, but the

Downloaded from https://www.everyspec.com

functional relations among them are not known. The advantage of dimensional analysis in such a situation is the fact that it yields the proper products of the variables and yields the minimum number of such products. The unknown functions are then usually determined by plotting one dimensionless product against another and obtaining the mathematical form of the relation by empirical methods.

Application of dimensional analysis to the Arrhenius model, as carried out below, may appear trivial. However, this analysis does, in fact, show that something is missing in the Arrhenius equation if it is to truly relate degradation rate to temperature. Specifically, the problem is to derive the Arrhenius relation,

$$R = e^{A-B/T} \quad , \tag{1.6}$$

by the methods of dimensional analysis. The following lists the physical variables and their dimensions in a system of units denoted by mass M, length L, time t, and temperature T:

Physical Quantity	Symbol	(M, L, t, T) Dimensions
Time rate of degradation	R	(0, 0, -1, 0)
Temperature	T	(0, 0, 0, 1)

Examination of the above tabulation shows that it is not possible to form any dimensionless products. This suggests that either there is no relation between the rate of degradation and temperature or some relevant physical quantities have been omitted. Because the Arrhenius model involves only R and T, it cannot be derived by dimensional analysis. This implies that an "essentially empirical" character is associated with the Arrhenius model. If the Arrhenius relation were based on "fundamental" principles of physics, it would be possible to obtain a form that "resembles" it by using only dimensional arguments.

1.5.1.2 A Mathematical Derivation of the Arrhenius Model

Considerable success has been obtained in applying the Arrhenius model to accelerated-test data for various components. Because of the empirical nature of these applications, it is desirable to attempt to relate the model to fundamental physical properties by means of theoretical concepts. It is shown in this section that the Arrhenius model is valid for a "perfect crystalline solid". It should be realized that this derivation does not constitute a theoretical physical basis for applying the Arrhenius model to electronic-component aging processes. However, it does serve to suggest the form of the theory of aging processes, and why the Arrhenius model appears to work so successfully.

A derivation of the Arrhenius model for a "perfect crystalline solid" uses several basic intermediate results of thermodynamics. These intermediate results are briefly described below; in each case more complete, or analogous, derivations are found in Reference 4.

Downloaded from https://www.everyspec.com

The Bohr-Sommerfeld approximation states that the cyclic integral of the momentum of a particle is given by

$$\oint (m\dot{x})\,dx = ih \quad , \tag{1.7}$$

where i is an integer and h denotes Planck's constant. This approximation is valid for the vibrational energy of a simple oscillator. Moreover, by replacing dx with $\dot{x}\,dt$, it is seen that the above integral becomes

$$\oint \frac{1}{2} m (\dot{x})^2 \, dt = ih/2 \quad , \tag{1.8}$$

and this may be shown to be twice the total energy of an oscillator. Thus, the energy associated with the intervals between successive values of i is given by $h\nu$, where ν denotes the frequency of the oscillation. The energy is then associated with the midpoints of the intervals having end points: 0, $h\nu$, $2h\nu$, $3h\nu$, . . ., so that

$$\epsilon_i = \left(i + \frac{1}{2}\right) h\nu \quad . \tag{1.9}$$

Suppose that the average energy is denoted by $\langle\epsilon\rangle$ and that P_i denotes the probability that the energy of the oscillator is equal to ϵ_i. The "least prejudiced" assignment of the probabilities P_i is that which maximizes the entropy,

$$= -K \sum_{i=0}^{\infty} P_i \ln P_i \quad , \tag{1.10}$$

under the constraints,

$$\sum_{i=0}^{\infty} P_i = 1 \quad , \tag{1.11}$$

and

$$\sum_{i=0}^{\infty} P_i \epsilon_i = \langle\epsilon\rangle \tag{1.12}$$

Using the method of Lagrange multipliers to solve for the P_i, the maximum entropy in Equation (1.10) under the constraints given by Equations (1.11) and (1.12) yields

$$P_i = e^{-\psi - \beta\epsilon_i} = \frac{e^{-\beta\epsilon_i}}{\sum_{i=0}^{\infty} e^{-\beta\epsilon_i}} \tag{1.13}$$

Downloaded from https://www.everyspec.com

and

$$\psi = \ln \sum_{i=0}^{\infty} e^{-\beta \epsilon_i} \tag{1.14}$$

where ψ and β are the Lagrange multipliers.* The function ψ is the Massieu function and is the natural logarithm of Z, the partition function; that is,

$$\psi = \ln Z \quad , \tag{1.15}$$

where

$$Z = \sum_{i=1}^{\infty} e^{-\beta \epsilon_i} \quad .$$

With $\epsilon_i = (1 + 1/2)h\nu$, it may be shown that

$$Z = 1/[2 \sinh (\beta h \nu/2)] \quad . \tag{1.16}$$

Although this result is derived for a perfect diatomic gas, it appears that the same result holds for the perfect crystalline sold. The Massieu function becomes

$$\psi = \ln Z = -\ln [2 \sinh (\beta h \nu/2)] \quad . \tag{1.17}$$

Assume that the solid consists of 3N oscillators, which all have the same frequency, ν. This is the Einstein solid. Adding the ψ-functions for the oscillators yields

$$\psi = -\sum_{n=1}^{3N} \ln [2 \sinh (\beta h \nu/2)]$$

or simply

$$\psi = -3N \ln [2 \sinh (\beta h \nu/2)] \tag{1.18}$$

*The functional to be maximized is

$$F = -K \sum_{i=0}^{\infty} P_i \ln P_i - \psi \left(\sum_{i=1}^{\infty} P_i - 1 \right) - \lambda \left(\sum_{i=1}^{\infty} P_i \epsilon_i - \langle \epsilon \rangle \right) \quad .$$

In these terms, the Lagrangian multiplier β is equal to the quantity $(\lambda - 1)$. This substitution serves to simplify the calculations in the following developments.

Downloaded from https://www.everyspec.com

Setting $\beta = 1/(kT)$ for the solid gives

$$\beta h\nu = 2k\beta\left(\frac{h\nu}{2k}\right) = \left(\frac{2k}{kT}\right)\left(\frac{h\nu}{2k}\right) \quad . \tag{1.19}$$

Then, setting $(h\nu/2k) = \theta$, the Einstein temperature, gives $\beta h\nu = 2\theta/T$, and

$$\psi = -3N \ln [2 \sinh (\theta/T)] \quad . \tag{1.20}$$

Returning to a single oscillator of the Einstein solid, it can be seen that

$$\psi = - \ln [2 \sinh (\theta/T)] \quad . \tag{1.21}$$

Now, for the P_i, it is seen from Equations (1.12), (1.16), and (1.20) that

$$P_i = \frac{e^{-\beta\epsilon_i}}{e^{\psi}} \quad ,$$

which becomes

$$P_i = [2 \sinh (\theta/T)]\, e^{-\beta\epsilon_i} \quad , \tag{1.22}$$

and with $\epsilon_i = (i + 1/2)h\nu$, $i = 0, 1, 2, \ldots$, the equation becomes

$$P_i = [2 \sinh (\theta/T)]\, e^{-\beta(i+1/2)h\nu} \quad . \tag{1.23}$$

From the above [Equation (1.19)], it is seen that the exponent in Equation (1.23) is

$$\beta (i+1/2) h\nu = (2)(i+1/2)(\theta/T) \quad ,$$

so that

$$P_i = [2 \sinh (\theta/T)]\, e^{-(i+1/2)(\theta/T)(2)} \quad , \tag{1.24}$$

The probability P_i is expressed as a function of θ/T, the ratio of the Einstein temperature to absolute temperature. Roughly, this distribution of P_i over i appears as shown in Figure 1.1.

Smooth curves have been used to approximate the discrete distribution, which is actually defined only at integral values of i.

Downloaded from https://www.everyspec.com

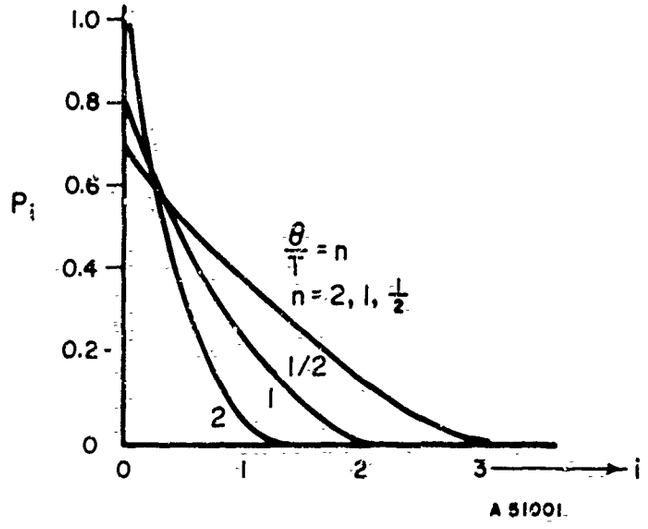


FIGURE 1.1. PROBABILITY DISTRIBUTION OVER THE ENERGY STATES FOR VARIOUS VALUES OF θ/T

Now, suppose (θ/T) is fixed so that one of the above distributions is under consideration. At a fixed time, consider a large population of oscillators and select one of these at random. The above distribution represents the probability of finding the randomly selected oscillator in energy state ϵ_i. Then assume that if the selected oscillator is operated over a sufficiently long period of time, its energy states will fluctuate from one energy state to another in accordance with this probability distribution. This constitutes an "ergodic" assumption. Now, consider the fraction of the time for which the oscillator is in an energy state ϵ_i, where i exceeds some fixed integer (say i_o). The fraction is given by the sum of the probabilities to the right of i_o in Figure 1.1. This probability is given by

$$\underline{P} = \sum_{i=i_o}^{\infty} P_i = 2 \sinh (\theta/T) \sum_{i=i_o}^{\infty} e^{-(i+1/2)(\theta/T)(2)} \quad (1.25)$$

To evaluate this probability, let

$$Z_o = \sum_{i=i_o}^{\infty} e^{-(i+1/2)(\theta/T)(2)} \quad (1.26)$$

and find

$$Z_o = e^{-i_o(\theta/T)(2)} \sum_{j=0}^{\infty} e^{-(j+1/2)(\theta/T)(2)} \quad (1.27)$$

Downloaded from https://www.everyspec.com

The sum has been represented above by Z so that

$$Z_o = Z\, e^{-i_o(\theta/T)(2)} \quad ,$$

or

$$Z_o = \frac{e^{-i_o(\theta/T)(2)}}{2 \sinh(\theta/T)} \quad . \tag{1.28}$$

Thus,

$$\underline{P} = \sum_{i=i_o}^{\infty} P_i = 2 \sinh(\theta/T)\, Z_o \quad , \tag{1.29}$$

and elimination of Z_o yields the result:

$$\underline{P} = e^{-i_o(\theta/T)(2)} \quad . \tag{1.30}$$

Next, consider two oscillators at two absolute temperatures T_1 and T_2 with $T_1 < T_2$. The right "tails" of the energy distributions are denoted by $\underline{P}_1$ and $\underline{P}_2$, respectively. The ratio of these probabilities is given by

$$\underline{P}_2/\underline{P}_1 = e^{-i_o(\theta/T_2)(2)}/e^{-i_o(\theta/T_1)(2)}$$

or

$$\underline{P}_2/\underline{P}_1 = e^{-i_o\theta[(1/T_2)-(1/T_1)](2)} \tag{1.31}$$

This equation has the Arrhenius form given by

$$\ln(\underline{P}_2/\underline{P}_1) = -B[(1/T_2)-(1/T_1)] \quad , \tag{1.32}$$

where the slope constant, B, is equal to $(i_o\theta)(2) = (ih\nu/2k)(2) = i(h\nu/k)$.

In order to identify this result with the Arrhenius model, it must be shown that the $\underline{P}$ corresponds to either degradation rate or a failure rate, or both. Instead of attempting a direct proof of this, the following assumptions regarding failure mechanisms are made. The failure rate (or degradation rate) of a component part is proportional to the probability that the energy state of the part exceeds some fixed "abusive" energy state. After a sufficiently long time of operation, the accumulation of time in the "abusive" energy states will exceed some "critical value" and produce a failure (or degradation exceeding some limit). Prior to such a failure, the "degradation" simply reflects the gradually accumulating time in the abusive state. This "theory of failure" is distinguished by the fact that it requires "long" operating times to produce "failures", and these "failures" are "inevitable" in the sense that long operating times combined with extremely small probabilities of abusive energy states are certain to produce observable failures. The general principle

Downloaded from https://www.everyspec.com

is suggestive of the assumptions that frequently underlie the Poisson distribution: a very large number of opportunities for an event of very small probability will yield actual occurrences of the event.

The "ergodic" assumption and the accumulation of "abuse" as the failure theory then establish the connection with the Arrhenius model.

A further formal connection with the Arrhenius model is given by the following argument. The Arrhenius model may be said to require the "logarithm of a time rate" to be a linear function of a "reciprocal stress". To interpret the $\underline{P}$ as a time rate, note that after t hours of operation the component part has accumulated $(\underline{P}t)$ hours of abuse. The "time rate of accumulation of abuse" is then given by

$$\frac{\text{hours of abuse}}{\text{total hours of operation}} = \frac{(\underline{P}t)}{t} = \underline{P} \quad .$$

Thus, $\underline{P}$ is numerically equal to the time rate of accumulation of abuse. Consequently, if the Arrhenius model holds:

$$\ln \underline{P}_1 = A - B(1/S_1) \tag{1.33}$$

and

$$\ln \underline{P}_2 = A - B(1/S_2) \quad , \tag{1.34}$$

where S_1 and S_2 denote two generalized stresses. Then

$$(\underline{P}_2/\underline{P}_1) = e^{-B[(1/S_2) - (1/S_1)]} \quad . \tag{1.35}$$

Consider the amount of operating time required to accumulate 1 hour of abuse. This is seen to be equal to

$$\frac{\text{total hours of operation}}{\text{hours of abuse}} = \frac{t}{\underline{P}t} = \frac{1}{\underline{P}} \quad .$$

The acceleration factor may be defined as follows:

$$\tau = \frac{\dfrac{\text{total hours of operation at normal stress, } (S_1)}{\text{hours of abuse at normal stress}}}{\dfrac{\text{total hours of operation at high stress, } (S_2)}{\text{hours of abuse at high stress}}} = \frac{(1/\underline{P}_1)}{(1/\underline{P}_2)} = \underline{P}_2/\underline{P}_1 \quad . \tag{1.36}$$

This shows that the equation for the acceleration factor based on the Arrhenius model is also obtained as a result of these arguments.

Downloaded from https://www.everyspec.com

1.5.1.3 Interpretation of the Constants of the Arrhenius Model

The slope constant, B, of the Arrhenius rate equation is proportional to the "experimental activation energy" of a component degradation process. Although this interpretation suggests a theoretical explanation derivable for an individual component, it is important to recognize the empirical nature of the constant B. It is a measure of average behavior of a set of nominally identical components and is estimated from an experiment on a sample of such devices. Thus, in an empirical sense, B is proportional to the incremental energy in excess of a "ground energy state" that must be acquired in order for a particular degradation process to occur.

The constant A in the Arrhenius equation is the intercept of the plot of the logarithm of degradation rate versus reciprocal temperature. In physical terms, this constant could be associated with the frequency factor of a degradation process in an analogous manner to the relationship of the constant B to activation energy.

1.5.1.4 Application of the Arrhenius Model in Reliability Physics

The Arrhenius model has been used to relate component failure rates and other related reliability measures to operating temperature. More recently, the model has been used as a mathematical basis for generating and analyzing accelerated-test data. The value of this model in accelerated testing is that it provides a basis for determining over what temperature ranges specific failure mechanisms are dominant. Where a single mechanism is dominant over a given temperature range, "true" acceleration factors can be calculated relating 1 hour of component operation at some increased temperature level to τ hours of operation at a reference temperature level.

Application of the Arrhenius model to accelerated testing is treated in detail in Section 4. In that section, the properties, limitations, and step-by-step analysis procedures associated with the Arrhenius model are described.

1.5.2 The Eyring Reaction-Rate Equation

The basic objection to the Arrhenius equation is that, because of its empirical nature, the constants of the model cannot be theoretically interpreted in terms of the chemical properties of the reactants involved in a given rate process. The Eyring reaction-rate equation, because of its foundation in quantum theory, alleviates this problem. This model was derived by Eyring in 1935, based on contributions to the quantum theory of specific heat. These contributions, due principally to Einstein, Debeye, and Van Karman, were developed approximately 25 years after the Arrhenius equation was postulated.

The importance of the Evring model to reliability physics stems from the hope that the constants of the model can ultimately be related to component materials and structure. If the hope can be realized, it is obvious that the Eyring model will constitute a powerful tool for achieving specified component reliability through appropriate initial design.

1.5.2.1 Application of Dimensional Analysis to the Eyring Model

This section includes a more precise statement of the method of dimensional analysis and demonstrates that the Eyring model "responds" to this approach. The basic principle

Downloaded from https://www.everyspec.com

of dimensional analysis is known as Buckingham's Pi Theorem.[5,6] The theorem may be stated as follows:

Theorem. Let $x_1, x \ldots, x_n$ have the n x m dimensional matrix of rank r = n-k:

$$A = \begin{bmatrix} P & R \\ Q & S \end{bmatrix} , \tag{1.37}$$

where P is a nonsingular r x r matrix. Then, if $f(x_1, \ldots, x_n)$ is dimensionally homogeneous, it follows that

$$f(x_1, \ldots, x_n) = 0 \tag{1.38}$$

is equivalent to

$$f(1, \ldots, 1, \pi_1, \ldots, \pi_k) = 0 , \tag{1.39}$$

in which

$$\pi_1 = x_1^{e_{i1}} x_2^{e_{i2}} \ldots x_n^{e_{in}} , \quad i = 1, \ldots, k,$$

are k = n - r independent and dimensionless quantities with the k x n matrix of exponents

$$E = (-QP^{-1}, I_k) , \tag{1.40}$$

where I_k denotes the k x k unit matrix.

This theorem is applied to the following variables, which are listed together with their dimensions in a system of units denoted by mass M, length L, time t, and temperature T:

Physical Quantity	Symbol	[M, L, t, T] dimensions
Temperature	T	[0, 0, 0, 1]
Boltzmann's constant	k	[1, 2, -2, -1]
Planck's constant	h	[1, 2, -1, 0]
Time rate of degradation (failure)	R	[0, 0, -1, 0]
Activation energy	ΔE	[1, 2, -2, 0]

Downloaded from https://www.everyspec.com

The matrix A of dimensions is given by

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 2 & -2 & -1 \\ 1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 0 \\ 1 & 2 & -2 & 0 \end{bmatrix} . \tag{1.41}$$

It is easily verified that the rank of this matrix is equal to 3. By reordering the columns to get a nonsingular matrix in the upper left corner of A, the following matrix is obtained:

$$A = \begin{bmatrix} 0 & 0 & 1 & . & 0 \\ 2 & -2 & -1 & . & 1 \\ 2 & -1 & 0 & . & 1 \\ . & . & . & . & . \\ 0 & -1 & 0 & . & 0 \\ 2 & -2 & 0 & . & 1 \end{bmatrix} = \begin{bmatrix} P & R \\ Q & QP^{-1}R \end{bmatrix} , \tag{1.42}$$

where the submatrices P, Q, R, and $QP^{-1}R$ correspond to the denoted portions of the matrix A. Now, assume that $f(T, k, h, R, \Delta E)$ is dimensionally homogeneous. Buckingham's Pi Theorem leads to the conclusion that

$$f(T, k, h, R, \Delta E) = 0 \tag{1.43}$$

is equivalent to

$$f(1, 1, 1, \pi_1, \pi_2) = 0, \tag{1.44}$$

where

$$\pi_1 = T^{e_{11}} k^{e_{12}} h^{e_{13}} R^{e_{14}} \Delta E^{e_{15}} ,$$

and

$$\pi_2 = T^{e_{21}} k^{e_{22}} h^{e_{23}} R^{e_{24}} \Delta E^{e_{25}} ,$$

Here, π_1 and π_2 are independent, dimensionless quantities having a matrix of exponents given by

$$E = (-QP^{-1}, I_2) . \tag{1.45}$$

Downloaded from https://www.everyspec.com

Reference to the partitioned form of the matrix A shows that

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 2 & -2 & -1 \\ 2 & -1 & 0 \end{bmatrix} . \tag{1.46}$$

The inverse matrix is found to be

$$P^{-1} = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & 1 \\ -1 & -1 & 1 \\ 1 & 0 & 0 \end{bmatrix} , \tag{1.47}$$

so that

$$QP^{-1} = \begin{bmatrix} 0 & -1 & 0 \\ 2 & -2 & 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & 1 \\ -1 & -1 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

or

$$QP^{-1} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \end{bmatrix} . \tag{1.48}$$

Thus, the matrix of exponents is given by

$$E = (QP^{-1}, I_2) = \begin{bmatrix} -1 & -1 & 1 & 1 & 0 \\ -1 & -1 & 0 & 0 & 1 \end{bmatrix} , \tag{1.49}$$

and we have

$$\pi_1 = T^{-1} k^{-1} h^1 R^1 \Delta E^0 = Rh/kT \tag{1.50}$$

and

$$\pi_2 = T^{-1} k^{-1} h^0 R^0 \Delta E^1 = \Delta E/kT. \tag{1.51}$$

Downloaded from https://www.everyspec.com

Consequently,

$$f(1, 1, 1, Rh/kt, \Delta E/kT) = 0,$$

or more simply,

$$f(Rh/kT, \Delta E/kT) = 0. \tag{1.52}$$

Buckingham's Pi Theorem has reduced the number of variables from 5 tc 2. If the assumption is made that the preceding equation can be solved explicitly for π_1, it follows that

$$(Rh/kT) = g(\Delta E/kT), \tag{1.53}$$

or

$$R = (kT/h)\, g(\Delta E/kT) \quad . \tag{1.54}$$

In a typical application of dimensional analysis, the functional form of g would be determined by empirical means. If an exponential decay were found for g, it would follow that

$$g(\Delta E/kT) = ae^{-\Delta E/kT} \quad , \tag{1.55}$$

so that

$$R = a(kT/h)\, E^{-\Delta E/kT} \quad . \tag{1.56}$$

This equation has the form of Eyring's rate equation. Thus, it has been shown that, in contrast to the Arrhenius model, the Eyring model is responsive to the approach of dimensional analysis. This means that the physical quantities involved in the Eyring model satisfy a necessary condition for the relation to be "fundamental", rather than "empirical". This fundamental nature would appear to be essential to success in relating the basic material properties of an electronic component part to the "physics of the aging process".

1.5.2.2 A Derivation of the Eyring Model as Applied in the Theory of Component Failure

It has been found empirically that rates of failure of certain electronic components can be related to the operating conditions by an equation of the Arrhenius type:

$$k = Ae^{-E/RT} \quad . \tag{1.57}$$

Here, k is the reaction rate, A is a constant of proportionality, E is an activation energy, R is the gas constant, and T is the absolute temperature. This result is of rather general character, for it follows as a good approximation to the Eyring reaction-rate equation.[7,8] The latter equation will be derived below in a heuristic manner so that some of the principles that underlie the application will be clearer.

Downloaded from https://www.everyspec.com

Before the derivation is given, however, the concepts used in classical mechanics regarding potential-energy surfaces should be reviewed. If a collection of particles of some type exists, the instantaneous configuration of the system can be specified by giving the values of the coordinates of its constituents. Since each configuration has associated with it a potential energy, a potential energy "contour" or "surface" can be associated with the entire system of particles. A simple graphic example of such a contour is provided by a ball at the surface of the earth. The position of the ball can be completely described by specifying its distance from the center of the earth and its position on the surface. Its potential energy is just that due to gravity. Then, the energy contour of the ball on the ground is just the actual earth's surface. The tendency of the ball to move to the lowest neighboring point is an instance of the tendency of any system to progress to the configuration in which its potential energy is minimized. If the system has N particles, 3N coordinates must be specified in order to specify the configuration of the system. The potential-energy surface will involve 3N-1 coordinates in this space. In order to reduce the mathematical complexity that arises due to the dimensionality of the coordinate system, in practice the original coordinate system, say $(x_1, \ldots, x_n)$ is transformed into a new system $(x_1', \ldots, x_n')$. The x_i' of the transformed system for which the potential energy surface is given by $y = f(x_1', x_2', \ldots, x_n')$ can be described as a superposition of n independent motions in terms of the normal coordinates $x_1', x_2', \ldots, x_n'$.

Consideration of the previous example can provide further insight. Suppose the region of the earth's surface is hilly and that the ball is near the bottom of a shallow depression on the side of one of these hills. If the ball (i.e., the system) is only displaced (activated) a small amount from the bottom of the depression (equilibrium), it will merely roll about within the depression. If it is further displaced (activated), say by a series of impulses, it will eventually have enough energy that, if it heads in the correct direction, it can roll over the lip of the depression and on down the side of the hill. The ball, while at the bottom of the depression, is said to be in a state of metastable equilibrium, since it is stable under small displacements but unstable under sufficiently large ones. It will be in a state of stable equilibrium only at the bottommost point on the surface.

The application of this notion to a theory of failure rates is quite direct. Consider a typical component. It will, like any other system, have a potential-energy surface associated with the possible configurations of its constituent particles. Suppose the quality control is sufficiently ideal that, when the component is newly manufactured, it is in one specified configuration. After it has been operating for a period of time, its behavior alters, indicating a change in its configuration. Eventually, it ceases to operate satisfactorily and finally fails. In terms of its position on the potential-energy surface, it passes from a state of metastable equilibrium when new, to a state of lower energy when it has failed. The path on the potential-energy surface that is traced out as the component passes through these configurations will be termed the failure mode. If the various failure modes that the components follow do not differ substantially from one another, then they can be conceived of as a sort of valley. The component starts out in a local depression in this energy valley. By means of thermal or other excitations, it eventually gets enough energy to leave the depression and go to a lower state – that where it is said to have failed. (It is assumed that the failure is permanent, i.e., the component does not cure itself spontaneously.) Suppose that, as the component progresses down this valley, its instantaneous states are always at the lowest point of a normal cross section of the valley.

Downloaded from https://www.everyspec.com

Such a path then represents a normal coordinate axis since there is no tendency of the system to move to either side, and a single coordinate, ξ, suffices to specify the motion. This ideal failure mode is pictured in Figure 1.2.

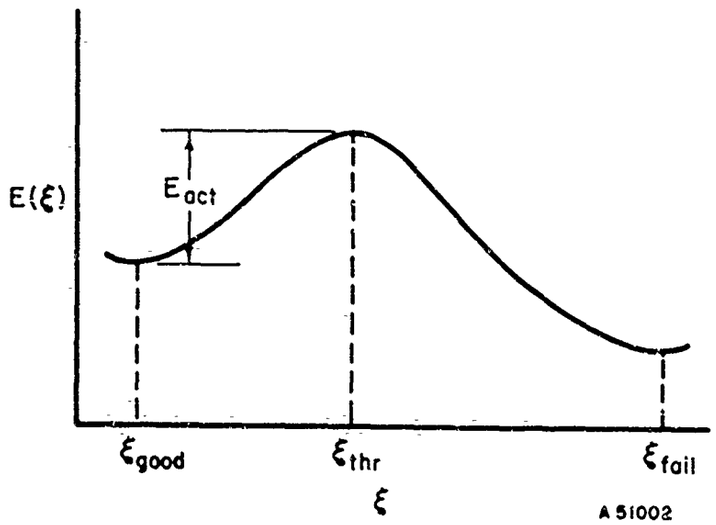


FIGURE 1.2. AN IDEAL FAILURE MODE

The potential energies along the failure mode, $E(\xi)$, are plotted against the failure-mode coordinate, ξ. The energy E_{act} must be supplied to the component before it can pass from the state of satisfactory performance, ξ_{good}, to the state of failure, ξ_{fail}. The threshold of failure is at ξ_{thr}.

Ordinarily, since failure occurs rather slowly, it is reasonable to assume that the system is very nearly in equilibrium at any point. Then the probability distribution for the position, ξ, and rate, $\dot{\xi}$, of the component along its failure mode is essentially the Maxwell-Boltzmann distribution[9]:

$$f(\xi, \dot{\xi})\, d\xi\, d\dot{\xi} = Ne^{-[E(\xi) + \frac{1}{2}m^* \dot{\xi}^2]/RT}\, d\xi\, d\dot{\xi} \quad , \tag{1.58}$$

where $f(\xi, \dot{\xi})$ is the probability density function, N is a normalization factor, and m^* is a parameter related to the inertia of the component along the failure mode. The gas constant R is used in (1.58) rather than the Boltzmann constant, κ, because it is assumed that the microscopic constituents have energies obtained by averaging the macroscopic energies $E(\xi)$ and $\frac{1}{2} m^* \dot{\xi}^2$ (the kinetic energy) over all the constituents. It is assumed that the entire component can be divided into a number of subcomponents that are small macroscopically but large microscopically.

Downloaded from https://www.everyspec.com

The fraction of these subcomponents within ξ_{good} and $\xi_{good} + d\xi$ is obtained by integrating the expression in (1.58) over all values of the $\dot{\xi}$, with ξ held at ξ_{good}:

$$\int_{-\infty}^{\infty} f(\xi_{sat}, \dot{\xi})\, d\xi\, d\dot{\xi} = Ne^{-E\,\xi_{good})/RT}\, d\xi \int_{-\infty}^{\infty} e^{-m^*\dot{\xi}^2/2RT}\, d\dot{\xi}$$

$$= N\left(\frac{2\pi RT}{m^*}\right)^{1/2} e^{-E\,(\xi_{good})/RT}\, d\xi. \qquad (1.59)$$

The fraction of the subcomponents that reach the threshold of failure, ξ_{thr}, multiplied by the mean speed with which they fail, is obtained by multiplying (1.58) by $\dot{\xi}$, evaluating at ξ_{thr}, and integrating over all positive values of $\dot{\xi}$:

$$\int_0^{\infty} \dot{\xi} f(\xi_{thr}, \dot{\xi})\, d\xi\, d\dot{\xi} = Ne^{-E\,(\xi_{thr})/RT}\, d\xi \int_0^{\infty} \dot{\xi} e^{-m^*\dot{\xi}^2/2RT}\, d\dot{\xi}$$

$$= \frac{N}{2}\left(\frac{2RT}{m^*}\right)^{3/2} e^{-E\,(\xi_{thr})/RT}\, d\xi\ . \qquad (1.60)$$

The failure rate, k, is given by the ratio of (1.60) to (1.59). Thus,

$$k = \frac{RT}{m^*\pi^{1/2}}\, e^{-E_{act}/RT}, \qquad (1.61)$$

where the activation energy, E_{act}, is the difference between $E(\xi_{thr})$ and $E(\xi_{good})$. The expression in (1.61) is the Eyring reaction-rate formula. Since the exponential function varies much more rapidly with T than does the linear factor in (1.61), the Arrhenius formula (1.57), is a very good approximation to (1.61) over small ranges of temperature.

It is predicted by (1.61) that an increase in the temperature of a component will increase the failure rate. The application of other stresses, such as a high voltage on a capacitor, is represented through changes in E_{act}. The nature of these changes depends on the manner in which the stress affects the potential-energy surface, and generally they need to be determined empirically.

It should be realized that the arguments presented above are not fully rigorous, and (1.61) should not be expected to have universal validity. However, the arguments do seem to be sufficiently suggestive of physical cases so that it can be expected that (1.61) will be a useful tool, at least as a first approximation, in the correlation of data on failure rates.

1.6 Extensions of the Arrhenius and Eyring Models

Treatment of the Arrhenius and Eyring models in Section 1.5 was concerned principally with the derivation of the models in terms of chemical rate processes. The extension of these models to reliability physics requires consideration of how they can be adapted to specific problems, including application to component-failure data and model

Downloaded from https://www.everyspec.com

formulation for simultaneous and sequential aging processes. In the following paragraphs, these problems are formulated in terms of the Arrhenius and Eyring models.

1.6.1 Application of the Arrhenius Model to Failure Data

1.6.1.1 Application to Hazard Rates

The most common method for calculating acceleration factors for components having exponentially distributed failure times is given by:

$$\tau = \frac{\lambda'}{\lambda} \quad , \tag{1.62}$$

where λ' denotes the hazard rate at some high stress level, and λ denotes the hazard rate at a normal (reference) stress level. It is shown below that Equation (1.62) for the acceleration τ is obtained through application of the Arrhenius model.

To obtain acceleration factors using the Arrhenius model,

$$R(T) = e^{A-B/T}, \tag{1.63}$$

it is required to obtain a measure, f(Q), linear over time, of the number of components failing per unit time. For failure data a suitable quality parameter, Q, is the cumulative proportion failing over time, F(t). Under the assumption of exponentially distributed failure times, it is seen that

$$1-F(t) = e^{-\lambda t} \quad . \tag{1.64}$$

Hence, the required transformation f(Q) is seen to be

$$f(Q) = -\ln [1-F(t)] = \lambda t \quad . \tag{1.65}$$

Thus, from (1.63) and (1.65), it is seen that

$$R(T)\,t = e^{A-B/T}\,t = \lambda t \quad . \tag{1.66}$$

Now, suppose test times under a normal stress, T, and some increased stress level, T', are such that the same degradation of quality has occurred; i.e., f(Q) = f'(Q). Then from Equation (1.66)

$$\left(e^{A-B/T}\right)\,t = \left(e^{A-B/T'}\right)\,t' \quad , \tag{1.67}$$

and solving for t yields

$$t = e^{-B\left(\frac{1}{T'} - \frac{1}{T}\right)}\,t' \tag{1.68}$$

Downloaded from https://www.everyspec.com

Defining the acceleration factor τ by the relation $t = \tau t'$, therefore, yields

$$\tau = e^{-B\left(\frac{1}{T'} - \frac{1}{T}\right)} = R(T')/R(T) \tag{1.69}$$

Finally, from Equations (1.66) and (1.69),

$$\tau = \frac{\lambda'}{\lambda} \quad . \tag{1.70}$$

Thus, the ratio of the hazard rates obtained at temperature stress levels T and T' is the appropriate acceleration factor based on the Arrhenius model.

1.6.1.2 Application to Weibull-Distributed Failure Data

In applying the Arrhenius model to data where times to failure are exponentially distributed, the required function for linear degradation over time was easily obtained by algebraic manipulation of the terms in the cumulative distribution function. However, for more general distributions, this requirement is more difficult to satisfy. In these cases, it is either necessary to make additional assumptions or resort to approximations in order to apply the Arrhenius model. In the following, it is seen that, under one additional assumption, the Arrhenius model can be applied to Weibull-distributed failure data.

The cumulative Weibull distribution function is given by

$$F(t) = 1 - e^{-t^{\beta}/\alpha} \quad . \tag{1.71}$$

Reducing (1.71) to obtain a linear function over time yields

$$\left\{-\ln\left[1 - F(t)\right]\right\}^{1/\beta} = \alpha^{-1/\beta}\, t \quad . \tag{1.72}$$

The left side of Equation (1.72) depends on the parameter β; hence, the equation cannot be plotted and the rate $\alpha^{-1/\beta}$ obtained, unless β is known and is independent of the applied stress.

Under the assumption that the parameter β of the Weibull distribution is independent of stress over the range of stress levels involved in testing, the acceleration obtained by testing at some increased stress level, T', can be calculated in the same manner as in Equations (1.65) through (1.70) of Section 1.6.1.1. The resultant formula for τ is

$$\tau = (\alpha/\alpha')^{1/\beta} \quad . \tag{1.73}$$

Downloaded from https://www.everyspec.com

1.6.2 Extensions of the Eyring Model

1.6.2.1 Simultaneous Occurrence of Eyring Processes

Current knowledge of the physics of the aging process of transistors suggests the possibility that several aging processes may occur simultaneously in the device. One or more "surface" processes may occur, with each one having its own partition constant and activation energy. Similarly, there may be several "bulk" processes. With this conceptual framework, the over-all aging process is a composite one made up of various sub-processes. Much of the work on accelerated tests has involved the assumption that "true acceleration" has occurred if the points representing a sequence of stresses "line up" on the Eyring plot. Because of the importance of this assumption, and the possibility of the simultaneous occurrence of several aging processes, this subject is considered in greater detail.

If a component quality parameter degrades according to two Eyring processes, and it is assumed that the effects of the processes are independent, then the resultant degradation process will not be an Eyring process. To see this, let

$$R_1(T) = A_1 \frac{kT}{h} e^{-\Delta E_1/kT} \tag{1.74}$$

and

$$R_2(T) = A_2 \frac{kT}{h} e^{-\Delta E_2/kT} \tag{1.75}$$

denote two Eyring processes acting simultaneously in a device. The constants A_1 and A_2 denote the frequency factors of processes 1 and 2, respectively; and ΔE_1 and ΔE_2 denote the activation energies of the respective processes. Then superposition of the two processes yields

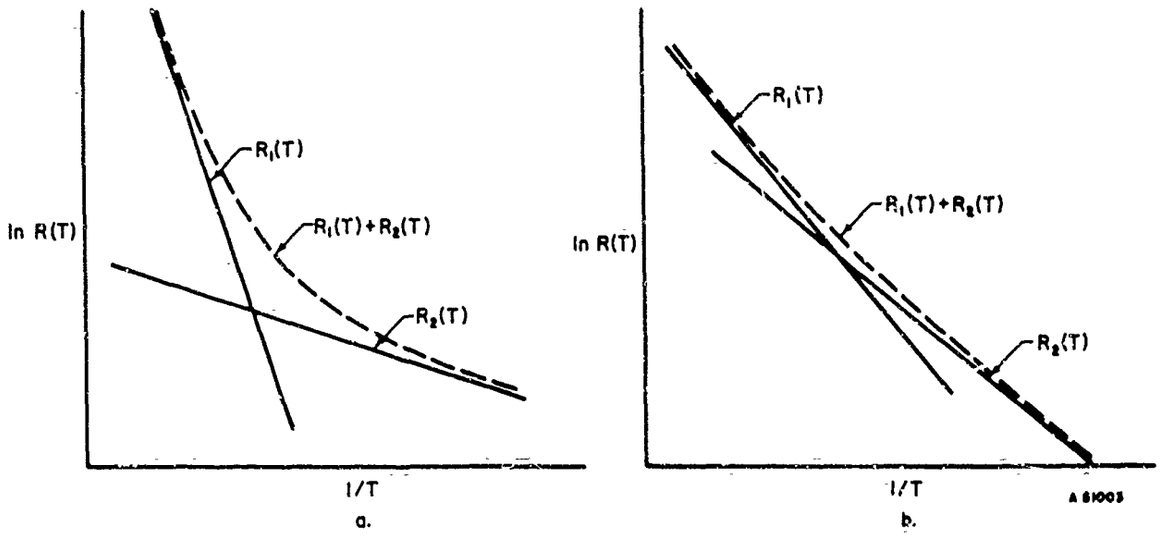
$$\begin{aligned} R &= R_1(T) + R_2(T) \\ &= A_1 \frac{kT}{h} e^{-\Delta E_1/kT} + A_2 \frac{kT}{h} e^{-\Delta E_2/kT} \quad . \end{aligned} \tag{1.76}$$

Taking logarithms yields

$$\ln R = \ln \frac{kT}{h} \left[A_1 e^{-\Delta E_1/kT} + A_2 e^{-\Delta E_2/kT} \right] \quad . \tag{1.77}$$

It is clear from Equation (1.77) that a linear relation between ln R and 1/T does not obtain; hence, the occurrences of two simultaneous Eyring processes does not yield an Eyring process (Figure 1.3). Moreover, it follows as a simple extension of the above that the superposition of n Eyring processes will not be an Eyring process.

Downloaded from https://www.everyspec.com



a. b.

FIGURE 1.3. SUPERPOSITION OF TWO SIMULTANEOUS EYRINC DEGRADATION PROCESSES

Some insight can be obtained regarding the data-analysis problem, when more than one process is taking place simultaneously, by rewriting Equation (1.77) in the form

$$R = A_1 \frac{kT}{h} e^{-\Delta E_1/kT} \left[1 + \frac{A_2}{A_1} e^{-(\Delta E_2 - \Delta E_1)/kT}\right] . \qquad (1.78)$$

If the two processes have significantly different activation energies (i.e., $\Delta E_2 >> \Delta E_1$), then the empirically generated Eyring plot will yield a curve which changes slope readily over the temperature region where the $R_1(T)$ is close to $R_2(T)$. In this case it is relatively easy to construct the two separate processes and to estimate the temperature domain over which each process is dominant (Figure 1.3a). On the other hand, if ΔE_1 and ΔE_2 are close in numerical value, it is seen from Equation (1.78) that an empirical plot of R versus $1/T$ may appear to be approximately linear with intercept $(A_1 + A_2)$ and slope $(\Delta E_1 + \Delta E_2)/2k$ (Figure 1.3b). The magnitude of the difference $(\Delta E_2 - \Delta E_1)$ for which such an erroneous result may be obtained will depend on the degree of randomness in the experiment. Such randomness may result from variations in the physical characteristics of the components under test, variation in experimental control of the test environment, and variations in parameter measurements due to instrumentation errors. One way of potentially increasing the probability of catching an error of this type is to extend the temperature range for which the Eyring plot is generated.

1.7 Summary

The treatment of mathematical models in reliability physics in the preceding paragraphs has centered on the Arrhenius and Eyring reaction-rate equations. Concentration has been on these particular models because of their repeated application in several

Downloaded from https://www.everyspec.com

sections of this Notebook. It should not be construed that there is a general lack of models of rate processes. On the contrary, many mathematical models have been reported in the reliability literature. In fact, a major problem is to determine the relationships among existing models in terms of the assumptions from which they are derived and the range of conditions over which they are applicable.

In some instances, several particular models can be derived from a single general model. For example, it is shown in Reference 10 that the Delbruck model for random failures in semiconductor materials is "derivable" from the expression

$$S \cong k \ln D \quad , \tag{1.79}$$

where S denotes entropy, k is Boltzmann's constant, and D is a measure of disorder. This equation arises in statistical physics from the work of Boltzmann and Gibbs. From (1.79),

$$\ln \frac{1}{D} = -\frac{S}{k} \quad .$$

Now, defining a constant W_1 by

$$S = -W_1/kT \quad ,$$

and letting $D = \theta/\sigma$ where θ denotes component mean-life, and σ denotes a second constant, the Delbruck model,

$$\theta = \sigma e^{W_1/kT} \quad , \tag{1.80}$$

is obtained. To obtain the Arrhenius model, it is sufficient to set $\theta^{-1} = \lambda = R(T)$, $W_1/k = B$, and $\sigma^{-1} = A$. Thus, substituting into Equation (1.80) yields the Arrhenius model

$$R(T) = Ae^{-B/T} \quad . \tag{1.81}$$

Many models have been developed to fit specific types of aging processes in particular classes of components. Oxidation processes occurring in thin-film devices are an excellent example of this type of modeling. In Reference 11, a failure model for a thin-film device is derived, based on the oxidation-rate equation

$$R = aP^b e^{-E/kT} \quad ,$$

where a and b are constants, P is the pressure, and E is the activation free energy. On this basis, it is shown that, if failure occurs when a change in resistance, r_f, is realized, the time to failure is given by

$$\lambda = C \left(\frac{r_f}{1 + r_f} \right) a^{-1} P^{-b} e^{E/kt} \quad , \tag{1.82}$$

Downloaded from https://www.everyspec.com

where C denotes film thickness. Various generalizations of this model are developed, based on a generalized expression for the rate equation. In Reference 12, a similar model is derived through the equation

$$Q = 4a^2 \frac{\partial n}{\partial x} e^{-E/kT} \quad , \tag{1.83}$$

where Q denotes the rate of diffusion of a gas through a metal, a is the activation distance, and $\partial n/\partial x$ is the concentration gradient at a distance x from the metal-air interface. Further, in Reference 13 an expression for rate of resistance change based on an oxidation process is given as

$$R(T) = Ae^{-E/R_g T} \quad , \tag{1.84}$$

where R_g denotes the universal gas constant. Moreover, it is shown that (1.84) can be formulated as a special case of the equation

$$\theta = T(a + \ln t) \quad , \tag{1.85}$$

which is a general relationship between time and temperature. This function has been applied extensively in the field of metallurgy.

A third type of modeling in reliability physics centers on the requirement for analyzing experimental data. In Reference 14, the mathematical properties of various generalized reliability models are treated in depth. Particular emphasis is placed on two models: (1) a simple wear model and (2) the generalized Allen model. The simple wear model is characterized by the equation

$$\frac{d[\mu(t)]}{dt} = \mu'(t) = -W[S(t)] \quad , \tag{1.86}$$

where $\mu'(t)$ denotes rate of dissipation of some attribute μ of a device, and W denotes a wear-rate function that has as its argument the stress function S(t). It is shown that the Arrhenius function can be expressed as a special case of (1.86) wherein

$$W[S(t)] = Ae^{-B/S(t)} \quad . \tag{1.87}$$

The generalized Allen model can be described in terms of the ratio of the hazard function under a stress environment, $\varphi_j(t)$, to the hazard function under a use-stress environment $\varphi_u(t)$ by

$$\frac{\varphi_j(t)}{\varphi_u(t)} = \frac{W^*[S_j(t)]}{W^*[S_u(t)]} \quad . \tag{1.88}$$

W is defined according to (1.86), and the mathematical form of $\varphi(t)$ follows from the assumed distribution of component failure times. Letting

$$W^*[S_i(t)] = [V_i(t)]^n \quad , \tag{1.89}$$

Downloaded from https://www.everyspec.com

where $V_i(t)$ denotes applied voltage as a function of time, the power-law function is obtained that is commonly employed in capacitor-aging experiments:

$$\frac{\varphi_j(t)}{\varphi_u(t)} = \left[\frac{V_j(t)}{V_u(t)}\right]^n \quad . \tag{1.90}$$

These brief examples of the kinds of mathematical models currently available in reliability physics suggest the close interrelationships that exist. The choice of which model, if any, is applicable in a particular instance depends on the assumptions associated with the model, the range of conditions over which the model is valid, and the intended purpose of the analysis.

REFERENCES

(1) Roberts, N. H., Mathematical Models in Reliability Engineering, McGraw-Hill Book Company, Inc., New York (1964), 129-134.

(2) Tukey, J. W., "The Propagation of Errors, Fluctuations, and Tolerances. Basic Generalized Formulas", Technical Report No. 10, Statistical Techniques Research Group, Princeton University.

(3) Thomas, R. E., and Gorton, H. C., "Research Toward a Physics of Aging of Electronic Component Parts", Physics of Failure in Electronics, 2, Edited by Goldberg and Vaccaro, Cato Show Printing Company (1964), 25-60.

(4) Tribus, M., Thermostatics and Thermodynamics, D. VanNostrand Company, 1961.

(5) Buckingham, E., "On Physically Similar Systems", Phys. Rev., 4, 354-376 (1914).

(6) Brand, L., "The Pi Theorem of Dimensional Analysis", Archive for Rational Mechanics and Analysis, 1, 33 (1957).

(7) Glasstone, S., Laidler, K. J., and Eyring, H., The Theory of Rate Processes, McGraw-Hill Book Company, Inc., New York (1941).

(8) Hill, T. L., Introduction to Statistical Thermodynamics, Chapter 11, Addison-Wesley (1960).

(9) Kennard, E. H., Kinetic Theory of Gases, McGraw-Hill Book Company, Inc., New York (1938).

(10) Pershing, A. V., and Hollingsworth, G. E., "Derivation of Delbruck's Model for Random Failure (for Semiconductor Materials); Its Identification With the Arrhenius Model; and Its Experimental Verification", Physics of Failure in Electronics, 2, Edited by Goldberg and Vaccaro, Cato Show Printing Company (1964), 25-60.

Downloaded from https://www.everyspec.com

(11) Goldberg, M., Horberg, A., Stewart, R., and Levinson, D., "Comprehensive Failure Mechanism Theory – Metal Film Resistor Behavior", Physics of Failure in Electronics, 2, Edited by Goldberg and Vaccaro, Cato Show Printing Company (1964), 25-60.

(12) Wortman, J. J., and Burger, R. M., "A Fundamental Failure Mechanism in Thin-Film Metal Dielectric Structures Observable as a Generated Voltage", Physics of Failure in Electronics, 2, Edited by Goldberg and Vaccaro, Cato Show Printing Company (1964), 25-60.

(13) Bretts, G., Kozol, J., and Lampert, H., "Failure Physics and Accelerated Testing", Physics of Failure in Electronics, 2, Edited by Goldberg and Vaccaro, Cato Show Printing Company (1964), 25-60.

(14) Winter, B. B., Denison, C. A., Hietala, H. J., and Greene, F. W., "Accelerated Life Testing of Guidance Components", Technical Documentary Report No. AL TDR64-235, Research and Technology Division, Wright-Patterson Air Force Base, September 3, 1964.

Downloaded from https://www.everyspec.com

SECTION 2. AGING AND FAILURE MECHANISMS

Downloaded from https://www.everyspec.com

TABLE OF CONTENTS

	Page
2. Aging and Failure Mechanisms	2-1
2.1 Introduction.	2-1
2.1.1 Definitions of Aging and Failure Processes	2-1
2.1.2 Definitions of Failure Terms.	2-2
2.2 Review of Physical Processes Pertinent to Reliability Physics.	2-2
2.2.1 Chemical Reactions	2-2
2.2.1.1 Oxidation	2-2
2.2.1.1.1 Oxidation of Metals	2-3
2.2.1.1.2 Anodic Oxidation.	2-5
2.2.1.1.3 Oxidation of Semiconductors	2-6
2.2.1.2 Reduction	2-6
2.2.1.3 Corrosion	2-6
2.2.2 Lattice Imperfections	2-7
2.2.2.1 Introduction	2-7
2.2.2.2 Frenkel Defects	2-7
2.2.2.3 Schottky Defects	2-7
2.2.2.4 Impurity Defects	2-7
2.2.2.5 Color Centers.	2-8
2.2.2.6 Dislocations	2-8
2.2.2.7 Twin Boundaries	2-9
2.2.2.8 Stacking Faults	2-9
2.2.3 Surfaces.	2-10

Downloaded from https://www.everyspec.com

TABLE OF CONTENTS
(Continued)

	Page
2.2.4 Diffusion	2-13
2.2.4.1 Diffusion From an Infinite Source	2-14
2.2.4.2 Diffusion From a Finite Source	2-15
2.2.4.3 Diffusion Mechanisms	2-16
2.2.5 Ionization	2-17
2.3 Review of Failure Modes in Selected Electronic Parts	2-17
2.3.1 Semiconductor Devices	2-17
2.3.1.1 General Aspects of Semiconductor-Device Reliability	2-17
2.3.1.2 Review of Construction of Various Semiconductor Devices	2-18
2.3.1.2.1 Surface-Barrier Transistors	2-18
2.3.1.2.2 Grown-Junction Transistors	2-22
2.3.1.2.3 Alloy-Junction Transistors	2-22
2.3.1.2.4 Microalloy Transistors	2-23
2.3.1.2.5 Surface-Alloy Transistors	2-23
2.3.1.2.6 Diffused-Junction Transistors	2-23
2.3.1.2.7 Double-Diffused Transistors	2-24
2.3.1.2.8 Triple-Diffused Planar Transistors	2-24
2.3.1.2.9 Alloy-Diffused Transistors	2-24
2.3.1.2.10 Drift and Microalloy-Diffused Transistors	2-24
2.3.1.2.11 Mesa Transistors	2-25
2.3.1.2.12 Planar Transistors	2-25
2.3.1.2.13 Epitaxial Transistors	2-26

Downloaded from https://www.everyspec.com

TABLE OF CONTENTS
(Continued)

	Page
2.3.1.2.14 Unijunction Transistors	2-26
2.3.1.2.15 Field-Effect Transistors	2-26
2.3.1.2.16 Point-Contact Diodes	2-27
2.3.1.2.17 Gold-Bonded Diodes	2-27
2.3.1.2.18 Alloyed and Diffused Diodes	2-27
2.3.1.2.19 Tunnel and Backward Diodes	2-27
2.3.1.3 Influence of Crystal Imperfections on Semiconductor-Device Reliability	2-28
2.3.1.3.1 Edge Dislocations	2-28
2.3.1.3.2 Twin Boundaries	2-30
2.3.1.3.3 Stacking Faults in Epitaxial Layers	2-30
2.3.1.3.4 Bulk Diffusion of Impurities	2-31
2.3.1.3.5 Thermal Effects in High-Power Transistors and Diodes	2-31
2.3.1.3.6 Recombination Radiation as a Cause of Failure	2-33
2.3.1.4 The Influence of Surface Effects on Semiconductor-Device Reliability	2-33
2.3.1.5 Mechanical Defects	2-36
2.3.1.5.1 Presence of Foreign Particles Inside the Container	2-36
2.3.1.5.2 Container Imperfections	2-37
2.3.1.5.3 Imperfect Wire-Lead Bonds	2-37
2.3.1.5.4 Imperfect Bonding of Die	2-37
2.3.1.5.5 "Purple Plague"	2-37
2.3.1.5.6 Tool Damage	2-37

Downloaded from https://www.everyspec.com

TABLE OF CONTENTS
(Continued)

	Page
2.3.2 Failure Mechanisms at Conductor-Insulator Interfaces	2-38
2.3.3 Failure Mechanisms in Thin-Film Resistors	2-40
2.3.3.1 Background	2-40
2.3.3.2 Review of Failure Possibilities in Thin-Film Resistors . .	2-41
2.3.3.2.1 Oxidation	2-41
2.3.3.2.2 Selective Oxidation or Evaporation of Solute . . .	2-43
2.3.3.2.3 Agglomeration	2-43
2.3.3.2.4 Precipitation of Solute	2-43
2.3.3.2.5 Order-Disorder Transformations	2-44
2.3.3.2.6 Homogenization	2-44
2.3.3.2.7 Stress Relief	2-44
2.3.3.2.8 Vacancy Condensation	2-44
2.3.3.3 Specific Failure Studies	2-45
2.3.3.3.1 Nickel-Chromium ("Evanohm") Thin-Film Resistors	2-45
2.3.3.3.2 Tin Oxide Film Resistors.	2-48
2.3.3.3.3 Tantalum Thin-Film Resistors	2-49
2.3.4 Integrated Circuits	2-50
2.3.4.1 General Aspects of Integrated Circuit Reliability . .	2-50
2.3.4.2 Bulk Effects	2-54
2.3.4.3 Surface Effects	2-54
2.3.4.4 Contacts	2-55
2.3.4.5 Isolation	2-56

Downloaded from https://www.everyspec.com

LIST OF FIGURES

		Page
Figure 2.1.	Types of Surface Oxidation of Metals.	2-4
Figure 2.2.	Cross Section of an Edge Dislocation	2-9
Figure 2.3.	View of a Screw Dislocation.	2-9
Figure 2.4.	Effect of Surface Charge on the Energy Band of an n-Type Semiconductor	2-11
Figure 2.5.	Energy-Level Diagram of Metal-Semiconductor Interface. . . .	2-12
Figure 2.6.	Experimental and Theoretical Variation of Zinc Concentration With Distance at 1000°C in GaAs	2-16
Figure 2.7.	Structure of Surface-Barrier Transistor	2-22
Figure 2.8.	Cross Section of Crystal Bar at the Base Contact in a Grown-Junction Transistor	2-22
Figure 2.9.	Configuration of an Alloy-Junction Transistor.	2-23
Figure 2.10.	Structure of a Triple-Diffused Transistor	2-24
Figure 2.11.	Structure of Mesa Transistor	2-25
Figure 2.12.	Planar Transistor Showing Oxide Passivation.	2-25
Figure 2.13.	Unijunction Transistor Structure	2-26
Figure 2.14.	Field-Effect Transistor Structure	2-26
Figure 2.15.	Diffusion Enhancement by the Dislocations of a Small-Angle Grain Boundary.	2-29
Figure 2.16.	Structure of Four-Layer, High-Power Diode With Grid Emitter Structure	2-32
Figure 2.17.	Surface Charge Separation on a Passivating Layer Under the Influence of the Fringing Field of a p-n Junction	2-35

Downloaded from https://www.everyspec.com

LIST OF TABLES

Table 2.1. Failure Modes From Lift-Test Failures of Mesa Transistors . . . 2-19

Table 2.2. Failure Modes During Life Test and Operation Typical of 2N696 and 2N697 Transistors 2-20

Table 2.3. Summary of Metallurgical Changes That Influence Electrical Resistance 2-42

Downloaded from https://www.everyspec.com

2. Aging and Failure Mechanisms

2.1 Introduction

The increased requirements for long-time, maintenance-free operation and the increased complexity of electronic systems impose severe demands on the reliability of electronic component parts. In meeting these demands, the reliability of electronic parts has been improved to the point that the determination of the levels of reliability achieved has become a major problem.

Conventional mass-testing techniques to determine mean time to failure of statistically significant sample sizes have become prohibitive with respect to both time and cost. A rapidly developing electronic technology has brought about the enigmatic situation that electronic parts are often replaced by improved models before the reliability of the parts can be determined. In addition, since no implications as to failure mechanisms can be made from statistical information on the probability of failure, such reliability information is of little value to the improvement of reliability.

Further, mass-testing techniques are particularly suited to the determination of failure rates that are strongly influenced by random factors. With the improvement of electronic technology, the random factor in part reliability, which is principally due to human error and material variability, is steadily diminished. Thus, with continued technological improvements, part reliability will become more strongly related to physical processes and mechanisms that are common to a given class of devices. In such circumstances, it should be possible to predict reliability from the basis of an understanding of the physical processes that tend to modify the electrical output of the devices.

2.1.1 Definitions of Aging and Failure Processes

Physical processes affecting part reliability may result in the gradual change of electrical properties, causing the device output to drift out of tolerance, or, upon reaching a critical level, may result in sudden and complete malfunction. In either case, wear-out mechanisms may be expected to have been operative over the entire life of the device, their observation often depending on whether they directly or indirectly affect the operating properties of the device. For example, the oxidation of the resistive element of a thin-film resistor will result in a change of resistance proportional to the oxidation rate, the value of resistance being affected from the initial operating time. On the other hand, the oxidation of the surface of a semiconductor may cause an increase in surface conductance, resulting in a conducting path in parallel with that of the bulk. The surface conductance may not become predominant until after many hours of operation and until then would not be manifest in the operating characteristics of the device. Thus, the early detection and measurement of the change in surface conduction would require special techniques involving more than the measurement of normal operating characteristics. In addition, physical processes leading to ultimate failure may occur that do not affect electrical properties until the moment of failure. Such physical processes generally result in mechanical failure of the device, i.e., the slow formation of the Au-Al-Si eutectic resulting in the fracture of a silicon device at the lead-semiconductor interface, or the gradual introduction of strain into a refractory material from repeated temperature excursions or vibration ultimately resulting in fracture of the material. It is difficult to differentiate between

Downloaded from https://www.everyspec.com

these types of catastropic failures and those not related to time-dependent processes, such as burn-out due to overload or fracture from physical impact.

2.1.2 Definitions of Failure Terms

Indicator of Failure. The electrical manifestation of device malfunction, such as the change over time of the output characteristics under conditions of constant input and constant stress.

Failure Mode. The electrical property of the device from which failure is defined, such as an open or short circuit, or excessive leakage current.

Failure Mechanism. The basic physical process or change at the atomic or molecular level responsible for the observed failure mode, such as ion adsorption which induces a conductive circuit path in the surface layer of a semiconductor.

2.2 Review of Physical Processes Pertinent to Reliability Physics

A relatively small number of physical processes may account for the majority of failures from time-dependent phenomena in the general class of solid-state electronic component parts. In some cases, notably diffusion and chemical reactions, rate equations have been developed from which the time dependence of the processes may be deduced. In this subsection, the pertinence of different kinds of physical processes to electronic-part reliability will be discussed, and, where appropriate, the rate equations associated with the process will be developed.

2.2.1 Chemical Reactions

2.2.1.1 Oxidation

Oxidation originally was defined as a chemical reaction in which oxygen took part as a reagent; however, it is also applied to many chemical reactions in which oxygen might play no part but in which a species of reagent might lose one or more valence electrons. For instance, in the reaction between sodium metal and chlorine to form sodium chloride, where a sodium atom gives up one valence electron to a chlorine atom, the sodium may be considered to have been oxidized.

However, most oxidation reactions of significance for electronic component-part reliability do involve reactions with either oxygen or water vapor. For instance, surface oxides on metal-film resistors play a part in determining the aging of the resistor. Similarly, oxides are used as surface passivating agents in semiconductor devices and as a dielectric medium in thin-film capacitors.

Downloaded from https://www.everyspec.com

2.2.1.1.1 Oxidation of Metals

Since oxidation of metals or alloys usually proceeds with the growth of an insulating surface film, the primary result of oxidation is a decrease in cross section and a consequent increase in resistance. Resistivity and temperature coefficient of resistance are not usually affected unless the oxidation is selective or unless a composition gradient exists through the film thickness.

An important parameter in the surface oxidation of a metal is the ratio of the volume of oxide produced to the original volume of metal or alloy that has undergone oxidation. For all metals, including beryllium but excluding the other alkali and alkali earth metals, this ratio is greater than one. Consequently, the oxide formed will tend to be in a state of compression. When, however, the ratio is less than one, the oxide formed is porous. Under these conditions, gaseous oxygen can readily penetrate the pores to the underlying metal, and passivation is impossible. The oxide then grows at a constant rate with time*.

It is a simple matter to find the volume ratio when only one metal is involved. However, in the case of alloys, the evaluation of volume ratios is more difficult. An approximate value can be calculated from a knowledge of the composition, and this might suffice for prediction of the probable time law.

The observation of a linear oxidation rate might indicate that the volume ratio is less than unity. This is true for alloys with a high content of a noble metal that does not participate in the oxidation reaction and therefore does not contribute to the increase of the volume of the metallic phase on oxidation.

A linear relation may also result from the formation of a volatile oxide. Even if the oxide does not disappear as rapidly as it is formed, the continuous breakup of the surface layer due to partial vaporization exposes new metal surfaces for attack by oxygen. The result is the same as with porous oxide layers in that the reaction takes place at the metal surface. This is the probable reason for the linear oxidation law found for tungsten and molybdenum under certain conditions.

When the volume ratio is greater than unity and volatile compounds are absent, a coherent surface film of oxide grows as a result of oxidation. If, in these layers, a diffusion process is rate determining, then, providing no aging effects alter the diffusion constant and the surface area remains constant, one may postulate that the rate of increase of film thickness is inversely proportional to film thickness, as

$$\frac{dx}{dt} = \frac{k_1}{2x} \quad , \tag{2.1}$$

where

k_1 = a constant

x = film thickness.

*This discussion assumes an unrestricted air ambient. In encapsulated components, such as resistive elements of a microcircuit, oxidation rates are, of course, controlled by a limited source of oxygen.

Downloaded from https://www.everyspec.com

Integrating,

$$x^2 = k_1 t \quad , \tag{2.2}$$

if $x = 0$ then $t = 0$.

This is the frequently encountered parabolic law of oxidation. In practice, it is often found that an additional constant (not an integration constant) is necessary to fit empirical data:

$$x^2 = k_1 t + k_2 \quad . \tag{2.3}$$

Occasionally, experimental data have been found to agree with a cubic relationship,

$$x^3 = kt \quad . \tag{2.4}$$

A logarithmic relationship, as

$$x = a \log (bt + 1) \quad , \tag{2.5}$$

has also been found to apply in certain cases. The different types of oxidation behavior are illustrated in Figure 2.1.

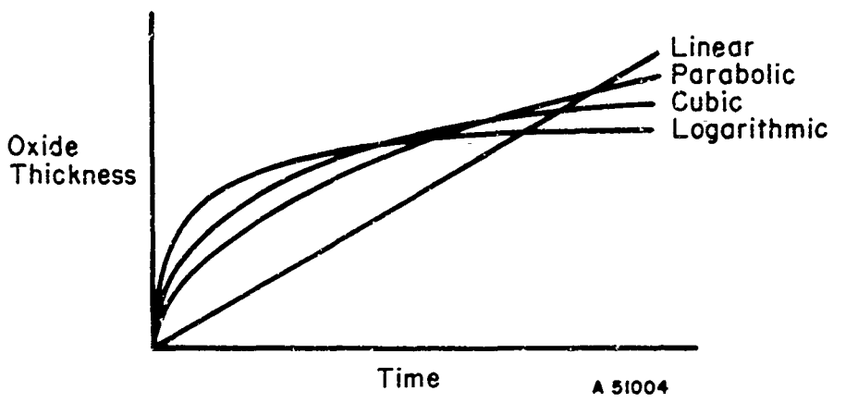


FIGURE 2.1. TYPES OF SURFACE OXIDATION OF METALS

Parabolic oxidation rates are the most common and so far have proved easiest to explain theoretically. The constant k_1 in the parabolic equation depends on temperature according to the Arrhenius equation:

$$k_1 = A\, e^{-\frac{B}{T}} \quad , \tag{2.6}$$

where B is proportional to the activation energy and can be obtained from the slope of a log k_1 versus T^{-1} plot.

Downloaded from https://www.everyspec.com

Eyring's reaction-rate theory, when applied to diffusion processes through a growing oxide film, assumes the presence of a transition state at the top of an energy barrier between the initial and final states of diffusion, the transition state being in equilibrium with the initial states. Two terms are introduced: a term $\frac{kT}{h}$ (k = Boltzmann's constant, h = Planck's constant) that arises from the average velocity of the activated species across the barrier and a term,

$$e^{-\frac{\Delta F}{RT}} ,$$

proportional to the number of activated species as a function of the free-energy barrier and absolute temperature, where R is the universal gas constant.

Since

$$\Delta F = \Delta H - T\Delta S \quad , \tag{2.7}$$

where H = enthalpy and S = entropy, it can be shown that

$$k_1 = \frac{2kT}{h} d^2 e^{\frac{\Delta S}{R}} e^{-\frac{\Delta H}{RT}} \quad , \tag{2.8}$$

where d is the closest cation-cation distance, assuming the diffusion of cations in the surface layer is the rate-determining process. Since the corresponding experimental activation energy, Q, is related to ΔH by the relation

$$Q = \Delta H + RT \quad , \tag{2.9}$$

we then obtain

$$k_1 = \frac{2kT}{h} d^2 \exp\left[1 + \frac{\Delta S}{R} - \frac{Q}{RT}\right] \quad . \tag{2.10}$$

2.2.1.1.2 Anodic Oxidation

A special type of oxidation often encountered in electronic components is anodic oxidation. Anodic oxidation is a process whereby a conducting part is made the anode in an electrolytic cell, and by passing current through the cell, the anode is oxidized as a result of reactions occurring between the anode and anion current carriers. The oxide film formed may be either porous or compact, depending both on the nature of the anode metal and the composition of the electrolyte. If the oxide is porous, its thickness usually increases indefinitely with the passage of current. However, in the case of compact oxides, which are only obtained with aluminum, tantalum, niobium, titanium, zirconium, hafnium, and silicon, the oxide quickly assumes a limiting thickness that is a function of applied voltage, temperature, and anode material. These compact oxides usually have an amorphous crystal structure and are good insulators. Recently, anodic oxidation has been accomplished by striking a glow discharge in oxygen at low pressure instead of using an electrolyte.(1)

Downloaded from https://www.everyspec.com

2.2.1.1.3 Oxidation of Semiconductors

Most of what has been discussed in the oxidation of metals applies equally to semiconductors; however, an additional effect arises in the case of semiconductors that is of no consequence in the case of metals. A semiconductor is usually doped with an impurity to impart a desired conductivity; in the subsequent growth of an oxide film on the semiconductor surface, the concentration of impurity in the oxide film will depend on the segregation coefficient of the impurity with respect to the semiconductor and its oxide. Depending on the value of the segregation coefficient, the impurity can either diffuse preferentially into the oxide or be rejected by the growing oxide film and pile up in the region of the semiconductor-oxide interface. For instance, aluminum accumulates preferentially in an oxide film growing on silicon, while phosphorus is rejected by the oxide.

2.2.1.2 Reduction

Reduction in its initial narrow sense is defined as the opposite of oxidation. In its broadest sense, it involves the donation of a valence electron to an atom or ion. For instance, many metal oxides are reduced to the metal by heating in hydrogen; the hydrogen is thereby oxidized to water vapor.

Reduction is not so significant a process as oxidation in affecting the reliability of electronic component parts because of the difficulty of removing such oxidizing agents as oxygen and water vapor from the gaseous ambient. Encapsulated electronic parts are often sealed in inert atmospheres of nitrogen or argon, or in a reducing atmosphere of hydrogen, to minimize the oxidation process. In a hydrogen atmosphere, a reduction process can, of course, occur. As with oxidation, the rate laws controlling the reduction process are of the Arrhenius or Eyring type.

2.2.1.3 Corrosion

Another phenomenon in electronic components dependent on the occurrence of a chemical reaction is corrosion. In almost all instances, the corrosion process requires the presence of small amounts of moisture. Nearly all metals except the noble metals, when exposed to a salt-water spray, are eventually corroded and the metal chloride is formed. Likewise, of the various fluxes used in soldering, a number are quite corrosive, especially in humid environments.

A type of corrosion occurring when two different metals are in contact in a humid environment is known as galvanic corrosion. When a film of water forms at the junction of the two metals, its appreciable conductivity enables an ordinary galvanic electrolytic cell to be set up with the higher metal in the electrochemical series acting as the anode. Since there are normally return paths for the currents that flow, electrochemical reactions are enabled to proceed with the effect of corroding the anode metal.

Under these circumstances, the rate at which the corrosion proceeds, although dependent on an activation energy, may be limited by the amount of moisture available in the ambient.

Downloaded from https://www.everyspec.com

2.2.2 Lattice Imperfections

2.2.2.1 Introduction

Models of the crystal structure of solids have played an important role in the development of solid-state physics. In these models, crystals are characterized by a regular spatial array of atoms, ions, or molecules repeated periodically in three dimensions to form the crystal lattice. The atoms, ions, or molecules situated at each lattice point are held together by binding forces and are continuously vibrating about their equilibrium positions.

2.2.2.2 Frenkel Defects

The concept of the periodic crystal lattice led to the successful accounting for many of the properties of solids, such as specific heat, X-ray diffraction, and propagation of sound. Its application to the quantum mechanical description of electronic conduction in solids was also successful. However, a number of physical properties defied description in terms of the ideal crystal model; for example, diffusion or ionic conduction. It was necessary to introduce the idea of imperfections or irregularities in the crystal lattice to explain these effects. This is not unreasonable since the kinetic theory of matter postulates that atoms or molecules are in some state of motion. This motion is mainly vibrational or rotational in solids and is dependent on temperature. It is predicted from statistical mechanics that a fraction of the atoms in the lattice will have values of energy that can be sufficient to enable the atoms to depart from their correct lattice positions.

If an atom moves into an adjacent interstitial position not belonging to the crystal lattice, a vacant lattice site is created, and the combination of vacancy and nearby interstitial atom, first described by Frenkel, is termed a Frenkel defect.

2.2.2.3 Schottky Defects

Defects of another type, named after Schottky who first discussed them, are often encountered in solids when the formation of interstitial atoms is energetically unfavorable. At the surface, an energetic atom may have insufficient energy to evaporate, but still may partially break away from its nearest neighbors without leaving the crystal surface. A vacancy is then formed that may diffuse into the interior of the crystal. Thus, a Schottky defect consists simply of a vacant lattice site.

2.2.2.4 Impurity Defects

An impurity defect results when a lattice site is occupied by an atom of an element other than the host element for that site. Apart from either adding or removing additional charge carriers in semiconductors, substitutional impurities distort the crystal lattice in their immediate neighborhoods and render the formation of other crystal defects energetically more favorable.

Downloaded from https://www.everyspec.com

2.2.2.5 Color Centers

Color centers are a type of defect found in insulating or semiconducting crystals. They usually appear in ionic crystals of binary compounds if the composition departs from stoichiometry. The best known examples occur in the alkali halides, which, when heated in the vapor of the corresponding metal or halogen, acquire a characteristic absorption band in the visible region of the spectrum. There is evidence that the defects responsible – known as color centers from their optical effects – consist either of electrons trapped at the sites of halogen vacancies or of holes trapped at the sites of metal vacancies. These vacancies are a result of an interaction between the heated halide and the surrounding vapor. If the vapor is of the alkali metal, halogen vacancies are created in the halide; conversely, if the vapor is the halogen, metal vacancies are formed in the halide. Another example of color center formation occurs in the semiconductor zinc oxide which loses oxygen on heating in vacuo and develops a yellow color associated with electrons trapped at oxygen vacancies. With other materials, the absorption maximum often is found in the infrared, and the presence of color centers only affect the electrical properties without affecting the color. Lead sulfide is an example; heating this in sulfur vapor induces p-type conductivity, the concentration of holes increasing with increasing vapor pressure of the sulfur. Conversely, on heating in a vacuum, sulfur is lost and n-type conductivity eventually results.

2.2.2.6 Dislocations

Dislocations are another type of imperfection found in crystals. The term signifies a particular type of disturbance in the crystal lattice due to the relative motion of crystal regions either during growth or as a consequence of plastic deformation.

Two distinct types of dislocations have been recognized – linear (or edge) dislocations and screw dislocations. An edge dislocation is a one-dimensional lattice defect (Figure 2.2). The most highly disturbed region in the slip plane is the center of the dislocation. The surrounding region, extending to the undisturbed part of the crystal, is termed the dislocation field. Edge dislocations and the slip vector are mutually perpendicular.

Screw dislocations (Figure 2.3), on the other hand, lie in a plane parallel to the slip vector. Slip, in the direction of the slip vector, occurs over the whole crystal thickness so that, instead of parallel atomic layers stacked one above the other, there is one continuous spiralling plane.

Examination of the different possibilities leads to the conclusion that dislocations must either form a closed circuit inside the crystal or terminate at the surface. They can then be detected by producing etch figures at their termination on the surface.

A dislocation is characterized by its Burgers vector, which is a measure of the gap in a complete circuit traversed in undistributed material surrounding the dislocation. This so-called Burgers circuit is conventionally traversed according to a right-hand screw rule. The Burgers vector is usually related to a translation vector of the lattice and coincides with the slip vector if the dislocation is caused by slip.

Dislocations may move in a crystal, either in the slip plane or at right angles to it. When a dislocation moves in the slip plane, the slipped region grows and plastic deformation results. When a dislocation moves at right angles to the slip plane, transport of

Downloaded from https://www.everyspec.com

matter, i.e., diffusion, occurs. The half-plane of atoms tends to complete itself by the addition of an interstitial atom or an atom from the next lattice site. The incomplete atomic plane may also shrink by loss of the edge atom to an interstitial site or to a vacancy, resulting in movement of the dislocation in the opposite direction.

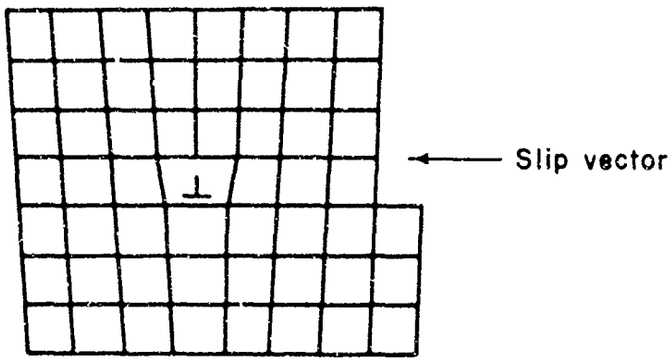


FIGURE 2.2. CROSS SECTION OF AN EDGE DISLOCATION

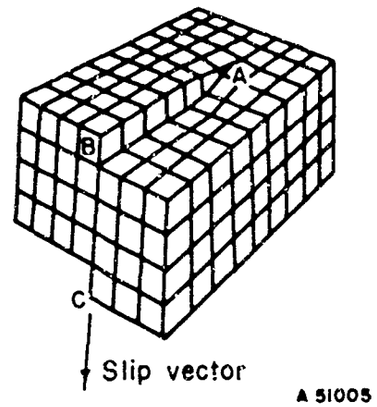


FIGURE 2.3. VIEW OF A SCREW DISLOCATION

The dislocation, one end of which is visible at A, is parallel to the line BC.

2.2.2.7 Twin Boundaries

There are two types of twin boundaries: "first order" or coherent twin boundaries and "higher order" or incoherent twin boundaries. In a coherent twin boundary, the common boundary plane is of (111) type, and atoms on either side of this boundary bear a mirror-image relation to each other. No dislocations, or broken or distorted bonds are required, and the associated energy necessary to create this defect is low – hence, its relatively frequent occurrence.

A frequent example of a higher-order twin boundary – termed a "second-order twin join" – is the interface between two grains that bear a nonparallel first-order twin relation to a third absent grain. The common plane is of (221) type.

2.2.2.8 Stacking Faults

A stacking fault is a disturbance of the regular order in which lattice planes are stacked upon each other in order to form a given crystalline lattice, which, in the case of silicon or germanium, is the diamond structure. The (111) crystallographic planes are stacked in an order that can be described as ABCABC... corresponding to the different atomic coordinates in each consecutive layer. If a plane should be missed, as in the case of ABCABABC, an "intrinsic" fault results. An extra plane, as in ABCABACABC, results

Downloaded from https://www.everyspec.com

in an "extrinsic" fault. Unlike edge dislocations referred to previously, which are one dimensional in extent, stacking faults extend in two dimensions. Stacking faults are bounded by partial dislocations. Partial dislocations have fault vectors that are not complete lattice vectors, contrary to the situation for normal dislocations. A special type of partial dislocation, termed a "stair-rod" dislocation, is formed when a fault bends from one (111) plane into another.

2.2.3 Surfaces

A crystal surface can properly be classified as a type of crystalline imperfection, since it is a boundary at which the regular three-dimensional array of the crystal lattice is terminated. Those atoms in the surface regions are, therefore, exposed to a somewhat different environment of neighboring atoms than those in the bulk of the crystal. Consequently, it is not surprising to find that, whereas in the undisturbed crystal lattice of the bulk of a semiconductor or insulator, values of electron energy between the valence and the conduction band are not allowed, the disturbance in the lattice at a surface introduces a number of allowed energy levels within the normally forbidden region. Impurity atoms and lattice vacancies on the surface will also produce additional localized energy levels in the forbidden zone similar to those produced by the same defects in the bulk of the crystal. Surface states can be related to different causes. They may result from the interruption in the periodic potential of the crystal lattice[2], from dangling bonds[3], from surface atoms, or from a change of crystal structure in the surface layer[4].

One might expect that the number of surface levels would be of the same order of magnitude as the number of atoms classifiable as surface atoms. According to the degree to which these possible surface levels are occupied, the surface acquires a trapped surface charge. Consequently, a space-charge region develops beneath the surface. The depth of this space-charge region will vary with crystal resistivity and the surface-charge density. The charge in the space-charge region must also be equal and opposite in sign to the charge on the surface.

Figure 2.4 illustrates four possible conditions for an n-type semiconductor as the charge on the surface is varied from positive to negative:

(a) An n-type accumulation layer is formed as a result of the attraction of electrons by the trapped surface charge.

(b) The "flat-band" condition corresponds to zero excess charge on the surface and no associated space charge.

(c) A depletion layer results from the repulsion of electrons by the trapped negative surface charge.

(d) An inversion layer results when, for large negative surface charges, a sufficient concentration of holes is attracted near to the surface region to invert the conductivity from n to p.

Corresponding diagrams can be drawn for p-type material.

Downloaded from https://www.everyspec.com

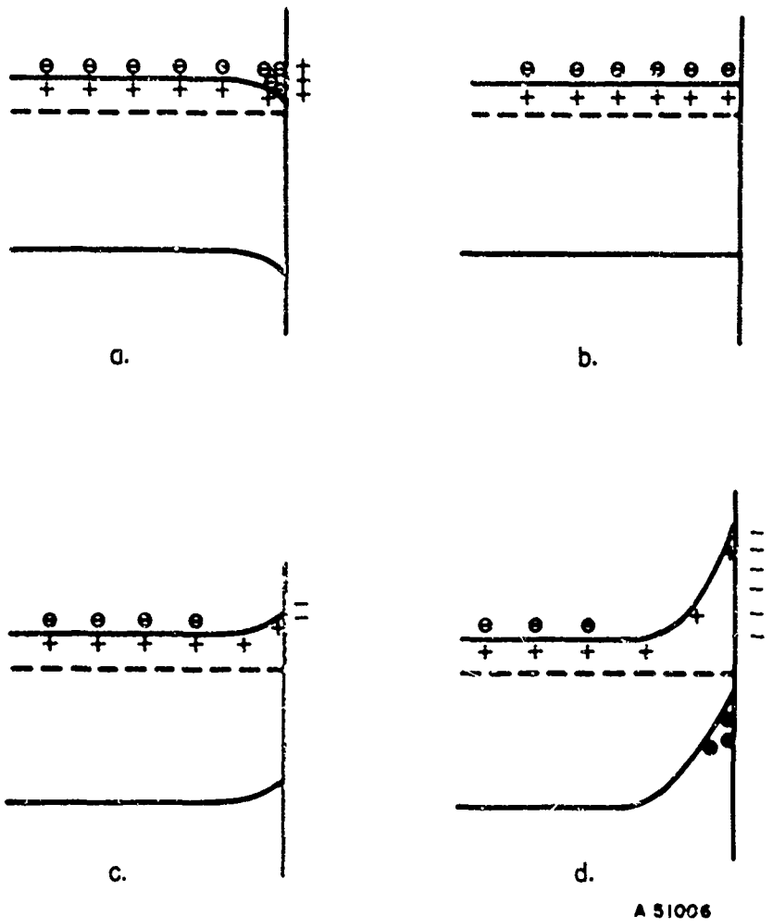


FIGURE 2.4. EFFECT OF SURFACE CHARGE ON THE ENERGY BAND OF AN n-TYPE SEMICONDUCTOR

These considerations have so far been applied to atomically clean semiconductor crystal surfaces in a perfect vacuum. However, a similar situation prevails when the crystal is located in a gaseous atmosphere or is in contact with a crystal of another material. Even at pressures in an ordinary high-vacuum system (10^{-4} to 10^{-6} torr), a monolayer of gases such as water vapor or oxygen will form on an exposed surface at room temperature in a matter of seconds. We must, therefore, consider the situation when the semiconductor crystal lattice is joined by a relatively discontinuous transition region to an alien crystal lattice of oxide, metal, or semiconductor.

In general, at equilibrium, a net flow of charge from one material to the other will have taken place. This results from the materials having different Fermi levels with respect to the standard zero point energy. In the material with the higher Fermi energy, a net loss of electrons occurs until its Fermi level becomes equal to that of the other material. The energy required to remove an electron from the Fermi level to a point of zero electrical potential is referred to as the work function of the material. This concept is usually encountered in connection with the thermionic emission properties of metals; it is well known that the work function can be changed by contamination of the surface, and advantage is taken of this fact when evaporating a monolayer of cesium metal on tungsten to reduce its work function.

Downloaded from https://www.everyspec.com

The influence of the relative work functions on the energy bands at the interface between a metal and semiconductor is shown in Figure 2.5.

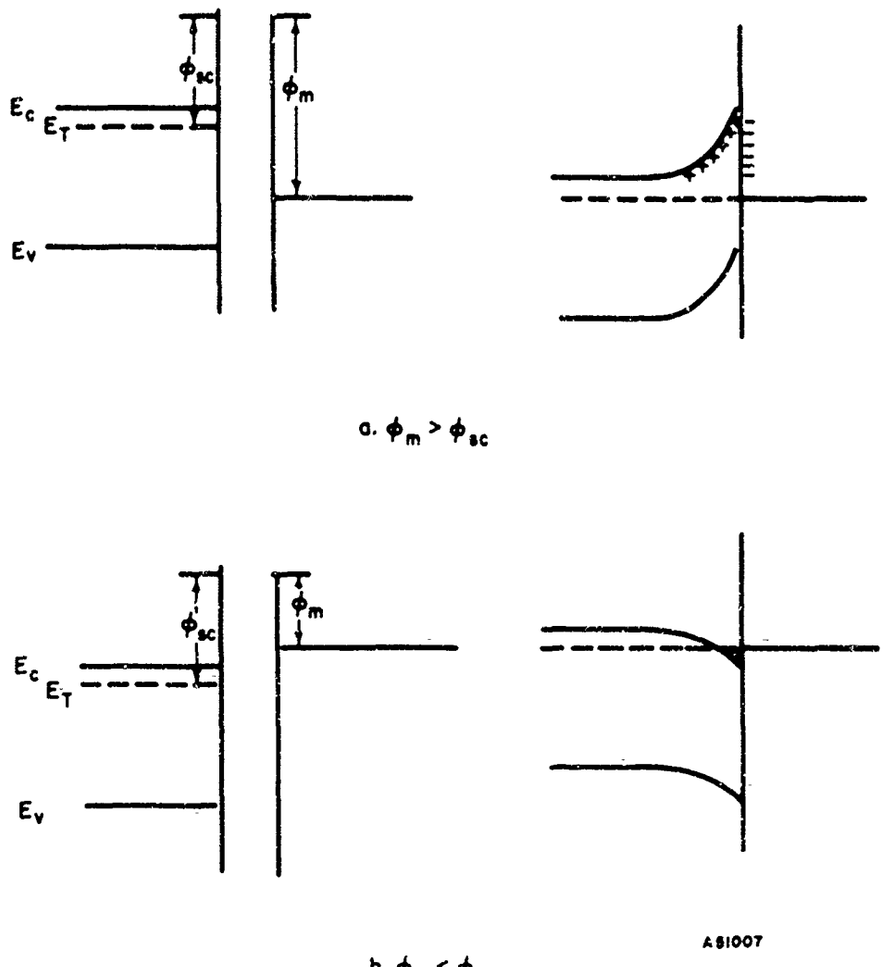


FIGURE 2.5. ENERGY-LEVEL DIAGRAM OF METAL-SEMICONDUCTOR INTERFACE

Because of the excessive density of conduction electrons in the metal compared with that in the semiconductor, the exchange of electrons has a negligible effect on the energy of the Fermi level in the metal. Thus, the Fermi level in the semiconductor assumes the energy of the Fermi level in the metal, and the energies of the conduction and valence bands are modified by the ionic charge near the interface resulting from the donation of electrons or holes to the metal.

It will be seen in the first case that the surface region of the semiconductor acquires a positive charge and in the second case, a negative charge. Similar situations are found when the semiconductor has a coating of oxide.

Downloaded from https://www.everyspec.com

Any disturbance in the equilibrium filling of the surface states, as already discussed, is removed within microseconds; these states are correspondingly known as fast states. A second group of states, termed slow states, has been observed on etched germanium and silicon surfaces. Their densities are larger than fast states, and associated relaxation times may vary from seconds to hours. These slow states seem to disappear when an oxide film is grown on either silicon or germanium. Their origin is not well understood.

2.2.4 Diffusion

The phenomenon of diffusion is a consequence of the kinetic nature of temperature, as a result of which any localized high concentration of a given species of material in contact with another tends to reduce the concentration value. Diffusion phenomena are governed by Fick's two laws. The first law states that the rate of flow per unit area per unit time is proportional to the concentration gradient and in the opposite direction:

$$\underline{J} = -D\nabla C \quad , \tag{2.11}$$

where

$\underline{J}$ = flux

D = diffusion constant

C = concentration.

In cgs units, the diffusion constant has the dimensions of $cm^2\ sec^{-1}$.

The above equation is not always valid; the relation between the flux and concentration gradient may depart from linearity. In that case, in order to use an equation of the form above, it is usual to regard D not as a true constant but dependent on concentration.

Fick's second law is readily derived from the first by considering the rate at which the concentration in a small elementary volume subject to a concentration gradient changes with time:

$$\frac{\partial C}{\partial t} = D\ \nabla^2 C = D\left(\frac{\partial^2 C}{\partial x^2} + \frac{\partial^2 C}{\partial y^2} + \frac{\partial^2 C}{\partial z^2}\right) \quad . \tag{2.12}$$

This assumes an isotropic diffusion medium. If the medium is not isotropic, one can write

$$\frac{\partial C}{\partial t} = D_x \frac{\partial^2 C}{\partial x^2} + D_y \frac{\partial^2 C}{\partial y^2} + D_z \frac{\partial^2 C}{\partial z^2} \tag{2.13}$$

Downloaded from https://www.everyspec.com

Then, after 100 seconds, this particular concentration level has reached a depth of 46 μ from the surface; after 10^4 seconds, 430 μ. Iron is a very rapid diffuser in silicon.

2.2.4.3 Diffusion Mechanisms

Diffusion of an impurity through a solid may occur as the result of a number of different mechanisms. It is important to appreciate that the value of the diffusion constant for a given matrix material is dependent on the mechanism of diffusion. The diffusing atoms may move through the lattice occupying interstitial sites between the host atoms, or they may replace the host atoms and move through the lattice by occupying vacant lattice sites. Diffusion constants are usually much higher for interstitial diffusion than for substitutional diffusion.

In certain cases, such as zinc in gallium arsenide, diffusion takes place both interstitially and substitutionally. The concentration of interstitial atoms depends on the concentration of substitutional atoms, and the ratio of interstitial to substitutional diffusion constants is very high. A mathematical model has been developed for the dual diffusion mechanism, which is in excellent agreement with experiment.[5] Figure 2.6 shows the experimental and theoretical diffusion curves of zinc into gallium arsenide at 1000°C.

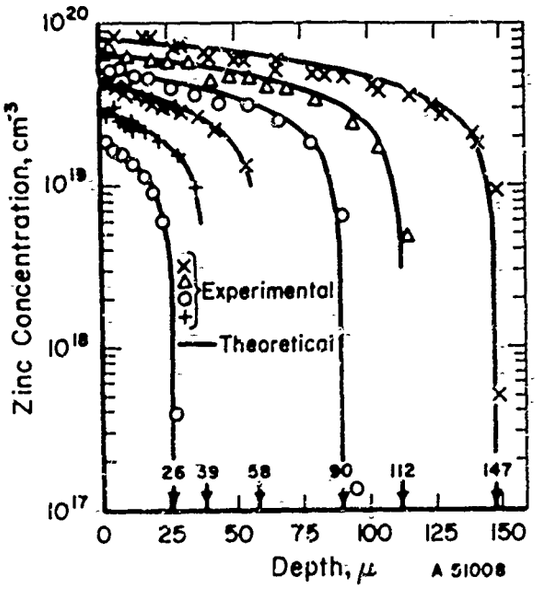


FIGURE 2.6. EXPERIMENTAL AND THEORETICAL VARIATION OF ZINC CONCENTRATION WITH DISTANCE AT 1000°C IN GaAs

Also, the diffusion mechanism will be affected in the crystalline region around dislocations in single-crystal solids and around grain boundaries in polycrystalline solids. This occurs as a consequence of the distortion of the crystal lattice in these regions, which, in the case of lattice expansion, will facilitate the passage of the diffusant atoms somewhat more readily. Thus, enhanced diffusion along an edge dislocation can destroy the uniformity of a desired diffusion front upon which the correct functioning of many semiconductor devices depends.

Downloaded from https://www.everyspec.com

Grain-boundary diffusion occurs more readily than lattice diffusion (e.g., the activation energies for grain-boundary and lattice diffusion of thorium in tungsten[6], for instance, are 90 and 120 kcal per mole, respectively).

Since atoms or ions can move through a solid lattice or along a grain boundary, it is not difficult to imagine more rapid migration along surfaces. Surface diffusion as a phenomenon is most frequently encountered in the condensation and aggregation of films deposited on solid surfaces. Consequently, it plays an important part in determining whether a thin film will be amorphous, polycrystalline, partially oriented, or fully oriented (epitaxial).

2.2.5 Ionization

Since ionization involves the removal of one or more valence electrons from an atom, it requires the availability of a certain minimum energy – the ionization energy. The temperature at which normal electronic components operate corresponds to kT values of less than 0.05 electron-volt. Although this energy is sufficient to ionize certain impurity atoms in the lattices of semiconductors, most gases require at least 1 electron-volt. Therefore, appreciable thermal ionization of the gas ambient in a semiconductor device, for instance, would not be expected to occur. However, the necessary energy for ionization can be supplied in other ways, such as by applying a sufficiently high electric field or by collision with energetic nuclear irradiation. Where the ionization rate is temperature dependent, it is an activated process, with the activation energy equal to the ionization energy.

2.3 Review of Failure Modes in Selected Electronic Parts

2.3.1 Semiconductor Devices

2.3.1.1 General Aspects of Semiconductor-Device Reliability

From the inception of the transistor, semiconductor devices have always appeared inherently capable of much greater reliability than vacuum tubes, since there are no obvious parts to wear out. There are no apparent "wear-out" modes in good semiconductors properly used. One could conclude that such items would survive until the equipment becomes obsolete. The level of reliability then would depend on the ability to develop adequate manufacturing controls, screening techniques, and care in application.

Semiconductor devices with the highest removal rate from field use are mixer crystals.[7] Field reporting of causes of failure or removal is limited to broad categories. A typical distribution for field removals is[7]:

Electric tolerance	25 per cent
Shorts	10 per cent
Opens	15 per cent
Mechanical	10 per cent
No apparent defect	40 per cent

Downloaded from https://www.everyspec.com

The achievement of high reliability in semiconductor devices starts at the design stage, involves all phases of manufacturing, and requires proper application by the user. Effective quality-control methods in manufacturing include adequate control of materials and processing, proper calibration and use of test equipment, and the use of failure-analysis techniques. Determination of the mechanism of a repetitious failure mode permits correction of the manufacturing process that causes or contributes to the failure. This procedure should eliminate the failure mode and result in a more reliable device.

Three aspects of a semiconductor device play a role in determining its reliability:

(1) Bulk-material properties, e.g., the effect of dislocations, twin boundaries or stacking faults in the starting material, uniformity of doping level, presence of strained material, spurious impurities, etc.

(2) Surface phenomena

(3) Defects of a mechanical or miscellaneous nature.[8]

The two latter aspects are illustrated in Tables 2.1[9] and 2.2, which are intended to give a general feeling for the type of problems commonly encountered in practice. Failure Modes I and J in Table 2.2 are not due to any lack of reliability in the device, but are a result of subjecting it to improper bias conditions imposed either through human error or by faulty equipment.

2.3.1.2 Review of Construction of Various Semiconductor Devices

In order to fully appreciate the modes of failure in semiconductor devices, it is necessary to have some knowledge of the chemical nature and geometrical configuration of the constituents of the devices. Therefore, a review is given below of the different types of semiconductor devices produced since 1954.

All transistors in use today are made from silicon or germanium. Gallium arsenide transistors have been made but are not available commercially. Silicon is preferred for most industrial and military applications, notably in computer switching. However, germanium transistors still have some use in industrial or military applications, especially when high-frequency operation is required or when design has been frozen.

Transistors may be conveniently classified according to the manner of formation of the emitter-base and base-collector p-n junctions. The junctions can be distinguished as surface barrier, grown, alloy, or diffused.

2.3.1.2.1 Surface-Barrier Transistors

The surface-barrier transistor consists of a chip of n-type germanium, into each side of which pits have been electrochemically etched. A metal, usually indium or cadmium, is then plated into each pit, as in Figure 2.7. Etching and plating are sequentially accomplished electrochemically by means of jets of electrolyte impinging simultaneously on each side of the chip, and are controlled by the polarity of bias applied between the electrolyte and the germanium. The thin germanium section, typically 4 μ thick, serves as a base region, while the plated dots function as emitter and collector.

Downloaded from https://www.everyspec.com

TABLE 2.1. FAILURE MODES FROM LIFE-TEST FAILURES OF MESA TRANSISTORS

External Problem	Internal Defect	Possible Production Fault
Open, no continuity	Lead wire-metal or metal-silicon bonds open	Poor alloying; improper surface cleaning
	Wire leads open	Electrical overload; work hardening of lead wire under vibration
	Wire off post	Poor weld
	Die off header	Improper cleaning of die or header; poor header plating
High I_{CBO} or I_{EBO}	Surface leakage	Contamination on junction; bad hermetic seal
Low BV_{CBO} or EB_{EBO}	Surface breakdown	Conducting particles across junction; improper junction cleaning
	Bulk leakage	Conduction through a flaw in the Si – usually a crack
	Presence of inversion layers	Unknown
C-B or E-B diode shorts	Bulk leakage	Electrical overload; surface or internal shorts
High $V_{CE}(sat)$ or $V_{BE}(sat)$	Degradation of h_{FE}	Migration of metal atoms; change in surface state
	Increase of emitter or base resistance	Bad metal surface, poor lead attachment

TABLE 2.2. FAILURE MODES DURING LIFE TEST AND OPERATION TYPICAL OF 2N696 AND 2N697 TRANSISTORS[a]

Failure Mode	Description of Failure	Failure Indicators	Types of Stress to Which Device is Most Susceptible	Intermediate Cause(s) of Failure
A	Channeling	High I_{CES} High I_{EBO} High I_{CBO}	Operating life; high-temperature storage; humid ambient	(1) Contaminant reaches surface of Si through porous oxide or flaw (2) Contaminant reaches surface of Si by solid-state diffusion (3) Ionic contaminant on SiO_2 surface induces inversion layer
B	Exposed or unprotected junction	High I_{CES} High I_{EBO} High I_{CBO}	Operating life; high-temperature storage; temperature cycling; humid ambient	(1) Mask misalignment causes removal of oxide over junction (2) Lifting of photoresistant mask during etching causes oxide removal (3) Chipping or cracked die exposes junction
C	Microplasma	High I_{CES} High I_{EBO} Low BV_{EBO}, BV_{CBO}	Operating life; high-temperature storage	Regions of high impurity concentration cause lower V_B than is typical of bulk
D	Beta drift	h_{FE} differs from previous reading	Operating life; high-temperature storage	(1) E-B junction has high reverse current (2) Base region surface has changed
E	Surface contamination	I_{CBO}, I_{EBO} versus V_{CB}, V_{EB} have ohmic slopes	Operating life; high-temperature storage; humid ambient	Charred organic material or metal oxides on the surface of SiO_2 form a conducting path between junction contact areas
F	Second breakdown	h_{FE} measures "O" on pulse test but is actually very high	Operating life; high-temperature storage	Local heating at points of extremely narrow base width
G	Formation of Au-Al-Si alloy	High V_{CE}	High-temperature storage; temperature cycling	Resistance of contact is increased by formation of Au-Al-Si alloy
H	Poor contact	High V_{CE}	Vibration; shock; centrifuge	Broken or poorly attached lead introduces base or emitter resistance
I	C-B, C-E short caused by gold spike through die	C-B short (or open base) C-E short (or open emitter)	Operating life; operation in improper or defective circuits; improper testing	(1) BV_{CB} exceeded due to insufficient current limiting (2) C-E short caused by overdriving (3) C-E short caused by exceeding LV_{CEO} in use
J	Emitter-base arc-over	B-E short (or base and/or emitter open)	Improper testing; operation in incorrectly designed circuit	(1) BV_{EBO} exceeded in testing (2) Certain operating conditions allow a runaway condition that results in destructive E-B current buildup

Downloaded from https://www.everyspec.com

TABLE 2.2. (Continued)

Failure Mode	Basic Physical Phenomenon Involved	Production Faults Causing Failure	Failure-Analysis Technique
A	(1) General contamination (2) Nonhermetic can weld (3) Oxide contamination (4) Silicon surface contamination	Improper housekeeping; reagent impurities; faulty can or header; improper adjustment of welder	Electrical examination of junction characteristics, microscopic examination of uncapped units; HF cleaning and oxide removal
B	(1) Stressing of die due to mismatch of thermal expansion coefficients of Si and Au-Si eutectic (2) Oxide mask did not cover junction (3) Die cracked by mechanical pressure	Improper mask alignment; damaged mask; improper removal of photoresistance, excessive pressure in die mounting; inadequate mechanical inspection	Microscopic examination of uncapped and etched unit
C	Contaminant diffuses through porous or nonuniform oxide at junction	Damaged masks (pinholes), incomplete KPR removal; KRP housekeeping; atmosphere control in diffusion area	Microscopic examination of uncapped unit with reverse bias applied
D	Same as for Modes A, B, and C	Same as for Modes A, B, and C	Electrical, microscopic examinations
E	Nonhermetic can weld allows contaminants to enter; organic contaminants char in sealed can when heated	Faulty parts (header, etc.); welder improperly adjusted, faulty final cleaning of transistor and can	Curve tracer observation of junction characteristics; leak check can; microscopic examination
F	Nonuniform diffusion profile	Reagent impurities; diffusion process control	Observation of I_{CE} versus V_{CE} curve
G	Intermetallic compound formation	Formed by high temperature	Microscopic examination of lead bond area
H	Faulty lead attachment	Inadequate inspection; welder maladjustment	Visual examination of opened unit
I	High local current flow causes heating to the extent that the gold lead alloys with the silicon to form a spike	Improper equipment design, improper testing; inadequate operator inspections	Electrical measurement, microscopic examination of opened unit; etching to expose spike
J	E-B junction is back biased to the point that a permanent high current path results	Same as for Mode I	Microscopic examination of arc scar

(a) Courtesy Raytheon Company, Components Division.

Downloaded from https://www.everyspec.com

Downloaded from https://www.everyspec.com

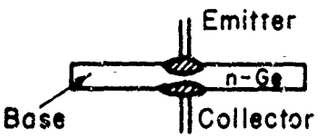


FIGURE 2.7. STRUCTURE OF SURFACE-BARRIER TRANSISTOR

2.3.1.2.2 Grown-Junction Transistors

Grown-junction transistors are normally in the shape of a bar with the transistor base in a transverse plane at the center, as in Figure 2.8. The cross section of the bar varies from 0.2 to 0.5 mm on a side. The bars are cut out of a single-crystal ingot into which correctly doped regions for emitter, base, and collector have been incorporated by a number of techniques, such as rate growing, melt back, double doping, and grown diffusion.

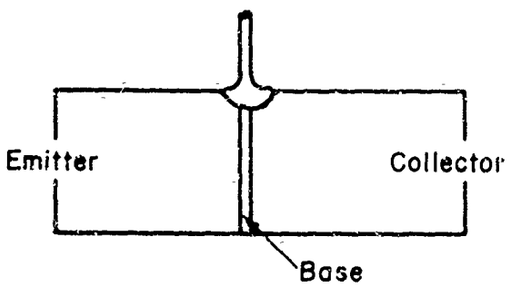


FIGURE 2.8. CROSS SECTION OF CRYSTAL BAR AT THE BASE CONTACT IN A GROWN-JUNCTION TRANSISTOR

Whereas the base widths of grown diffused devices are typically 2 to 3 μ, they might be 10 to 15μ for rate-grown, melt-back, and double-doped devices.

The main problem in all grown-junction devices is the contact to the base region; since the minimum diameter of a practical contact wire is about 25μ, there will be overlap onto emitter and collector regions. To prevent shorting, the composition of the wire must be such that p-n junctions are formed between the wire and the emitter and collector regions.

2.3.1.2.3 Alloy-Junction Transistors

Alloy-junction transistors are made by placing an appropriate metal in contact with the semiconductor base and heating both to the eutectic temperature of the metal-semiconductor alloy. A common example is that of indium on n-type germanium, as shown in Figure 2.9. The combination is heated to 500°C for perhaps 1 minute, and the liquid alloy formed contains about 10 per cent germanium. On cooling, the germanium, highly doped with indium, crystallizes out as a single-crystal layer. Thus, a junction with

Downloaded from https://www.everyspec.com

the original n-type base material results. Some indium diffusion may occur in the solid beyond the liquid-solid interface, depending on alloying time and temperature. However, the distance of diffusion is small in comparison with the alloy depth.

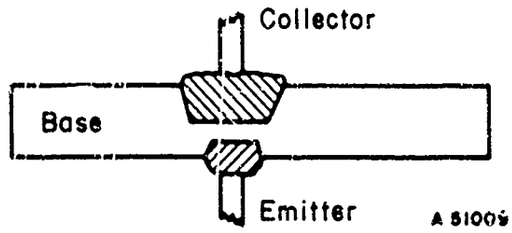


FIGURE 2.9. CONFIGURATION OF AN ALLOY-JUNCTION TRANSISTOR

Both emitter and collector junctions are formed in this manner, with the collector dot usually about twice the size of the emitter dot. Aluminum is commonly used to form an alloy junction with n-type silicon, and antimony is commonly used with both p-type silicon and p-type germanium.

2.3.1.2.4 Microalloy Transistors

Essentially a variation of the alloy-junction transistor, the microalloy transistor is made by forming a surface-barrier transistor and then alloying the emitter and collector dots to about 0.1 μ depth in the n-germanium base.

2.3.1.2.5 Surface-Alloy Transistors

In the surface-alloy transistor, n-type silicon is chemically etched to form locations for the emitter and collector. Aluminum is then vacuum deposited into the depressions and alloyed by appropriate heating.

2.3.1.2.6 Diffused-Junction Transistors

Graded junctions, which have better high-frequency characteristics than step junctions, can be obtained by the diffusion of selected impurity atoms into a semiconductor wafer. These impurities are introduced at the surface of the wafer, either from the vapor phase or from a solid containing the diffusant. Elevated temperatures and long diffusion times are needed.

In germanium, usual donor impurities are arsenic or antimony; acceptor diffusants are commonly gallium or indium. Donors diffuse more rapidly than acceptors in germanium. For silicon, the reverse is true; acceptors diffuse more rapidly than donors. Usual donors in silicon are phosphorus or antimony; acceptor impurities are boron or aluminum.

The substrate is normally the collector. The base is formed by diffusion of an appropriate impurity, and the emitter can then be formed by alloying. Alternatively, the

Downloaded from https://www.everyspec.com

emitter and base can be formed in one step by taking advantage of the different diffusion rates of donors and acceptors.

2.3.1.2.7 Double-Diffused Transistors

In these devices, a base is formed by diffusion into a collector substrate. Then the device is masked with oxide to reveal only the desired emitter region, and a second diffusion step is performed to form the emitter. However, while this process is relatively easy to carry out with silicon, since it has a stable oxide useful for masking against in-diffusing impurities, it is not used on germanium devices since germanium lacks such an oxide.

2.3.1.2.8 Triple-Diffused Planar Transistors

Triple-diffused transistors were designed to reduce collector series resistance and yet maintain high base-collector breakdown voltages. As the name implies, there are three different steps. First, the substrate is diffused so that the surface is degenerate, and a concentration gradient is set up inside the chip. Next, the degenerate region is lapped off from one side, and emitter and base junctions are diffused into the lapped surface, as shown in Figure 2.10.

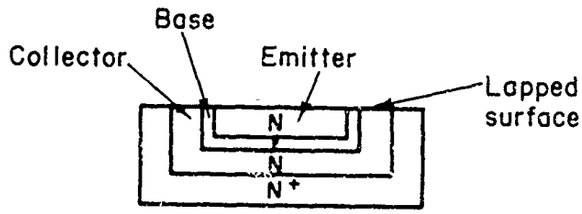


FIGURE 2.10. STRUCTURE OF A TRIPLE-DIFFUSED TRANSISTOR

2.3.1.2.9 Alloy-Diffused Transistors

Alloy-diffused transistors are formed by alloying a pellet of aluminum containing 1 per cent antimony and 2 per cent gallium into either n-type germanium or p-type silicon. The high-temperature portion of the process is maintained long enough to allow the faster-diffusing impurity to diffuse into the collector and form the base region. To permit contact to the base region, a thin surface layer must be diffused into it.

2.3.1.2.10 Drift and Microalloy-Diffused Transistors

The drift transistor and microalloy-diffused transistor are made by first diffusing from the emitter side of the chip, thus grading the bulk material. The emitter and collector are then formed by conventional alloying techniques, forming a p-n-i-p device. The impurity gradient across the base region results in a built-in electric field, which assists the transport of charge carriers across the base.

Downloaded from https://www.everyspec.com

2.3.1.2.11 Mesa Transistors

In the development of high-frequency and switching transistors, the time constant, R_BC_C, must be as small as possible. R_B is the base lead resistance. It can be kept small by control of the impurity level and concentration gradient in the base region. C_C is the capacitance of the collector-base junction and can be reduced by reducing the junction area. This is done in the mesa transistor by etching away the corners of a completed double-diffused or diffused-alloy transistor chip to form a mesa structure, as shown in Figure 2.11.

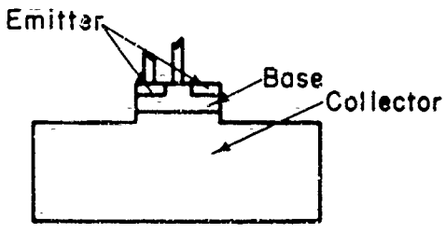


FIGURE 2.11. STRUCTURE OF MESA TRANSISTOR

2.3.1.2.12 Planar Transistors

In fabricating a planar transistor (see Figure 2.12), the silicon is first coated with a silicon dioxide film by high-temperature oxidation. Using photolithographic techniques for masking, windows are etched in the oxide by hydrofluoric acid. Boron and phosphorus are diffused through the windows to form the base and emitter junctions. Contact to the base and emitter regions is made with vacuum-evaporated aluminum. Since the intersection of the junctions with the surface occurs under the protective oxide coating, surface leakage currents are kept to very small values. Also, small values of collector-base junction capacitance for high-frequency performance are readily obtained by this technique.

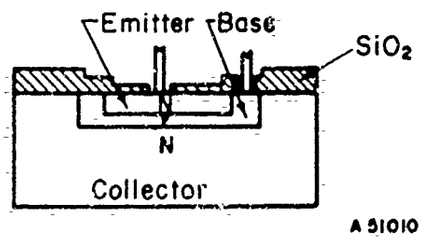


FIGURE 2.12. PLANAR TRANSISTOR SHOWING OXIDE PASSIVATION

Downloaded from https://www.everyspec.com

2.3.1.2.13 Epitaxial Transistors

If the collector junction is to have a high breakdown voltage, the collector must be of high-resistivity material. However, this causes the collector series resistance to be undesirably high. This problem was solved in epitaxial transistors by starting with a low-resistivity substrate and depositing a thin film of high-resistivity material onto the substrate by chemical vapor deposition. The deposited film is epitaxial, that is, it is single crystal and adopts the crystal orientation of the underlying substrate. Base and emitter regions are then diffused sequentially into the epitaxial layer by means of planar or mesa techniques. The triple-diffusion technique is an alternative method for reducing collector series resistance.

2.3.1.2.14 Unijunction Transistors

The unijunction transistor, also known as the double-base diode, is made from a single n-type semiconductor bar, usually silicon, with ohmic contacts to each end forming the two bases (see Figure 2.13). The emitter is formed by a single alloyed junction on one side of the semiconductor bar at some point along it. Current flowing from one base contact to the other induces a potential drop along the bar. The emitter junction may be either forward or reverse biased, depending on the relative emitter potential with respect to the potential of the bar in the region of the alloy.

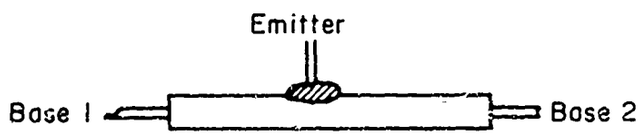


FIGURE 2.13. UNIJUNCTION-TRANSISTOR STRUCTURE

2.3.1.2.15 Field-Effect Transistors

The field-effect or unipolar transistor depends for its operation upon the modulation of the conductivity of a narrow channel by the depletion layers of two reverse-biased p-n junctions (see Figure 2.14). The width of the depletion layers is controlled by applying a bias to the junctions. Since the source-to-drain current does not flow in the depletion layers, the cross section of the channel available to mobile charge carriers can be greatly varied by biasing the p-n junctions. The channel can, in fact, be completely "pinched off" by application of sufficient bias so that the two depletion layers meet. The channels can be prepared either by alloying or by diffusion techniques.

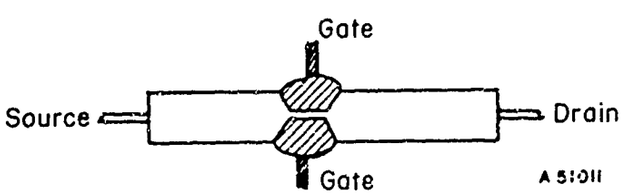


FIGURE 2.14. FIELD-EFFECT-TRANSISTOR STRUCTURE

Downloaded from https://www.everyspec.com

In the MOS (metal-oxide-semiconductor) device, channel modulation is effected from one side only by the bound charge of a charged dielectric film of thermally grown silicon dioxide, on the outer surface of which the gate electrode has been evaporated. This type of operation is described as a depletion mode of operation, since the application of bias serves to widen the depth of the depletion layer in the channel.

Another MOS device has been described that operates in what is termed an enhancement mode. In this case, source and drain are not connected by a channel when the gate bias is zero. Considering n-type source and drain regions in p-type material, interposed between these regions is the gate electrode on the surface oxide, as already described for the depletion-mode MOS. By applying positive bias to the gate, an n-inversion channel can be generated in the surface of the p-region; the gate is designed to overlap source and drain areas so that the generated channel extends from source to drain. The dimensions of the channel, and hence its resistance, are dependent on the gate bias applied. This mode of operation is referred to as the enhancement mode, since the gate bias serves to increase the conductance between source and drain rather than to reduce it.

Although enhancement mode devices using n-type material are theoretically possible, all such devices described to date have been constructed of p-type material.

2.3.1.2.16 Point-Contact Diodes

Historically, the first semiconductor devices, point-contact diodes, are now only used for microwave mixing purposes. N-type single crystals of both germanium and silicon are used. The point material may be tungsten, phosphor-bronze, or platinum-iridium. These diodes cannot withstand high power.

2.3.1.2.17 Gold-Bonded Diodes

The gold-bonded diode has replaced the point-contact diode in computer logic circuit applications. In the gold-bonded diode, a preformed whisker of gold is pressed against an n-type germanium chip and alloyed into the germanium by controlled pulses of current. The gold whisker is doped with 1 per cent gallium. These diodes are used in fast-switching applications.

2.3.1.2.18 Alloyed and Diffused Diodes

Practically all the techniques used with transistors also have been used with diodes. Thus, in alloy-junction diodes, an indium dot may be alloyed with an n-type germanium die or an aluminum dot with an n-type silicon die. Likewise, there are diffused diodes, mesa diodes, planar diodes, and epitaxial diodes.

2.3.1.2.19 Tunnel and Backward Diodes

Both tunnel and backward diodes are heavily doped, two-terminal devices. Backward diodes are used with tunnel diodes to impart to a circuit the unidirectional impedance characteristic that tunnel diodes themselves lack.

Downloaded from https://www.everyspec.com

In addition to being heavily doped, tunnel diodes must have an abrupt junction. This is normally accomplished by alloying. For example, using n^+ germanium doped with gallium, a dot of arsenic-doped lead is alloyed to the germanium to form the tunnel diode.

2.3.1.3 Influence of Crystal Imperfections on Semiconductor-Device Reliability

The subject of crystal imperfections has been discussed in Section 2.2.2. The implications of these imperfections for semiconductor-device reliability have been studied in the case of edge dislocations, twin boundaries, and stacking faults. The imperfections will be discussed in that order.

2.3.1.3.1 Edge Dislocations

An illustration of an edge dislocation in a primitive cubic lattice has been given in Figure 2.2. Inserted into an initially perfect crystal is an extra half-plane of atoms, which terminates at the dislocation. Silicon and germanium have the more complicated diamond structure, and the arrangement of atoms around the dislocation line has not been decided beyond doubt.

Since it is possible to obtain silicon crystals that are free of dislocations[9], it might appear that, for highly reliable devices, one should work exclusively with such material. Unfortunately, improved device parameters have not resulted from using dislocation-free starting material, since the processes of diffusion[10,11] and oxidation[12] introduce dislocations into the crystal. However, it has been found that the direct electrical effects of dislocations on device properties are surprisingly weak. Indirect effects, however, may be appreciable.

Although relatively high concentrations of diffusing impurities are necessary to cause "diffusion-induced" dislocations, such concentrations are commonly used in the production of semiconductor devices. Other processes, such as alloying, thermocompression bonding, and scribing, can also introduce damage and dislocations into the crystal lattice.

The fact that edge dislocations have negligible direct electrical effects has been noted from the properties of p-n junctions containing a small-angle grain boundary that traverses the junction region.[13,14] A small-angle grain boundary consists of a regular array of very closely spaced edge dislocations. No significant differences were found between junction characteristics of diodes containing grain boundaries and others from the same wafer. However, it has been established that dislocations reduce the lifetime of minority carriers.[13]

Edge dislocations attract impurity atoms and surround themselves with a concentration of these impurities. This results from the regions of dilatation and compression caused by the dislocation in its immediate neighborhood. Atoms larger than silicon then tend to concentrate in the regions of dilation, while smaller atoms are more easily accommodated in regions of compression. Impurity atoms also diffuse faster along edge dislocations than through an undisturbed lattice.[14,15] Consequently, a small-angle grain boundary will

Downloaded from https://www.everyspec.com

destroy the uniform front so often desired in diffusion processes. An example is shown Figure 2.15. In devices with narrow base regions, this could result in collector-to-emitter shorts if the "spike" extended through the base region.

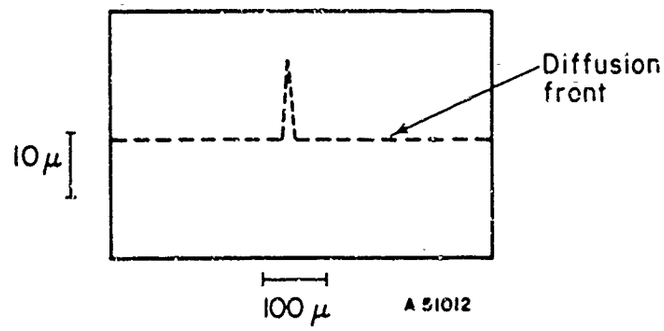


FIGURE 2.15. DIFFUSION ENHANCEMENT BY THE DISLOCATIONS OF A SMALL-ANGLE GRAIN BOUNDARY

The beveled and stained section shows a spike in the diffusion-front profile.[20]

In principle, isolated dislocations should have similar but much less pronounced effects than those in the dislocations at the grain boundary. However, even slight nonuniformity in doping can have serious consequences, especially in high-frequency units with base widths less than 1μ. Such nonuniformities in the impurity-concentration profile can lead to local concentrations of current through the base, resulting in "hot spots" that limit transistor operation.[16]

Apart from causing nonuniform concentration profiles, there is another serious consequence of impurity buildup around dislocations. Contaminants, especially heavy metals, are frequently present in semiconductors. So long as these are present in solid solution, they are comparatively innocuous. However, dislocations provide nuclei upon which the dissolved metals may precipitate. Should metal precipitates form in a p-n junction, the effect is immediately obvious.[14,17,18] The reverse current is increased, perhaps to the point where the device is useless. The breakdown characteristic also becomes soft.

Twin boundaries also provide favorable nucleation sites for precipitates and have been shown to be responsible for soft p-n junction characteristics.[14,19] Thus, the precipitation of dissolved impurities at crystal defects can lead to failure of initially good devices. Operating temperatures are always below those required for appreciable diffusion of standard donor or acceptor impurities, but fast-diffusing interstitial metals can migrate freely enough, especially along dislocations, to result in harmful concentration of precipitates during the device's lifetime. Precise numerical data on diffusion constants for these metals are lacking.

There is evidence that the concentration of dissolved oxygen can build up around a dislocation, giving rise to photovoltaic effects.[14,20] Changes in polarity and magnitude

Downloaded from https://www.everyspec.com

of the photoresponse indicate that the electrical nature of different boundary dislocations is variable. These effects parallel those found in crucible-grown silicon[22,23,24], which contained up to 10^{18} atoms of oxygen/cm^3. The oxygen existed as dissolved SiO_4^+ complexes or as precipitated SiO_2 aggregates.

Thus, a dislocation must be considered in combination with its local impurity environment in determining its over-all influence on device properties.

2.3.1.3.2 Twin Boundaries

Coherent twin boundaries exert no harmful influence where they intersect a p-n junction.[14,19] However, "softness" in junction characteristics due to precipitation of metallic impurities has resulted where an incoherent twin boundary intersected a junction. Such devices can be made hard by gettering with "glassy" oxide layers[17] when this softness is due to metallic precipitation.

No evidence of diffusion enhancement along (221) twin boundaries has been found.[14,19]

2.3.1.3.3 Stacking Faults in Epitaxial Layers

A specific type of lattice defect frequently observed in epitaxial silicon is the stacking fault. This defect is not commonly found in crystals grown from the melt.

The geometry of fault structures depends on the orientation of the substrate. On a (111) plane, the faults grow as tetrahedra with the apex at the substrate-film interface. Consequently, with increase in film thickness, this tetrahedron grows in size, a fact that has been used in determining film thickness.[14,25] The intersection of film surface and tetrahedron results in triangular figures. Ribbonlike figures are sometimes observed, and are considered to be stacking faults bounded by dislocations.

The total lattice mismatch at a stacking fault is similar to that of a coherent twin boundary. Since very little deleterious influence is found for these boundaries, it is to be expected that similar results hold for the stacking-fault planes. No enhancement of diffusion nor any detrimental effects in p-n junctions associated with stacking faults have been observed.[14,26,27,28]

The effects of the associated stair-rod dislocations on p-n junction properties are expected to be more pronounced. It has been observed[26,27,28] that, on application of reverse bias to a p-n junction grown in an epitaxial layer, preferential microplasma breakdown occurs at the corners of the triangles, i.e., at the stair-rod dislocations. This breakdown can be pinpointed because of the emission of visible light associated with microplasma conduction. However, the effect is not observed on all stair-rods, indicating that the dislocation alone does not cause the effect but that something else must be necessary, i.e., some precipitated species. P-n junctions fabricated in faulted epitaxial material often have soft reverse I-V characteristics. This constitutes a serious failure mechanism and has been correlated with the presence of preferential microplasma breakdown at the stair-rod dislocations.[14,26,27,28]

Downloaded from https://www.everyspec.com

2.3.1.3.4 Bulk Diffusion of Impurities

The diffusion constants of the normally used dopants in semiconductor devices are such that appreciable diffusion would not take place to any measurable extent at normal operating temperatures in the course of a century. However, it has already been described how certain metallic impurities can diffuse at appreciable rates to the regions around dislocations and perhaps eventually precipitate there.

A type of degradation which has been found in surface barrier transistors, has been attributed to normal solid state diffusion.[29] In freshly prepared devices, the junction is abrupt and, consequently, the junction capacitance varies with bias as $(V - V_D)^{-1/2}$, where V_D is the built-in potential difference. On aging, the C-V characteristics will assume a form $C \alpha (V - V_D)^{-n}$, where $1/2 > n > 1/3$ and n slowly approaches the value 1/3, the expected value for a linear-graded junction. In addition to causing changes in junction capacitance, the diffusion of metal impurities through the metal-semiconductor interface physically shifts the position of the junction (defined as the plane at which net ionized impurity concentration is zero) from its original position at the interface into the body of the semiconductor, resulting in a decrease in "punch-through" voltage. It was tentatively concluded that the impurity was indium from the plated collector electrode. The diffusion constant was measured as 10^{-17} cm^2/sec at 100°C and 10^{-21} cm^2/sec at 30°C.

2.3.1.3.5 Thermal Effects in High-Power Transistors and Diodes

When semiconductor devices dissipate appreciable quantities of power, it has been found that a number of bulk failure mechanisms occur that involve thermal factors. For instance, when power transistors exceed a certain critical operating temperature, current density and temperature tend to increase in a localized region of the device. The temperature of the "hot spot" may be greatly in excess of the value expected on the basis of uniform current distribution over the working area of the device. Local alloying or diffusion phenomena occur, leading to early failure.[14,30]

"Hot spots" are also believed to be responsible for the phenomenon known as "secondary breakdown", wherein the voltage between the collector and emitter decreases abruptly from its normal operating level to a much lower value.[31] Defects such as thin spots in base layers or imperfect alloying to the heat sink are normally responsible for thermal instability. However, base-collector breakdown was not always observed to occur at the site of an existing hot spot.[14,32,33] Possibly, local defects provide high densities of generation-recombination centers that initiate breakdown.

Lack of consistency in hot-spot development has been observed occasionally also; for instance, one might observe on successive pulses of current that the hot spots develop in different locations.[14,32,33]

Thermal "runaway" is of more importance in germanium than in silicon, since the intrinsic carrier density is greater in germanium than in silicon. Thermal runaway is caused by thermal-generated current flowing into the base region from the collector.

Downloaded from https://www.everyspec.com

Emitter current is consequently induced the same as if external base current were flowing. Dependent on circuit configuration, a condition can result in which an increase in junction temperature, and hence an increase in thermally generated current, results in an increase in emitter current and power dissipation that is more than sufficient to sustain the rise in junction temperature. The current then continues to increase with increasing junction temperature until the device is destroyed or the current is limited by an external load.

Problems of hot-spot development have been encountered in high-power four-layer diodes[34,35], but are minimized by using a large heat sink. A diagram is shown in Figure 2.16. Junction J_3 is shunted as a result of the geometry of the emitter grid structure. Junction J_3 is forward biased when the device is forward biased, and provides some emitter current to flow into the p-base region.

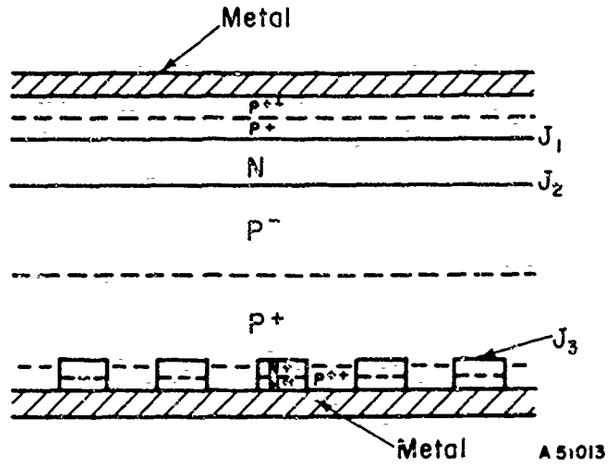


FIGURE 2.16. STRUCTURE OF FOUR-LAYER, HIGH-POWER DIODE WITH GRID EMITTER STRUCTURE

Failure as a result of nonuniform turn-on can occur when the diode is pulsed. Failure of the diode occurs when turn-on initiates at a local spot and does not spread over the device area before the current pulse builds up. Localized turn-on may be caused by:

(1) Local structural defects that lead to breakdown of the second junction, J_2, below the designated switching voltage

(2) An imperfection that has resulted in a nonuniform diffusion front, similar to a thin base spot in power transistors, thus causing a localized high current gain.

Failure occurs as the local temperature reaches that at which silicon-contact metal alloy melts, short-circuiting the device. This failure mode is minimized by the use of a grid structure for the emitter and by appropriate triggering of the current pulse.

Downloaded from https://www.everyspec.com

Thermal fatigue, leading to the formation of microcracks in the silicon, was found to result from the thermal strains generated and relieved due to the repeated passage of current through the device; the strains were dependent on the amount of mismatch of coefficients of expansion of the silicon and the contacting material. In some cases, the expansion-contraction sequence produced audible clicks. However, this failure mode could be almost completely avoided by the use of gold-germanium eutectic for soldering to the silicon.

2.3.1.3.6 Recombination Radiation as a Cause of Failure

Studies of gallium arsenide tunnel diodes show that early failure is induced by biasing above the valley voltage.[36-44] Investigation of the phenomenon has led to the conclusion that the prime failure mechanism is one whereby crystal defects are formed by the removal of host and/or impurity atoms from normal lattice sites. This process is entirely dependent on whether the device is operating above the recombination radiation threshold, which suggests that the photons produced by recombination are a causative factor in the failure mechanism[35,36,39,41,42], at least in gallium arsenide. The rate of deterioration increases with the output of recombination radiation.

The temperature of the device is only important in a minor way in that the threshold voltage increases with temperature. Thus, one could operate a device at a constant current just above the threshold voltage. On raising the temperature, with the current held constant, degradation ceases as soon as the applied voltage falls below the threshold value. The failure manifests itself by an increase of valley current and a decrease in junction capacitance.

The situation in regard to indirect band-gap semiconductors, such as gallium antimonide, silicon, or germanium, is not yet clear. Certainly, tunnel diodes in these materials degrade at rates many orders of magnitude less than those in gallium arsenide.[36,37,42,43]

2.3.1.4 The Influence of Surface Effects on Semiconductor-Device Reliability

The effects of water vapor and other surface contamination on the properties of semiconductor devices are now well documented, if not thoroughly understood. There are a number of possible surface effects; one is a straightforward effect in which a layer of contaminant is sufficiently conducting to provide a current leakage path across a reverse-biased p-n junction. Alternatively, moisture or oxygen in the ambient or surface contaminants can induce changes in the electrical properties of the surface. These changes may correspond to a change in the resistivity of the material or even inversion to a conductivity of the opposite type. If these changes take place at the intersection of a p-n junction with the surface, the net effect is to alter the shape of the junction at the surface, which may result in increased leakage currents and junction capacity. In extreme cases, the base of a transistor structure may be stimulated by an inversion layer, thereby providing a junctionless current path from emitter to collector.

When these effects were first encountered in germanium devices, surfaces were stabilized by using varnishes, silicone rubber or grease, etc., as well as special cleanup

Downloaded from https://www.everyspec.com

etches. This was the case with silicon also until advantage was taken of the thermally grown oxide to passivate the surface. Although surface passivation considerably improved reverse leakage current, water vapor and contaminants on the outer surface of the passivating oxide still influence device parameters.

While absence of chemical contaminants depends upon careful handling in processing, the exclusion of moisture requires, in addition, adequate hermetic sealing of the device container. Although the initial moisture content of the ambient gas might be acceptably low, it is possible, especially when the unit is operated at elevated temperatures, for sufficient moisture to be generated by outgassing from the inner canister surfaces to cause deterioration of device characteristics. Therefore, special moisture getters, usually of Vycor, are often incorporated into the enclosure to maintain a low moisture level.

Furthermore, at the extremely low temperatures normally required in military specifications, the internal dew point of the ambient inside the can may be reached. The saturation vapor pressure of ice at -55°C is 21×10^{-3} millibars; the pressure of the ambient may be 840 millibars. Hence, for moisture concentrations in excess of 25 ppm, the internal dew point will be reached, with undesirable effects on device performance. Gettering should enable concentrations below 1 to 2 ppm to be obtained.

There has been considerable study recently[34,45-54] concerning the effects of charged contaminant species under the influence of the fringing electric fields associated with a reverse-biased p-n junction. The accumulation of surface charge and the fringing field induce an opposite charge in the silicon surface, thereby causing resistivity changes and, in extreme cases, inversion in the surface layer of silicon.

Perhaps the best-known example involving the deleterious effects of accumulated surface charge was that of Telstar I where, in the space-radiation environment, failure of certain silicon transistors resulted.[55] It was quickly established that failure was due to ionization of the gas in the canister. Under the presence of the fringing electrical field, the ions were attracted to the edges of the depletion layers. However, contamination seemed to play a role also in that, when the effect was simulated in the laboratory using gamma rays as an ionizing source, not every transistor was affected. It was postulated that significant changes in transistor parameters would result only when contaminated species were present to retain the charges on the surface. In the absence of a radiation environment, charge accumulation has still been observed under simultaneous application of elevated temperature and applied reverse bias across the p-n junction.

The net result of the accelerated aging of planar devices is an increase in the reverse current of the collector base junction.[53] This current may not saturate until voltages not far from avalanche breakdown are reached. However, the "failed" devices recover their initially good characteristics, either by opening the can and exposing it to the atmospheric ambient or by heating it to 200°C for a few minutes. No bias of course, must be applied.

Positive ions move in the direction of the fringing field and pile up above the edge of the junction depletion region. Likewise, the negative charges are piled up above the opposite boundary of the depletion region. Piled-up positive charges over the p-region induce an increased concentration of electrons in the silicon underneath. This silicon,

Downloaded from https://www.everyspec.com

originally p-type of some desired acceptor concentration, now has in this surface region more electrons than before. The result is a decrease in the net effective acceptor concentration in the surface region (Figure 2.17). In the extreme case, if sufficient positive ions have piled up to induce more electrons in the underlying silicon than the original number of acceptor atoms, the region becomes n-type. Similar considerations apply in the case of negative charges piled up over the n-region depletion-layer interface.

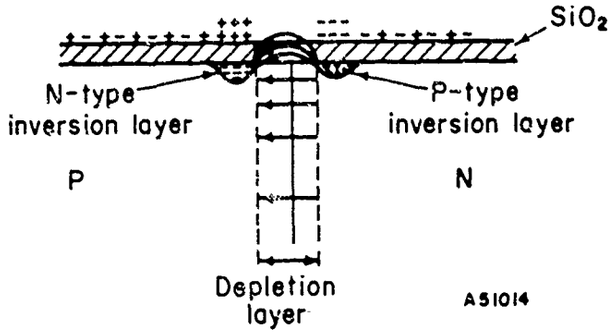


FIGURE 2.17. SURFACE-CHARGE SEPARATION ON A PASSIVATING LAYER UNDER THE INFLUENCE OF THE FRINGING FIELD OF A P-N JUNCTION

The parameters affected by these surface charges are leakage current, surface recombination velocity, beta, capacitance, and breakdown voltage. These effects are reversible. The ions may be either redistributed or driven away by heating the affected devices in the absence of an applied reverse bias. It has been calculated that 10^{-3} monolayer coverage of the passivating layer by ions is sufficient to invert 1 ohm-cm n-type silicon to a depth of 100 Å.(47)

Since the moisture content of the ambient atmosphere increases the mobility of the ions on the surface, processing procedures that result in a minimum of moisture inside the can should produce devices most resistant to this effect.

Apart from the presence of surface contamination, the presence of an oxide on silicon tends to induce an n-type charge under the surface.(56,57) Thus, there is a tendency for inversion layers to exist on very-high-resistivity p-type surfaces that have been oxidized. The silicon dioxide aids the formation of inversion layers on p-type surfaces and tends to inhibit the formation of inversion layers on n-type surfaces.

Silicon dioxide passivating layers are apparently not the homogeneous, amorphous layers they were formerly considered to be. There are grounds for believing that they possess a built-in space charge associated with oxygen vacancies and that a semiconducting layer exists at about 1000 Å from the silicon-silicon oxide interface.(49) It has also been demonstrated by deliberate contamination that sodium and lithium ions will migrate through the oxide under the influence of fringing electric fields at p-n junctions.(58) Sodium piles up at the silicon oxide interface, while the lithium diffuses further into the silicon. The diffusion of sodium or lithium through the oxide can be prevented by converting the outer layer of the oxide into a phosphorus-rich glass.(59)

Downloaded from https://www.everyspec.com

A similar effect was found when other glasses were used to passivate p-n junctions.[60] Only in the instance where alkalifree glasses were used was channeling avoided. In the case of planar devices, it was found that contaminant effects could be eliminated by rinsing the device in a polar solvent, such as a mineral acid[52]. Nonpolar solvents such as xylene or acetone were ineffective. The contamination was suspected in certain cases to be compounds of phosphorus released from the electroless nickel or gold plate on the cap or header.[52]

Another successful method[53] of avoiding surface effects with planar transistors was:

(1) Rinse in nitric acid (hot, then cold with ultrasonic agitation)

(2) Rinse in deionized water (hot, then cold with ultrasonic rinse)

(3) Bake in a dustfree dry ambient.

This procedure was repeated:

(1) Before aluminum evaporation

(2) Before wafer scribing

(3) After die bonding

(4) After wire bonding.

It seems safe to say that surface contamination remains the major limitation on semiconductor reliability today.

2.3.1.5 Mechanical Defects

The majority of the device defects that come under the heading of mechanical defects really belong to the area of production control and need only brief descriptions. These are given below.

2.3.1.5.1 Presence of Foreign Particles Inside the Container

Approximately 60 per cent of all transistor and diode designs have suffered from loose, internal, metallic-particle problems[61]. These may result from weld or solder splatter, peeling of plated coatings, tool damage, or use of gold epoxy. If the particles are conductive and are attracted by the fields across a p-n junction, they increase the leakage current of the junction, sometimes catastrophically.

Downloaded from https://www.everyspec.com

2.3.1.5.2 Container Imperfections

These comprise leaking weld seams and leaking metal-to-glass seals around the terminal; moisture from the outside air can enter through these openings and change device characteristics. Detergent and similar "bombing" techniques for gross leak detection have been generally ineffective, frequently leading to latent failure.(7) There have been instances also of poorly insulating terminals contributing can leakage currents to p-n junction leakage currents.

2.3.1.5.3 Imperfect Wire-Lead Bonds

Assuming that an open circuit does not immediately result, imperfect wire-lead bonds can contribute a large series resistance to the emitter, base, or collector terminals of a transistor. Approximately 50 per cent of all assembly failures and 30 per cent of all component failures have been due to interconnection or bonding difficulties.(61)

2.3.1.5.4 Imperfect Bonding of Die

The die is usually bonded to the header, either by soldering or by forming a eutectic at the interface between a gold-plated header and the silicon chip. If the surfaces to be joined are not sufficiently clean, then voids may result between the die and the header. Good electrical and mechanical contact may still result. However, voids are particularly undesirable when high power must be dissipated by the device and reliance is placed on a heat sink to prevent overheating. The presence of voids at the die-header interface interrupts heat flow and leads to hot spots in the die, often with catastrophic results.

2.3.1.5.5 "Purple Plague"

This descriptive term refers to a contact problem that develops between gold wire leads and aluminum metallization. It is now established(62,63) that this involves the growth of a brown or dark alloy of gold, aluminum, and silicon, referred to as the "black death", which is quite brittle, and hence mechanically weak, and which increases the electrical resistance of the leads. The term "purple plague" arose because the intermetallic compound $AuAl_2$ is purple. In apparently successful attempts to solve the problem of the "black death" alloy by using aluminum wires, it was observed that purple $AuAl_2$ developed where the aluminum wires were bonded to the gold-plated Kovar terminals, but nevertheless without any objectionable effects. The bond remained mechanically strong, and no significant increase in electrical contact resistance resulted.

The growth rate, both of "purple plague" and "black death", is only appreciable at temperatures above 100°C.

2.3.1.5.6 Tool Damage

It has been observed, especially in semiconductor functional blocks, that tool damage frequently occurs as a result of operator handling.(64) The usual sympton is a scratch

Downloaded from https://www.everyspec.com

across an evaporated interconnection, leading to a complete break in severe cases. This break may only develop after some time in operation, probably due to development of a hot spot.

2.3.2 Failure Mechanisms at Conductor-Insulator Interfaces

An electronic component often contains different materials in contact, such as the dielectric and electrodes in a capacitor where the change in composition at the interface is abrupt. Since the component was designed on this basis, any change in the situation is usually detrimental. An abrupt junction between two dissimilar materials is often thermodynamically unstable. The proper provocation, such as elevated temperatures or high electric fields, can serve as the activating mechanism for relief of the metastable state. This can result in interdiffusion of the two materials or chemical reactions between them. In particular, if the insulator is chemically reduced to a more conducting composition, its dielectric properties are degraded and failure results.

Examples of systems important to microelectronic applications are:

(1) $Ta\text{-}Ta_2O_5$

(2) $Si\text{-}SiO_2$

(3) Metal-SiO

(4) Metal-aluminum silicate and borosilicate glasses deposited by chemical vapor deposition.

It has been pointed out that, at temperatures high enough for the dielectric to have sufficient conductivity, a metal-dielectric-metal sandwich may function as a galvanic cell and therefore develop an internal EMF.[65] Chemical reactions occurring at the interfaces supply the necessary energy. However, further consideration of this phenomenon will be excluded, as it is considered unlikely that components will be stressed at sufficiently high temperatures for this to happen in situations of interest here.

A usual feature of all thin-film capacitor behavior is the presence of flicker in the d-c leakage-current values. This is observed particularly in anodic oxide films but is also evident to some extent in other thin-film dielectrics. It appears that the momentary high value of leakage current corresponds to instantaneous breakdown of a submicroscopic area, but normally the breakdown is self-healing in either one of two ways, at least if the energy available for dissipation is limited by a resistor in series with the capacitor. The first self-healing process to be considered occurs in solid- or liquid-electrolytic capacitors. During the current surge, oxygen ions flow into the breakdown site either from the liquid electrolyte or from the semiconducting manganese dioxide coating in solid electrolytic capacitors. This oxygen ion current oxidizes the underlying anode, and the resulting growth of oxide heals the breakdown. However, this type of healing is not possible where the cathode is nonoxidizing in nature, e.g., metallic. Then, if this is a very thin metal film, the current flowing through the breakdown site can be sufficient to vaporize locally the electrode and thus electrically isolate the breakdown site, resulting in apparent healing. However, a decrease in capacitance results as a consequence of a decrease in electrode area. This phenomenon on thin-film capacitors manifests itself by occasional transient sparking over the surface of the counterelectrode.

Downloaded from https://www.everyspec.com

It seems evident that, at the site of a local imperfection in a thin insulating film, increased leakage current would occur. In the aging of a thin-film dielectric, points of high leakage current might be spontaneously generated, resulting in local heating. Since conductivity in insulators increases with temperature, the current will then tend to increase and the voltage to drop, depending on the value of a series resistance or the output impedance of a power supply. If the voltage does not drop sufficiently to terminate the increase in power dissipation, then local heating is further increased and "thermal runaway" occurs, resulting usually in catastrophic device failure. It is clear then that, in this situation, values for device reliability have significance only when information is available on the effective value of the series resistance.

The presence of humidity in the ambient influences the electrical properties of $Ta\text{-}Ta_2O_5$ capacitors and $Al\text{-}SiO_2$ capacitors. SiO capacitors were affected only in an increase of breakdown events, but there was no apparent effect on capacity or dissipation factor.[66] The $Ta\text{-}Ta_2O_5$ system is normally rectifying; but it has been found that, on baking a tantalum capacitor at 400°C in vacuo, the rectifying properties disappear. However, they gradually reappear at a rate dependent on the porosity of the counterelectrode material (metal or MnO_2) on exposure of the unit to the normal humid atmospheric ambient. The crystal structure of these films is amorphous. However, it has been found that local crystallization increases leakage current. Crystallization has been observed to occur under the influence of electric fields and also around defects and impurity sites on the underlying metal.[67]

One aging mechanism in the $Ta\text{-}Ta_2O_5$ system has been identified[68-70], i.e., that of tantalum atoms across the oxide film under the influence of either temperature or the electric field. It has been established by radioactive tracer methods that a concentration gradient of tantalum exists in anodized Ta_2O_5 films. Starting from the $Ta\text{-}Ta_2O_5$ interface, the tantalum concentration decreased more or less exponentially to reach a substantially constant value about 600 Å into the oxide. This situation resulted in an approximately linear decrease of capacity against log frequency. After aging at higher temperatures and also under bias, the shape of the tantalum concentration profile changed. This has not been correlated with capacitance or dissipation-factor changes.

The capacity is normally bias independent. However, after heating to 400°C to 450°C for a short time, as is normal in the routine application of MnO_2 counterelectrodes, the capacitance becomes bias sensitive, decreasing with increasing bias.[71-73] This suggests that changes have occurred in the dielectric that were induced by high temperatures. Deposition parameters also can influence the electrical characteristics of thin-film capacitors, particularly in the treatment between deposition of the dielectric and final deposition of the counterelectrode. If the dielectric is exposed to a humid ambient before the final deposition of counterelectrode, capacitance values and dissipation factors are affected, capacitance values being lower and dissipation factors higher.

Silver migration on the dielectric surface was observed when silver was used as a counterelectrode material. Other failure modes are related to strain and other mechanical aspects of the dielectric film. Strain depends on deposition rate, substrate temperature during deposition, thickness, coefficient of expansion, and adhesive and cohesive properties. If adhesion is poor, peeling may occur above a certain film thickness. If cohesion is poor, the film may rupture or craze.

Downloaded from https://www.everyspec.com

2.3.3 Failure Mechanisms in Thin-Film Resistors

2.3.3.1 Background

In a thin-film resistor, a resistive path between two terminal points is provided by a thin film of resistive material deposited on an inert substrate that serves as a mechanical support. The thickness of the deposit will normally be less than 1 micron.

Common methods for depositing films onto a substrate are by evaporation in a vacuum (as in the case of nickel-chromium alloys); by sputtering (for example, tantalum); reactive sputtering (for example, tantalum nitride); or chemical decomposition (for example, the decomposition of hydrocarbon vapors onto a hot substrate to form pyrolytic carbon films or of tin (stannic) chloride to form tin oxide films).

In the case of a very pure annealed metal, the resistance decreases linearly with decreasing temperature, approaching zero as the temperature approaches absolute zero. Electrical resistance is caused by the scattering of conduction electrons by collision with the crystal lattice. The probability of scattering by the lattice vibrations is proportional to the mean square of the amplitude $(\bar{X})^2$ of the vibrations.

$$(\bar{X})^2 = \frac{h^2 T}{4\pi^2 km\theta^2} \quad , \tag{2.21}$$

where

h = Planck's constant

k = Boltzmann's constant

m = atomic mass

T = absolute temperature

θ = Debye characteristic temperature.

Taking resistivity, ρ_t, as proportional to $(\bar{X})^2$,

$$\rho_t \propto \frac{T}{m\theta^2} \quad . \tag{2.22}$$

In practice, exceptions are found to this behavior, since real metals are seldom of sufficient purity to approach zero resistance. All dilute solid solutions approach a limiting residual value of resistance due to additional scattering caused by solute-induced perturbations in the periodicity of the field of the metal lattice. Any lattice defect that produces a perturbation in the periodicity of the electric field of the lattice makes a

Downloaded from https://www.everyspec.com

contribution to the resistivity. Such defects include solute atoms, vacancies, dislocations, local mechanical fields, and stress. Thus, the resistivity found in practice has two components:

$$\rho = \rho_t + \rho_d \quad , \tag{2.23}$$

where

ρ_t = temperature contribution to resistivity

ρ_d = defect contribution to resistivity.

To the extent that defect concentration depends on temperature, the ρ_d term will be temperature dependent. The ρ_t term depends on temperature according to Equation (2.22). Since, for dilute solid solutions, the ρ_d term is largely independent of temperature, the temperature coefficient of resistivity (TCR) depends primarily on ρ_t:

$$\text{TCR} = \frac{1}{\rho}\frac{d\rho}{dT} = \frac{1}{\rho}\frac{d\rho_t}{dT} = \alpha \quad . \tag{2.24}$$

Therefore,

$$\alpha\rho = \frac{d\rho_t}{dT} \quad . \tag{2.25}$$

Thus, the product $\alpha\rho$ is a constant for dilute solid solution of a given parent metal. Equation (2.25) is found to be valid over a wide range of solid solutions.

Since high resistivity and low TCR are generally associated with one another, and since alloying increases the resistivity of a pure metal, metal resistors are usually made of alloys. One consequence of using alloy resistors is that the ρ_d term is strongly dependent upon composition; the addition of a few atomic per cent of a solute may double the resistivity. Also, if the solute concentration of an alloy decreases due to (1) precipitation of a solute-rich phase, (2) selective oxidation of the solute, or (3) selective evaporation, large decreases in resistivity (and corresponding increases in TCR) can be expected.

2.3.3.2 Review of Failure Possibilities in Thin-Film Resistors

Table 2.3 lists the pertinent metallurgical changes that influence the electrical resistance of resistive alloys and may, therefore, represent causative factors in the unreliability of thin-film resistors.[74] Each type of metallurgical process is discussed in detail below.

2.3.3.2.1 Oxidation

General aspects of the oxidation of metals, alloys, and semiconductors have been discussed in Section 2.1.1.1. Since oxidation of metals or alloys usually proceeds with the

Downloaded from https://www.everyspec.com

TABLE 2.3. SUMMARY OF METALLURGICAL CHANGES THAT INFLUENCE ELECTRICAL RESISTANCE[a]

Process	Effect on Resistivity	Effect on Temperature Coefficient	Effect on Component Resistance	Magnitude of Effect
Uniform oxidation	0	0	i	No limit
Selective oxidation of solute	d	i	d	May be very large
Uniform evaporation	0	0	i	No limit
Selective evaporation of solute	d	i	d	May be very large
Agglomeration	0	0	i	No limit
Precipitation of solute	d	i	d	May be very large
Long-range ordering of solute[b]	d	i	d	Very large
Disordering of solid solution	i	d	i	May be very large
Homogenization[c]	i or d	i or d	i or d	May be very large
Stress relief	d	i	d	< few pph
Mechanical stressing	i	d	i	< few pph
Vacancy condensation	Usually d	Usually i	~0	< few ppm
Clustering of solute[d]	Usually d	Usualiy d	i	Normally <1 pph
Solution of impurities[e]	i	d	i	May be very large

(a) i = increases; d = decreases; pph = parts per hundred; ppm = parts per million.
(b) Very rare - only a few specific alloys are known to produce long-range order and, because the order-disorder transformation is reversible, these would never be used for resistors.
(c) For example, in an evaporated film with a composition gradient through the thickness.
(d) For example, the formation of Guinier-Preston zones in Al-Cu alloys.
(e) Certain metals such as Ti, Zr, V, Nb, and Ta can dissolve appreciable oxygen. This portion of the oxidation process then corresponds to contamination by dissolution of impurities.

Downloaded from https://www.everyspec.com

growth of an insulating surface film, the primary result of oxidation is a decrease in cross section and consequent increase of resistance. Resistivity and temperature coefficient of resistance are not ordinarily affected unless the oxidation is selective (or unless a composition gradient exists through the film thickness).

Two other effects may be encountered when metal-film resistors suffer oxidation. The first follows from the fact that, if the metal-film thickness is sufficiently reduced by surface oxidation so that it approaches values comparable to the electronic mean free path, surface scattering becomes important. The effective resistivity becomes:

$$\rho = \rho_o \left(1 + \frac{3\ell}{8d}\right) , \tag{2.26}$$

where

ρ_o = bulk resistivity

ℓ = electronic mean free path

d = film thickness.

The second effect results from so-called internal oxidation. For example, in alloys of copper alloyed with a more reactive metal, underneath the surface oxide is a layer rich in copper (with oxygen in solution) containing small particles of the oxides of the alloying elements. The net effect is usually to increase the resistivity of the alloy film.

2.3.3.2.2 Selective Oxidation or Evaporation of Solute

These two processes similarly effect resistivity changes, since the remaining alloy contains less of the solute atoms. The changes obey the kinetics of the rate-limiting process, which is ordinarily the diffusion of an oxygen ion or of the atomic species that is oxidized or vaporized. Thus, diffusion-rate behavior is usually observed.

2.3.3.2.3 Agglomeration

Agglomeration is a phenomenon resulting from surface-tension forces in metal films on substrates that they do not readily wet. In extreme cases, the agglomeration becomes complete and the film becomes an array of unconnected islands. Such a film is, of course, electrically open. Gold has been observed to agglomerate fully on ground quartz but remains as a uniform film on alumina to about 400°C. The agglomeration apparently occurs by surface diffusion, obeying the same rate laws but at rates frequently 1000 times faster than bulk diffusion of the same metal.

2.3.3.2.4 Precipitation of Solute

Precipitation of excess solute from solid solution at temperatures below the solubility boundary is one of the most commonly observed sources of large changes in composition and resistivity in alloys. Decreases in resistivity by a factor of 2 are not uncommon.

Downloaded from https://www.everyspec.com

The formation and growth of a precipitate in alloy systems requires appreciabl time. Precipitation will not occur in solid solutions with less than the saturation quantity of solute.

2.3.3.2.5 Order-Disorder Transformations

Order-disorder transformations are only of theoretical, not practical, interest. Since the long-range order of the lattice electric field disappears when the solute atoms become randomly distributed, the effect on resistivity is quite marked.

The rate of approach of the resistance to an equilibrium value as the solute atoms become randomly distributed obeys the Arrhenius temperature dependence, A exp (-B/T), with the value of the constants A and B characteristic of those for bulk diffusion.

The alloy systems that exhibit this transformation (e.g., Cu-Au, Pt-Cu, Pd-Cu, Mg-Cd) are fully disordered above a critical temperature. Below this temperature, they progressively become ordered with time, as the compositions approach critical stoichiometric ratios. Partially ordered (short-range-ordered) conditions exist that have a correspondingly reduced effect on the resistivity of the alloy.

2.3.3.2.6 Homogenization

Homogenization is a process wherein composition gradients, which may result from nonuniform vapor compositions, are reduced to zero. The process involves diffusion, and its kinetics are therefore those of the diffusing species. The corresponding effects on resistivity may be complicated due to the compositional changes. This type of segregation of constituents should ordinarily be annealed out during manufacture and should not be a factor in the reliability of thin-film resistors. However, it may be a factor in microelectronic circuits where different active materials are deposited in layer fashion.

2.3.3.2.7 Stress Relief

Internal stress may arise in films that have been deposited onto hot substrates. Where a difference between thermal-expansion coefficients of the film and substrate exists, the film becomes stressed when the unit is cooled to room temperature. Such stresses may spontaneously relieve themselves as the result of physical changes in the film. The processes of stress relief are observed to precede recrystallization of cold-worked metals at elevated temperatures. The rates and temperature dependence of stress relief in pure metals indicate a mechanism of bulk self-diffusion.

2.3.3.2.8 Vacancy Condensation

Vacancy condensation can rarely produce more than barely detectable changes in the effective cross section of the resistive path, since their concentration in metals is very low. Such a decrease in cross section would occur by self-diffusion and would not necessarily produce an increase in resistance, since the vacant sites scatter electrons when dispersed, just like solute atoms. Since these effects tend to offset one another, no significant change is expected to result from the process.

Downloaded from https://www.everyspec.com

Recovery processes can proceed at surprisingly low temperatures. In copper, silver, and gold, appreciable recovery of resistivity occurs at -150°C, -180°C, and -120°C, respectively. The vacancy condensation process and subsequent collapse to form dislocation loops has received considerable attention because of its effect on the mechanical properties of pure metal crystals rather than because of a serious effect on resistivity.

2.3.3.3 Specific Failure Studies

2.3.3.3.1 Nickel-Chromium ("Evanohm") Thin-Film Resistors

Investigations[74-76] have been made on the aging characteristics of Evanohm, a film resistor deposited from a (75-20) nickel-chromium alloy with 2.5 per cent each of aluminum and copper. During vacuum evaporation of the alloy, fractionation occurs, so that thin resistor films contain relatively less nickel. A typical film composition was shown to be: 49 per cent nickel, 24 per cent chromium, 14.7 per cent aluminum, and 12.3 per cent copper (in weight per cent).

A study was made of the oxidation in air of an Evanohm film that had been deposited onto a quartz crystal. The growth of the oxide film was monitored by observing the change in oscillation frequency of the crystal resulting from the increase in mass due to oxidation. Oxidation was carried out at 250°C in oxygen. The frequency shift, Δf, was represented fairly well after 1 day by a logarithmic expression:

$$\Delta f = 740 \log_{10} t + 840 \quad , \tag{2.27}$$

where t = elapsed time in days and Δf is in hertz (Hz). In order to determine the oxide-film thickness from the shift in oscillation frequency of the quartz crystal, the following considerations must be applied. Let v be the velocity of propagation of a transverse electric wave in quartz. For an "AT"-cut quartz crystal, the propagation of the wave is in the direction of the crystal thickness (d), and the corresponding frequency of oscillation (f) is given by:

$$f = \frac{v}{2d} \quad . \tag{2.28}$$

If d is increased to $d + \Delta d$, the corresponding frequency shift is given by

$$\Delta f \simeq -\frac{v}{2d^2} \Delta d \quad . \tag{2.29}$$

It can be assumed as a first approximation that this change in frequency is solely from the increase in crystal mass (Δm) due to the thickness increment, Δd.

Therefore,

$$\Delta m = \rho A \Delta d = -\frac{\rho v A}{2f^2} \cdot \Delta f \quad , \tag{2.30}$$

Downloaded from https://www.everyspec.com

where

ρ = density of quartz

A = area of relevant crystal face.

If, however, Δm was due to the growth of a film of some other material onto the crystal face, the value of Δf is not appreciably changed.

For quartz at 25°C,

$$\rho = 2.66 \text{ g/cm}^3 \text{ and } v = 3.34 \times 10^5 \text{ cm/sec} \quad .$$

Therefore,

$$\frac{\rho v}{2} = 4.44 \times 10^5 \quad .$$

Typical values are: $A = 0.1 \text{ cm}^2$

$d = 0.167$ mm

$f = 10^7$ Hz = 10 MHz.

A value of Δf equal to 10 Hz is detectable. This corresponds to a Δm of 4.4×10^{-9} g and to a film thickness in the order of 1 Å.

In the present instance, Δm is the mass of oxygen that has been involved in oxidation of the Evanohm film. If a pure metal had been oxidized, it would then be a simple matter of chemistry to calculate the mass of metal involved in the oxidation; addition of the mass of metal and the mass of oxygen gives the mass of oxide. Division of this by its density gives its volume. Division of the volume by the area yields the film thickness (assumed uniform).

In the case of an alloy such as Evanohm, it is necessary, assuming the oxide film is of uniform thickness and composition, to determine the composition of the oxide and its density in order to obtain the mass of reacted metal, mass of the oxide film, and its thickness.

Further investigations showed that, after aging at 250°C, oxide particles began to appear, especially in the region of grain boundaries. These particles were tentatively identified as nickel oxide, NiO, by electron diffraction.[75]

The presence of a precipitation phenomenon was confirmed by aging resistor components in a vacuum of 10^{-} torr or better while monitoring resistance. Under these conditions, no appreciable oxidation can occur. Very thin films, consistently exhibited a resistance decrease, while thicker films first showed a resistance increase followed by a decrease to over-all lower values. This behavior is attributed to precipitation of an intermetallic compound, Ni_3Al. The precipitation is preceded by short-range ordering, as in the case with ternary alloys of Ni-Cr-Al. The additional presence of copper does

Downloaded from https://www.everyspec.com

not seem to influence the process. It also appears that the particulate stage of oxidation does not start until precipitation starts.[77] Consequently, this complicates the mathematical modeling of the failure aspects of the device.

In other studies of Evanohm[52], it was assumed that aging was due to some sort of activated process considered to be oxygen diffusion, but the particular process was not specified. An aging parameter can be empirically determined that is a function both of time and temperature and corresponds to a given amount of aging. Thus, a given resistor may change in value by 1 per cent after t_1 hours at temperature T_1 or t_2 hours at temperature T_2. These differing temperatures and times would still correspond to a constant value of the aging parameter. The amount of resistance change is a function only of the aging parameter.

This parameter can be derived from the Eyring theory of rate processes. The rate at which many processes proceed is related to temperature by the following equation:

$$R = A \exp\left(-\frac{Q}{RT}\right) \quad , \tag{2.31}$$

where

$$A \propto \frac{kT}{h} \exp \frac{\Delta S}{R} \quad . \tag{2.32}$$

The rate in this case is given by the ratio of relative resistance change to time where t is the time necessary to produce a given resistance change at temperature T.

$$\frac{1}{t} \propto A e^{-\frac{Q}{RT}} \tag{2.33}$$

or

$$T \log A t = \frac{Q}{R} \log e + \text{a constant} \quad . \tag{2.34}$$

This function T log A t is the aging parameter. Since $A \propto \frac{kT}{h} \exp \frac{\Delta S}{R}$, it is temperature dependent. In the work described, Evanohm resistors were subjected to constant stress and step-stress conditions. The relative amount of resistance change was then plotted against the aging parameter. "A" was taken as constant, since this led to only 5 per cent error in results. A value of $A = 10^{15}$ reciprocal days was found to give good agreement between the resistance changes and the aging parameter. T, of course, is the absolute temperature, normally measured in degrees Kelvin.

In a study of (78-20) nickel-chromium alloy film resistors[78], a curious behavior was found which, while not definitely a case of failure, should nevertheless be appreciated. The alloy films were evaporated onto different flat substrates coated with 1-μ-thick silicon monoxide; after film resistor deposition, a furthe protection coating of SiO was deposited on top of the resistor. It was found that a reddish-brown coloration of the SiO developed between adjacent turns of the film resistor at temperatures between 100°C and 300°C. On further aging, the outer SiO coating developed cracks and peeled off the resistor. Necessary conditions for the appearance of this behavior were (1) current flow in the

Downloaded from https://www.everyspec.com

resistor, so that a potential difference existed between adjacent turns, (2) a humid ambient (when the ambient was a vacuum or dry oxygen, the phenomenon was not observed), and (3) a minimum of alkali content in the substrate glass (Corning 0211) [the phenomenon did not appear when substrates of alumina (Alsimag 614) and Corning 7059 glass we used].

With the spalling of the protective coating, the resistance of the alloy film increased, presumably due to the onset of oxidation.

2.3.3.3.2 Tin Oxide Film Resistors

Tin oxide film resistors are fabricated by spraying stannic chloride onto a hot substrate. Stannic oxide is formed, together with small amounts of free tin and stannous oxide. The films are usually doped with small quantities of antimony. After deposition, the films are usually spiralled to form a helix, as in the case of Evanohm resistors.

Life tests, which have been run on tin oxide resistors with both step- and constant-stress levels, gave results that were in good agreement with the aging parameter, T log A t, referred to previously in connection with Evanohm film resistors.[52]

It was found that, at temperatures above 350 °C, anomalous behavior occurred. First, there was an increase in resistance with time and temperature; second, a rapidly decreasing resistance; and third, a rapidly increasing resistance. The rapid decrease in resistance is apparently the result of dissociation of stannous oxide into tin and stannic oxide, which takes place above 350 °C. Since accelerated-test data should be representative of operating conditions, temperatures should not be allowed to exceed 350 °C; a new failure mechanism becomes significant above this temperature.

The activation energy for the process operating at temperatures below the onset of stannous oxide decomposition was determined by taking a constant level of change in resistance observed at different temperatures, T, and plotting log (time) versus T^{-1}. A straight line should result, the slope of which is $-\frac{Q}{R}$, where Q is the activation energy and R is the molar gas constant. This was done for change levels from 0.1 per cent to 5 per cent. The average activation energy obtained was 35 kcal/mole. The upward drift in resistance, therefore, probably results from a diffusion-dependent reaction having an activation energy of 35 kcal/mole.

Investigations of 1/f noise in tin oxide resistors indicate that the detection of excessive 1/f noise is a good screening criterion.[79] Units with higher-than-average values of noise level were shown to have a much higher failure probability. In fact, manufactured units individually inspected for obvious defects revealed such things as scratches across the film path, cracks or checks in the substrate high-resistance terminations, and bridging of adjacent turns by foreign material as being associated with excessive 1/f noise. In those cases where it was possible to correct or reduce the extent of the defect, immediate reductions in noise levels resulted.

Downloaded from https://www.everyspec.com

2.3.3.3.3 Tantalum Thin-Film Resistors

Information on reliability of the tantalum film resistor is sparse. In work describing the development of a highly reliable thin-film resistor[80], the tantalum was evaporated by electron-beam melting, and an outer coating of SiO was found necessary for minimum resistance change. Various substrates were investigated, including soda lime glass, Pyrex, Corning 7059 glass, and Corning 7059 with a coating of SiO. The most reliable units were made on the Corning 7059 with a coating of SiO. Good results were also obtained with the Corning 7059 substrate without the SiO overlayer. Corning 7059 is an alkali-free glass. The best deposition rates were 1 Å/sec onto substrates at 260°C. The best contact materials were aluminum under copper applied by evaporation.

Very few catastrophic failures, e.g., substrate breakage or film rupture, occurred. Resistors that had drifted out of tolerance usually had foreign material accumulated at the negative terminal or at a thin area where it appeared that the resistive material had been depleted by some process. This failure mechanism was not observed on alkali-free glass substrates. Consequently, alkali glasses were rejected for use as substrates. The resistance of tantalum films increased logarithmically during aging.

Downloaded from https://www.everyspec.com

2.3.4 Integrated Circuits

2.3.4.1 General Aspects of Integrated Circuit Reliability

The potential advantages of the integrated approach to electronic circuits, in addition to the small size and weight achieved, are low cost, high reliability, and, in some cases, improved performance. There still appears to be general agreement, however, that the concept of integrated circuits does not automatically lead to high reliability. The circuit designer, for example, must play a much greater role if the high potential reliability of the integrated circuit is to be realized. Circuit design must take into account the wide range of unit-to-unit parameter variations encountered with semiconductor elements. Also, the relationship between consistency in processing techniques and reliability levels becomes more critical with higher packing densities. For the most widely developed integrated circuit approach, the silicon-based monolithic technique, reliability data available today are largely generic in nature. This stems from the fact that the fabrication techniques used borrow heavily from established silicon transistor technology. For the other major approach, the so-called thin-film technique, sufficient experience and use have not yet been accumulated to validly assess the reliability. This general lack of a suitable basis for reliability assessment is due to several factors:

a. Rapid pace of technology in this area.

b. Diversity of approaches being concurrently developed by industry (hence, the lack of extensive reliability analysis of any single approach).

c. Lack of an adequate basis for defining reliability at the circuit level.

d. Lack of central guidance, such as military specifications for most circuits.

e. Extremely small size and susceptibility to damage during test and evaluation.

f. Prohibitive test time and costs required to verify expected high levels of reliability.

g. Inconsistencies in existing test methods and test results.

For some time to come, then, it is clear that any claims or assumptions for the reliability of integrated circuits can only be made in terms of potential reliability. It should be pointed out, however, that some of the factors enumerated above also mitigate against the reliability assessment of any highly reliable electronic device.

As suggested above, there are two basic approaches to integrated circuitry--the silicon-based monolithic approach and the thin film technique. There are, of course, many variations and combinations of the two currently being investigated. Discussion in this section will be limited to the silicon-based technique since considerably more experience is available as a basis for reliability analysis of it than for any other approach. A more complete and detailed discussion of the various approaches is reserved for later editions of this handbook due to the rapidly evolving situation in integrated circuit technology, and the fact that there are several efforts currently in progress on reliability assessment of such circuits.

Before discussing the principal failure modes and mechanisms experienced in the silicon integrated circuit, it may be well to summarize the more common general construction features of this circuit type. Both active and passive elements are introduced into a single-crystal piece of silicon. In one approach, the resistive elements are doped silicon diffused into the silicon substrate. With a p-type substrate, for example, an n-type silicon layer is used for isolation with a p-type layer over it to act as the resistive layer. In another approach, thin resistive films of different materials are deposited on top of the oxide-covered silicon substrate. Capacitors for integrated circuits may be formed either by using the capacitance inherent in a large area p-n junction, or through the use of a thin film of silicon dioxide between

Downloaded from https://www.everyspec.com

an n-type silicon layer and the aluminum metalization. An alternative approach is to deposit a thin film capacitor on top of the oxide covered silicon substrate. In common with the planar technology of discrete semiconductor devices, photoprocessing, diffusion, and epitaxial growth are the principal processes used in the silicon monolithic approach to form active devices onto the silicon substrate. Diffusion is used to control the depth of structures, while photoprocessing controls the surface geometry. The planar technology permits the evaporation of a film of aluminum to form simultaneously the contacts and conductive interconnecting paths between elements.

The following list typifies some of the failure modes experienced in representative digital integrated circuits of the silicon-based type using diffused passive elements:[81]

a. Die (Bar) Degradation
   (1) Inversion
      (a) Channeling
      (b) Piping
   (2) Contamination
   (3) Improper surface oxide or passivation
   (4) Improper diffusion
      (a) Diffusion of gold into the silicon due to excessive temperature
      (b) Low breakdown voltage along major portion of junction
      (c) Low breakdown voltage along isolated portion of junction
   (5) Current concentration centers
   (6) KMER or photolithographic mask defect leading to improper diffusion

b. Die(Bar) Material Degradation
   (1) Material defects (bulk defects)
   (2) Cracks

c. Faulty Interconnections (Evaporated Lead Pattern)
   (1) Faulty evaporation or deposition
   (2) Cracks (fractures, scratches or other surface defects)
   (3) Separation from the die (bar) (evaporated pattern peeling from oxide)
   (4) Open evaporated lead pattern
      (a) Melted leads (open) due to excessive current flow
      (b) Associated with a scratch
   (5) Migration of gold into lead due to excessive temperature
   (6) Chemical reaction between materials composing evaporated lead pattern
   (7) Chemical reaction between lead pattern and reagent residue
   (8) Poor ohmic contact of evaporated lead to silicon

d. Faulty Resistor
   (1) Unstable resistance element
   (2) Current concentration centers

e. Faulty Bond to Die (Bar) Process
   (1) Faulty evaporation or deposition of contact (bonding pad)
   (2) Cracks in bond

Downloaded from https://www.everyspec.com

(3) Voids in bond
(4) Separation of bond
  (a) Faulty bond placement
  (b) Improper bonding conditions
  (c) Chemical reaction between bond material and lead material
(5) Separation of lead from bond
(6) Purple Plague
(7) Faulty preform to die (bar) bond
  (a) Voiding between preform and die (bar)
  (b) Die (bar) overhang of preform (area wetted by preform less than area of the die (bar)
  (c) Insufficient alloying of die to preform
(8) Faulty preform to case bond
(9) Voids in preform

f. Faulty External Terminal - Lead Wire Assembly
(1) Faulty terminal - lead connection (lead wire weld)
(2) Improper lead wire routing
(3) Sagging wires
  (a) Shorting to edge of die (bar)
  (b) Shorting to wrong evaporated lead
(4) Broken lead wire
(5) Lead wire shorted to case

g. Improper Packaging
(1) Faulty cap to case seal
(2) Faulty terminal to insulator seal
(3) Faulty insulator to case seal
(4) Terminal fatigue
(5) External surface contamination
(6) Marking deterioration
(7) Finish deterioration
(8) Foreign material inside package
(9) Void in Pyroceram

h. Improper Measurement and/or Test Procedures

i. Parasitic Transistor Action

j. Stressed Beyond Device Design Capabilities
(1) Mechanical
  (a) Broken wire
  (b) Open bond
  (c) Broken die (bar)
  (d) Broken case
(2) Thermal
  (a) Alloying of bond material into die (bar)
  (b) Alloying of bond material into evaporated lead

Downloaded from https://www.everyspec.com

(c) Alloying of evaporated lead into die (bar)

(3) Electrical

(a) Incorrect bias values

(b) Power dissipation limitations exceeded

Since the processes used to fabricate the silicon-based integrated circuit are generally the same as those used in fabricating planar epitaxial silicon transistors and diodes, it is apparent, as shown in the above list, that the failure modes and mechanisms are similar. The proximity of several elements on a common substrate and the longer conductive interconnect paths are, of course, features of the integrated circuit which would tend to aggravate some of the mechanisms encountered in discrete silicon devices. The major circuit design problems with the silicon-based approach are parasitics, due to poor isolation between components, poor dimensional tolerance, temperature and voltage sensitivity of diffused passive components, and an inability to realize useful inductive effects in diffused silicon bodies. Through a series of compromises and hybrid variations, all but the last of these problems have been resolved to a certain extent.

Several methods for assessing integrated circuit reliability have been proposed. In one such method, the Test Element Group (TEG),[82] each circuit element of the integrated circuit is simultaneously reproduced on the same silicon wafer and connected to its own separate leads on the external package. Each circuit element, representing its counterpart in the adjacent functional integrated circuit, can thus be individually subjected to test and measurement. In some cases, additional special devices, such as MOS structures, are added to the TEG to furnish an indication of surface oxide quality. In this manner, it is hoped to develop techniques for improving process and quality control, as well as for reliability assessment.

There are several advantages to the TEG approach:

1. It is possible to detect very small changes in the individual circuit elements of the microcircuit since contacts are made to each node as required. Therefore, effects much too small to be noted in the functional circuit, but which may gradually increase and cause failure, may be detected earlier.

2. Special devices which may not be included in the functioning circuit, but which are very sensitive to certain failure mechanisms (e.g., IGFET) may be included on the test pattern. These special devices may give clues to the life performance of a circuit in a short time.

3. It may be possible to test a much smaller number of circuits, but with a much greater amount of physical information.

Some of the assumptions involved in the TEG approach are:

1. That the test pattern on the wafer is identical in material properties and processing to its "siblings."

2. That a correlation can be made between aging performance of the test device and reliability of the sibling integrated circuits. In other words, will the changes measured in test pattern parameters reflect failure at a later time in the integrated circuits?

3. That valid extrapolation from high stress test conditions to use conditions is possible. (This assumption, of course, requires confirmation in high stress tests of any part.)

Another method for assessing integrated circuit reliability was used in the Minuteman Component Quality Assurance Program (CQAP).[83] This was primarily a part improvement program, but estimates of failure rates were made. This method requires that an initial estimate be made of the failure rate of the integrated circuit which is to be improved. This estimate is made on the basis of whatever test data

Downloaded from https://www.everyspec.com

and experience are available on the integrated circuit or its individual elements. A series of high stress tests is then employed to generate failures of the circuit being tested. After failure analysis and categorization of failure modes, corrective actions are taken to eliminate or minimize particular failure modes. It is then assumed that the initial failure rate estimate has been reduced by the percentage that the "eliminated" failure mode originally contributed to the total number of failures recorded during tests. This technique is more suited to improving rather than assessing reliability, however. Some of the obvious difficulties are:

1. The initial estimates of the failure rate may be difficult to obtain and justify.
2. No acceleration factors are derived for the failure modes observed which can be used to extrapolate failure mechanism rate at high stress to normal stress conditions.
3. High stress tests may introduce or make dominant failure mechanisms which are not prevalent at use conditions.

In the section which follows, failure mechanisms in integrated circuits are classified as bulk, surface, contact, and isolation. The discussion is largely limited to those aspects of failure mechanisms in integrated circuits not previously covered under discrete components.

2.3.4.2 Bulk Effects

As with discrete semiconductor devices, fabrication of silicon integrated circuits depends on good control of the type and distribution of chemical impurities introduced into the silicon substrate. Although the impurity atoms can move by diffusion at any temperature, this movement is negligible at operating temperatures for boron and phosphorous dopants. The presence of defects, such as dislocations, stacking faults, and other microstructural cracks and strains, may act as sinks for unwanted impurities and precipitates. Copper, for example, can substantially reduce the breakdown voltage of a junction and soften its reverse characteristics. Introduction of foreign materials can also cause softening of the isolation junctions used in a circuit. This leads to high leakage currents and increased parasitic effects. Several alternative routes to the back-biased p-n junction method of isolation are currently being adopted, however, in modern integrated circuit technology.

Failure mechanisms caused by bulk diffusion characteristics may be detected through high temperature storage tests. The presence of microstructural defects is often detected by the use of mechanical and thermal shock tests along with appropriate electrical parameter measurements. The distribution and density of defects on a silicon wafer is an important consideration when it is to be used for complex integrated circuits.

2.3.4.3 Surface Effects

In common with the discrete silicon semiconductor devices, once the gross mechanical and quality type defects have been screened out of the integrated circuit, surface contamination effects remain as a major reliability problem. While a thermally-grown passivating layer of SiO_2 gives a certain degree of protection from external contamination, it does not completely eliminate the field effects of the presence of charges on its surface. Also, since these passivating layers are amorphous structures with wide variability in lattice spacings, ions are diffused at relatively fast rates, even in the temperature ranges from 300°C to 500°C.

Compared with discrete semiconductor devices, integrated circuits tend to limit freedom in the design of variable current, or power, life tests because they are more complex and because their intended application is usually more specific. For these reasons, circuit bias conditions are often fixed, and circuit

Downloaded from https://www.everyspec.com

elements must operate at stress levels well below maximum ratings for their discrete counterparts. At lower current levels, the dominant failure modes will therefore be surface oriented. This suggests that a d-c biased operating life test at an elevated ambient temperature should be a suitable method of testing the long term stability of an integrated circuit.[84] However, this lack of freedom prevents testing at high current levels which have been used in the past (e.g., on transistors and diodes) to accelerate certain failure modes other than surface related ones. Lack of hign current tests may mitigate the apparent importance of these nonsurface failure modes.

Low resistance paths are commonly caused by various degrees of dielectric breakdown of the passivating surface layer. Conductive films can find electrical shorting paths when microscopic areas of the passivation layer become thinner than 1000 Å in high field regions, or when these areas become impregnated with foreign conducting material. Pinholes in the oxide under metal interconnecting films have caused shorting of junctions by metal migration through the pinholes. Pinholes in the oxide over active junctions have also caused failure by surface conduction through ionization of contaminants.

Pinholes can be detected by means of a "Zapp" test.[85] This test is designed to initiate catastrophic breakdown of any thin or degenerate microareas of the dielectric, while at the same time it is so restrained as to be nondestructive to normal oxide protective coatings. The "Zapp" test consists of the application of an electric field of appropriate voltage and polarity with suitable protective current limiting resistance to various inputs of the integrated circuit. The field will cause a permanent breakdown of dielectric which subsequent electrical tests will indicate as a passivation problem.

2.3.4.4 Contacts

Since there are not only fewer external connections but also fewer interfaces between different materials in integrated circuits than in their discrete element circuit counterparts, reliability should be greater. The interconnect problem, however, still requires attention. Longer conductive paths over the oxide layer are required in integrated circuits than for individual transistors. As discussed previously, faults in the oxide can short circuit the metalization leads to the substrate. Variations in thickness of the soft overlay metal film, due either to mechanical scoring and scratching, or to narrowing when crossing an oxide step, may lead to high resistance paths. Void areas resulting from improper metalization, nonadherence to the surface oxide, or a dissolution of interconnect material can also result in unwanted opens.

If gold is used for internal wire leads with aluminum interconnects, the usual intermetallic compound formations, probably catalyzed in the presence of silicon, are encountered at the higher temperatures. Heat treatment temperatures may also cause the aluminum metal to react with the surface oxide when defects are present in the oxide.

Thermal compression bond failures in integrated circuits have been attributed to:

a. Incomplete diffusion at the Au-Al interface.

b. Nonadherence of aluminum to silicon.

c. Volumetric changes in the diffusion region producing (1) an interface of the gold bond and the diffusion region, (2) an interface of two metallic phases, (3) an annular crack between two intermetallic phases; and (4) a surface crack in the silicon die.

Metallurgical weaknesses can be detected through high temperature storage tests in conjunction with mechanical and thermal shock tests. Typical parameter degradation effected by thermal compression bond failures are increased saturation voltage and increased output cutoff current.

Downloaded from https://www.everyspec.com

2.3.4.5 Isolation

Since all circuit elements, both active and passive, are located in close proximity on a common substrate, thermal, chemical, mechanical, and electrical coupling may be expected to be strong. Hence, the problem of isolation between elements in the silicon integrated approach creates a serious design problem. There are many schemes for achieving electrical isolation, and new ones are continually being developed. When reversed-biased p-n junctions are used for isolating transistors, there are several limitations. The disadvantages include capacitive coupling to all other isolated regions at high frequencies, large increases in collector spreading resistance, junction leakage, inversion layers, and undesirable transistor and p-n-p-n effects. Current developments tend toward the dielectric isolation method where an actual insulated layer separates the elements from the substrate. When this method is used, the isolation regions are, of course, subject to the same surface and bulk effects previously mentioned. It is clear that the problem of electrical isolation is a serious design problem, but very little has been done to show how parasitics affect the reliability of the total integrated circuit.

The principal mode of heat transfer from the circuit elements is conduction through the die to, typically, a ceramic case. The thermal conductance of the common silicon substrate is high enough so that intra-die gradients tend to be quite small. Strains may be transferred to the die, however, if a thermal mismatch of materials exists between the die and the case, as often occurs in power transistors A reduction in the thermal resistance of die-to-case bonds is still being sought.

REFERENCES

(1) Miles, J. L., and Smith, P. H., "Formation of Metal Oxide Films Using Gaseous and Solid Electrolytes", J. Electrochem. Soc., 110, 1240 (1963).

(2) Tamm, I., "A Feasible Model of Electron Bonding at Crystal Faces", Phys. Z. Soviet Union, 1, 733 (1932).

(3) Shockley, W., Dislocations and Edge States in the Diamond Crystal Structure", Phys. Rev., 91, 228 (1953).

(4) Lander, J. J., and Morrison, J., "Low Energy Electron Diffraction Study of Silicon Surface Structures", J. Chem. Phys., 37, 729 (1962).

(5) Weisberg, L. R., and Blanc, J., "Diffusion With Interstitial-Substitutional Equilibrium Zinc in GaAs", Phys. Rev., 131 (4), 1548-52 (19 3).

(6) Langmuir, I. J., "Thoriated Tungsten Filaments", J. Franklin Inst., 217, 543 (1934).

(7) "Reliability of Semiconductor Devices", Final Report on Contract NObsr 81304 (December, 1961), 18-19.

(8) Tamburrino, A. L., "Analysis of Requirements in Reliability Physics", PFE*, 2, 1-24 (1964).

*Note: Frequent reference will be made to Physics of Failure in Electronics, Edited by M. F. Goldberg and J. Vaccaro. This will be abbreviated to PFE, followed by the volume number.

Downloaded from https://www.everyspec.com

(9) Dash, W. C., "Growth of Silicon Crystals Free From Dislocations", J. Appl. Phys., 30. 459 (1959).

(10) Queisser, H. J., "Slip Patterns on Boron-Doped Silicon Surfaces", J. Appl. Phys., 32, 1776 (1961).

(11) Prussin, S., "Generation and Distribution of Dislocations by Solute Diffusion", J. Appl. Phys., 32, 1876 (1961).

(12) Brors, D. L., "Formation of Dislocations in Oxide Passivated Silicon", Electrochemical Society Meeting, Washington, D. C. (October, 1964).

(13) Queisser, H. J., "Dislocations and Semiconductor Device Failure", PFE, 1, 146-155 (1963).

(14) Queisser, H. J., "Failure Mechanisms in Silicon Semiconductors", Final Report RADC-TDR-62-533 on Contract AF 30(602)-2556 (ASTIA No. AD 297 033) (January, 1963).

(15) Queisser, H. J., Hubner, K., and Shockley, W., "Diffusion Along Small Angle Grain Boundaries in Silicon", Phys. Rev., 123, 1245 (1961).

(16) Scarlett, R. M., Shockley, W., and Haitz, R. H., "Thermal Instabilities and Hot Spots in Junction Transistors", PFE, 1, 194-203 (1963).

(17) Goetzberger, A., and Shockley, W., "Metal Precipitates in Silicon p-n Junctions", J. Appl. Phys., 31, 1821 (1960).

(18) Queisser, H. J., Final Report on Contract AF 33(616)-7786, "Solar Cell Parameter Study" (1962).

(19) Queisser, H. J., "Properties of Twin Boundaries in Silicon", J. Electrochem. Soc., 110, 51 (1963).

(20) "Reliability of Semiconductor Devices", Final Report on Contract NObsr 81304 (December, 1961), 18-19.

(21) Hooper, W. W., and Queisser, H. J., "Photoresponse of Small-Angle Grain Boundaries in Silicon", Bull. Am. Phys. Soc., 7, 211 (1962).

(22) Kaiser, W., Frisch, H. L., and Reiss, H., "Mechanism of the Formation of Donor States in Heat-Treated Silicon", Phys. Rev., 112, 1546 (1958).

(23) Kaiser, W., "Electrical and Optical Properties of Heat-Treated Silicon", Phys. Rev., 105, 1751 (1957).

(24) Lederhandler, S., and Patel, J. R., "Behavior of Oxygen in Plastically Deformed Silicon", Phys. Rev., 108, 239 (1957).

Downloaded from https://www.everyspec.com

(25) Dash, W. C., "A Method for Measuring the Thickness of Epitaxial Silicon Films", J. Appl. Phys., 33, 2345 (1962).

(26) Queisser, H. J., and Goetzberger, A., "Microplasmas at Stacking Faults in Silicon", Bull. Am. Phys. Soc., 7, 602 (1962).

(27) Queisser, H. J., "Stacking Faults and Failure of Silicon Devices", PFE, 2, 476-482 (1964).

(28) Queisser, H. J., and Goetzberger, A., "Microplasmas at Stair-Rod Dislocations in Silicon", Phil. Mag., 8, 1063 (1963).

(29) Johnson, M. C., "Solid State Thermal Diffusion: A Contributor to Degradation of Semiconductor Junction Devices", PFE, 2, 124-144 (1964).

(30) Schafft, H. A., and French, J. C., IRE Trans., ED-9, 129 (1962).

(31) Schafft, H. A., and French, J. C., "Characteristics of Second Breakdown and Transistor Failure", PFE, 1, 104-107 (1963).

(32) Scarlett, R. M., and Schroen, W., "Localized Thermal Effects in Silicon Power Transistors", PFE, 2, 285-303 (1964).

(33) Queisser, H. J., and Hooper, W. W., "Study on a Definitive Confirmation of Thermal Instability in Silicon Power Transistors", Final Report RADC-TDR-63-350 on Contract AF 30(602)-3084 (September, 1963).

(34) Schroen, W., and Hooper, W. W., "Failure Mechanisms in Silicon Semiconductors", Report Nos. RADC-TDR-64-153 and RADC-TDR-64-361 on Contract AF 30(602)-3016.

(35) Schroen, W., Beaudoiun, J., and Hubner, K., "Failure Mechanisms in High-Power Four-Layer Diodes", PFE, 3 (to be published).

(36) Nanavati, R. P., "Some Physical Mechanisms Contributing to Tunnel Diode Failure", PFE, 1, 214-222 (1963).

(37) Nanavati, R. P., "Important Mechanism Contributing to Tunnel Diode Failure", PFE, 2, 550-559 (1964).

(38) Anderson, R. L., "On the Degradation of Gallium Arsenide Tunnel Diodes", PFE, 2, 328-336 (1964).

(39) Henkel, H. J., "Aging Effects in Gallium Arsenide Tunnel Diodes", Z. Naturforschung, 17a, 358 (1960).

(40) Holonyak, N., Jr., Selig, T., and Smith, J., "Hump Current Dependence Upon Trapping Effects and the Relationship to Some Aspects of Forward Injection Failure of GaAs Tunnel Diodes", IRE Trans., ED-8, 427 (1961).

Downloaded from https://www.everyspec.com

(41) Gold, R. D., and Weisberg, L. R., "The Degradation of Gallium Arsenide Tunnel Diodes", IRE Trans., ED-8, 428 (1961).

(42) "Mechanisms of Failure in Semiconductor Devices", Report RADC-TDR-63-338 on Contract AF 30(602)-2778 (September, 1963).

(43) "Failure Mechanisms of Tunnel Diodes", Final Report RADC-TDR-64-313 on Contract AF 30(602)-2778 (September, 1964).

(44) "Theoretical and Experimental Studies Relating to Mechanisms of Failure of Semiconductor Devices", Final Report RADC-TDR-62-271 on Contract AF 30(602)-2177 (May, 1962).

(45) Statz, H., and de Mars, G. A., "Electrical Conduction via Slow Surface States on Semiconductors", Phys. Rev., 111, 169 (1958).

(46) Atalla, M. M., Bray, A. R., and Lindner, R. J., "Stability of Thermally Oxidized Silicon Junctions in Wet Atmospheres", Proc. IEEE, 106B, 1130 (1959).

(47) Metz, E. D., "Silicon Transistor Failure Mechanisms Caused by Surface Charge Separation", PFE, 2, 163-172 (1964).

(48) Skinner, S. M., Dzimianski, J. W., "Study of Failure Mechanisms", Final Report RADC-TDR-63-30 on Contract AF 30(602)-2558 (December, 1962).

(49) "Failure Mechanisms in Microelectronics", Reports Nos. RADC-TDR-63-425 (March, 1964), -64-61 (May, 1964), -64-125 (May, 1964), and -64-252 (August, 1964) on Contract AF 30(602)-3017.

(50) Dzimianski, J. W., Pemsel, F. R., Lytle, W. J., and Skinner, S. M., "Silicon Surface Passivation: Materials and Microproperties", PFE, 2, 450-465 (1964).

(51) Schnable, G. L., Schlegel, E. S., and Keen, R. S., "Failure Mechanisms in Reverse-Biased Oxide-Passivated Silicon Diodes", PFE, 3 (to be published).

(52) "Accelerated Testing of High Reliability Parts", Contract AF 30(602)-3415 (October, 1964).

(53) "Detailed Study of Deleterious Effects on Silicon Transistors", Reports No. RADC-TDR-64-111 and -64-352 on Contract AF 30(602)-3244 (April and October, 1964).

(54) Shockley, W., Queisser, H. J., and Hooper, W. W., "Charges on Oxidized Silicon Surfaces", Phys. Rev. Letters, 11, 489 (1963).

(55) Peck, D. S., Blair, R. R., Brown, W. L., and Smits, F. M., "Surface Effects of Radiation on Semiconductors", Bell Syst. Tech. J., 42, 95 (1963).

(56) Hofstein, S. R., Zaininger, K. H., and Warfield, G., "Frequency Response of the Surface Inversion Layer in Silicon", Proc. IEEE, 52 (8), 971 (1964).

(57) Thomas, J. E., and Young, D. R., "Space Charge Model for Surface Potential Shifts in Silicon Passivated With Thin Insulating Layers", IBM J., 8, 368 (1964).

Downloaded from https://www.everyspec.com

(58) Snow, E. H., Grove, A. S.. Deel, B. E., and Sah, C. T., "A Study of Ion Migration in Thin Insulating Films Using the MOS Structure", Abstract No. 129, Electrochemical Society Meeting, Washington, D. C. (October, 1964).

(59) Kerr, D. R., Logan, J. S., Burkhardt, P. J., and Pliskin, W. A., "Stabilization of SiO_2 Passivation Layers With P_2O_5", IBM J., 8, 376 (1964).

(60) Griffin, W. A., Mathews, J. R., and Olson, K. H., "Channel Generation in Glass Encapsulated Silicon Diodes", Abstract No. 134, Electrochemical Society Meeting, Washington, D. C. (October, 1964).

(61) Packard, C. C., "Reducing Reliability Risks Through Failure Analysis, Screening, and Process Improvement", Semiconductor Reliability, 2, Edited by W. H. Von Alven, Engineering Publishers, Elizabeth, New Jersey, 91-99 (1962).

(62) Longo, T. A., and Selikson, B., "Identification and Elimination of a Failure Mechanism in Semiconductor Devices", PFE, 2, 424 (1964).

(63) Selikson, B., "The Role of Compound Formation in Semiconductor Device Reliability", PFE, 3 (to be published).

(64) Anderson, G. P., and Erickson. R. H., "Failure Modes in Integrated and Partially Integrated Microelectronic Circuits", PFE, 2, 498-524 (1964).

(65) Wortman, J. J., and Burger, R. M., "Fundamental Failure Mechanism in Thin-Film Metal-Dielectric Structures Observable as a Generated Voltage", PFE, 2, 173-188 (1964).

(66) Klein, N., Gafni, H., and David, H. J., "Mechanism of D.C. Electrical Breakdown in Thin Silicon Monoxide Films", PFE, 3 (to be published).

(67) Vermilyea, D. A., "The Crystallization of Anodic Tantalum Oxide Films in the Presence of a Strong Electric Field", J. Electrochem. Soc., 102, 207 (1955).

(68) "Study of Failure Mechanisms at Surfaces and Interfaces", Final Report RADC-TDR-63-152 on Contract AF 30(602)-2593 (March, 1963).

(69) "Failure Mechanisms at Metal Dielectric Interfaces", Reports RADC-TDR-64-138, -64-359, and -64-459 on Contract No. AF 30(602)-3266.

(70) Johnson, N., and Greenough, K., "Diffusion Studies on Stressed Tantalum Oxide Capacitors", PFE, 2, 103-123 (1964).

(71) Shirn, G. A., and Smyth, D. M., "Some Failure Mechanisms at Insulator-Conductor Junctions", PFE, 2, 154-162 (1964).

(72) Smyth, D. M., Shirn, G. A., and Tripp, T. B., "The Heat Treatment of Anodic Oxide Films on Tantalum, I. The Effects on Dielectric Properties", J. Electrochem. Soc., 110, 1264 (1963).

Downloaded from https://www.everyspec.com

(73) Smyth, D. M., and Tripp, T. B., "The Heat Treatment of Anodic Oxide Films on Tantalum, II. Temperature Dependence of Capacitance", J. Electrochem. Soc., 110, 1271 (1963).

(74) "Study of Comprehensive Failure Mechanism Theory", Final Report RADC-TDR-63-431 on Contract AF 30(602)-2731 (August, 1963).

(75) "Study of Comprehensive Failure Theory", Final Report RADC-TDR-64-309 on Contract AF 30(602)-3054 (1964).

(76) Goldberg, M., Horberg, A., Stewart, R., and Levinson, D. W., "Comprehensive Failure Mechanism Theory - Metal Film Resistor Behavior", PFE, 2, 68-93 (1964).

(77) Levinson, D. W., and Stewart, R. G., "Experimental Confirmation of Precipitation and Oxidation in Evanohm Condensate Thin Films", PFE, 3 (to be published).

(78) Smith, P. C., and Genser, M., "An Investigation of Thin Film Resistor Failure", PFE, 3 (to be published).

(79) Curtis, J. G., "Current Noise Measurement as a Failure Analysis Tool for Film Resistors", PFE, 1, 204-213 (1963).

(80) Weston, T. C., and Tomlinson, J. L., "Physics of Failure Principles Applied to a Device Development Program", PFE, 3 (to be published).

(81) Autonetics CQAP Exhibit 384-3, Revision 5, March 3, 1965.

(82) General Electric Contract AF 33(615)-2716, July 1965 (see Progress Reports when available).

(83) Fleming, D.C., "A Method for the Determination of Failure Rates for WS-133 B Electronic Component Parts," Autonetics Company Official Report, October 1964.

(84) Motorola Letter Report No. 10, Contract AF 30(602)-3299, December 16, 1964.

(85) Westinghouse Final Report on Contract DA-36-039-AMC-03617E, April 1965.

NOTE: Proceedings of the Annual Symposia on "Physics of Failure in Electronics" may be obtained as follows:

First Annual (1962): Available at a cost of $7.50 from Spartan Books, Inc., 1106 Connecticut Avenue, N.W., Washington, D.C. 20036.

Second Annual (1963): Available at a cost of $7.00 from the Clearing House for Federal Scientific and Technical Information (CFSTI), Department of Commerce, Washington, D.C. 20230, citing identification number AD 434329.

Third Annual (1964): Available at a cost of $8.00 from CFSTI, citing identification number AD 617715.

Downloaded from https://www.everyspec.com

SECTION 3. PHYSICAL PROPERTIES OF MATERIALS AND PROCESSES PERTINENT TO RELIABILITY PHYSICS

Downloaded from https://www.everyspec.com

TABLE OF CONTENTS

	Page
3. Physical Properties of Materials and Processes Pertinent to Reliability Physics	3-1
3.1 Energy Gap and Carrier Mobility of Semiconductors	3-2
3.2 Energy Levels of Chemical Impurities	3-4
3.3 Diffusion Coefficient	3-4
3.4 Eutectic Melting Points	3-12
3.5 Vapor Pressure	3-13
3.6 Intrinsic Carrier Concentration	3-14
3.7 Thermal Conductivity	3-16
3.8 Emissivities of Various Materials	3-18
3.9 Thermal Coefficient of Linear Expansion	3-19
3.10 Elastic Constants	3-22
3.11 Dielectric Properties	3-25
3.12 Properties of Solder	3-28
3.13 Properties of Encapsulants and Potting Compounds	3-30
3.14 Resistivity and Temperature Coefficient of Resistance of Metals and Alloys	3-36

LIST OF TABLES

Table 3.1	Energy Gap and Carrier Mobility of Semiconductors	3-2
Table 3.2	Energy Levels of Chemical Impurities - Germanium	3-5
Table 3.3	Energy Levels of Chemical Impurities - Silicon	3-6
Table 3.4	Energy Levels of Chemical Impurities - Gallium Arsenide	3-6
Table 3.5	Diffusion Constants of Impurities - Germanium	3-8
Table 3.6	Diffusion Constants of Impurities - Silicon	3-9

Downloaded from https://www.everyspec.com

LIST OF TABLES
(Continued)

		Page
Table 3.7	Diffusion Constants of Impurities - Gallium Arsenide	3-10
Table 3.8	Eutectic Melting Points	3-12
Table 3.9	Vapor Pressure	3-13
Table 3.10	Thermal Conductivity	3-16
Table 3.11	Emissivity	3-18
Table 3.12	Thermal Coefficient of Linear Expansion	3-20
Table 3.13	Elastic Coefficients	3-23
Table 3.14	Dielectric Properties	3-26
Table 3.15	Properties of Solders	3-28
Table 3.16	Typical Cured Properties of Basic Encapsulants	3-31
Table 3.17	Solvents, Trade Names, and Manufacturers of Encapsulating Materials	3-33
Table 3.18	Resistivity and TCR of Metals and Alloys	3-36

LIST OF FIGURE

Figure 3.1	Intrinsic Carrier Concentration - Germanium and Silicon	3-15

Downloaded from https://www.everyspec.com

3. Physical Properties of Materials and Processes Pertinent to Reliability Physics

The preceding section has presented descriptions, models, and mathematical treatments of the physical processes known to be predominant factors in the reliability of electronic parts. This section presents the numerical values of material properties relevant to physical processes that are active in electronic parts.

In forming the following tables, a limited selection of materials and physical properties was made. Materials chosen were those that are now used or have real use potential in advanced electronic equipment. This category includes not only materials for transistors, diodes, thin-film resistors, and thin-film capacitors, but also materials used for conducting leads, contacts, protective overcoats, substrates, and potting materials.

In some cases, a process or property implies the presence of more than one material, e.g., diffusion and diffusion constant. The emphasis in forming tables of appropriate properties in such cases has been on those materials that can be expected to interact because of their relative positions in a system.

Most material properties are temperature sensitive. In these tables, it is assumed that the maximum temperature applicable for electronic systems will be approximately 500°C, and the values tabulated are limited accordingly. Exceptions are made to this procedure in cases where it is felt that the information might be useful to an engineer in eliminating processes from consideration as possible causes of a reliability problem.

Numbers in parentheses next to each datum in the tables are the references from which the datum was taken. References pertinent to each table immediately follow the table.

Downloaded from https://www.everyspec.com

3.1 Energy Gap and Carrier Mobility of Semiconductors

Two properties of semiconductors basic to the solution of most semiconductor problems are energy gap and carrier mobility. These properties, along with the temperature and pressure dependences of the energy gap, are given in Table 3.1.

TABLE 3.1. ENERGY GAP AND CARRIER MOBILITY OF SEMICONDUCTORS

Material	Energy Gap at 0°K, ev	Energy Gap at Room Temperature, ev	Temperature Dependence of Energy Gap, $\beta \times 10^4$, ev/°K[a]	Pressure Dependence of Energy Gap, $\gamma \times 10^6$, ev/atm[b]	Electron Mobility[5], cm^2/volt-sec[c]	Hole Mobility[5], cm^2/volt-sec[c]
Si	1.21[1]	1.09[2]	-4.1[1]	-2.0[4]	1,900	500
Ge	0.785[1]	0.66[2]	-4.4[1]	8.0[4]	3,900	1,900
GaAs	1.53[6]	1.41[7]	-5.0[3]	9.4[4]	7,200	680
CdS	2.6[10]	2.38-2.42[10]	-5.2[3]	3.3[4]	250	10
CdTe	1.606(T~0)[11]	1.44[12]	-2.3 at 77°K[3] -5.4 at 800°K[3]		800	100
CdSe		1.7[2]	-4.6[3]			
InSb	0.2357[6]	0.180[8]	-2.7[3]	15.2[6]	80,000	4,000
InAs	0.46[9]	0.360[6]	-3.5[3]	5.4[6]	30,000	240

(a) $E_G = E_{G(T=0)} + \beta T$.
(b) $E_G = E_{G(P=0)} + \gamma P$.
(c) The mobilities listed are typical of the highest values observed. In any particular sample, mobility values will be dependent upon free-carrier density, crystal perfection, and other factors.

Downloaded from https://www.everyspec.com

REFERENCES, TABLE 3.1

(1) Hannay, N. B., Semiconductors, Reinhold Publishing Corporation, New York (1959), p 332.

(2) Hannay, N. B., loc. cit., p 52.

(3) Bube, R. H., Photoconductivity of Solids, John Wiley and Sons, Inc., New York (1960), Table 7.8-1, p 237.

(4) Bube, R. H., loc. cit., Table 7.8-2, p 238.

(5) Bube, R. H., loc. cit., Table 8.5-1, p 269.

(6) Paul, W., "Band Structure of the Intermetallic Semiconductors From Pressure Experiments", J. Appl. Phys., Suppl. to 32 (10), 2082 (1961).

(7) Ehrenreich, H., "Band Structure and Transport Properties of Some III-V Compounds", J. Appl. Phys., Suppl. to 32 (10), 2155 (1961).

(8) Moss, T. S., Optical Properties of Semiconductors, Butterworth's Scientific Publications, London (1959), p 224.

(9) Hilsum, C., and Rose-Innes, A. C., Semiconducting III-V Compounds, Pergamon Press (1961), p 131.

(10) Neuberger, M., "Cadmium Sulfide Summary Review and Data Sheets", EPIC Report No. DS-124, 63, April, 1963.

(11) Thomas, D. G., "Excitons and Band Splitting Produced by Uniaxial Stress in CdTe", J. Appl. Phys., Suppl. to 32 (10), 2298 (1961).

(12) Neuberger, M., "Cadmium Telluride Data Sheets", EPIC Report No. DS-101, 13, June, 1962.

Downloaded from https://www.everyspec.com

3.2 Energy Levels of Chemical Impurities

The electrical properties of a semiconductor are determined to a large degree by the concentration and chemical species of impurities in the semiconductor. An impurity atom may be capable of contributing either a free electron or a free hole when ionized, or it may be electrically inactive. The amount of energy required to ionize an impurity and, consequently, provide a charge carrier, depends upon certain properties of both the host crystal and the impurity atom, and is different for different combinations of them. A donor-type impurity atom, when ionized, contributes an electron to the conduction band of the crystal. A crystal containing predominantly donor-type impurities is called an n-type crystal. An acceptor-type impurity, when ionized, contributes a hole to the valence band of the crystal. A crystal containing predominantly acceptor-type impurities is called a p-type crystal. One type of atom may contribute either holes or electrons, depending upon its mode of occupation of the host crystal.

Tables 3.2, 3.3, and 3.4 list impurities known to affect the electrical conductivity of germanium, silicon, and gallium arsenide. The quantity "$E_C - E_I$" is the energy difference between the valence electron energy of the impurity atom in the unexcited state and the minimum energy of electrons in the conduction band of the host crystal. The quantity "$E_I - E_V$" is the energy difference between the maximum energy of a valence electron of the host crystal and a valence electron of the impurity atom in the host crystal. The "maximum solubility" column tells the number of impurity atoms of a given type that may be added to the host crystal without precipitation or agglomeration.

3.3 Diffusion Coefficient

The concept and ramifications of diffusion coefficients have been discussed in Section 2 of this Notebook. Tables 3.5, 3.6, and 3.7 contain diffusion coefficients, D, of impurities in silicon, germanium, and gallium arsenide calculated from the expression:

$$D = D_o e^{-\frac{\Delta H}{RT}},$$

where D_o is a constant for the impurity in the crystal, and ΔH is the activation energy for diffusion.

Downloaded from https://www.everyspec.com

TABLE 3.2.. ENERGY LEVELS OF CHEMICAL IMPURITIES - GERMANIUM

Impurity	Ionization Energy[1], ev $E_C - E_I$	Ionization Energy[1], ev $E_I - E_V$	Conductivity Type[1]	Maximum Solubility[2, Except P], atoms/cm^3
Li	0.0093		D	7×10^{18}
Cu		0.040	A	3.5×10^{16}
Cu		0.31	A	--
Cu	0.26		A	--
Au	0.04		A	2×10^{16}
Au	0.20		A	--
Au		0.15	A	--
Au		0.053	D	--
Zn		0.029	A	2.5×10^{18}
B		0.0104	A	--
Al		0.0102	A	4×10^{20}
Ga		0.0108	A	5×10^{20}
In		0.0112	A	4×10^{18}
Tl		0.014	A	--
P	0.0120		D	1×10^{20}[3]
As	0.0127		D	2×10^{20}
Sb	0.0096		D	1×10^{19}
Bi	0.012		D	--
Mn	0.35		A	--
Mn		0.16	A	--
Fe	0.27		A	--
Fe		0.34	A	1.5×10^{15}
Co	0.31		A	--
Co		0.25	A	2×10^{15}
Ni	0.30		A	--
Ni		0.22	A	8×10^{15}
Pt	0.20		A	--
Pt		0.04		5×10^{14}
Te	0.10		D	--
Se	0.14		D	--
S	0.18		D	--
Ag	0.29		A	--
Ag	0.09		A	--
Ag		0.13	A	--
Zn		0.09	D	--
Cd		0.05	D	--
Cd		0.16	D	--

Downloaded from https://www.everyspec.com

TABLE 3.3. ENERGY LEVELS OF CHEMICAL IMPURITIES - SILICON

Impurity	Ionization Energy[4], ev $E_C - E_I$	Ionization Energy[4], ev $E_I - E_V$	Conductivity Type[4]	Maximum Solubility[2, Except In], atoms/cm^3
Li	0.033		D	6×10^{19}
Cu		0.49	A	--
Cu		0.24	D	1.3×10^{18}
Au		0.33	D	1.2×10^{17}
Au	0.54		A	--
Zn	0.55		A	6×10^{16}
Zn		0.30	A	--
B		0.045	A	6×10^{20}
Al		0.057	A	2×10^{19}
Ga		0.065	A	4×10^{19}
In		0.16	A	6.7×10^{20}(3)
P	0.044		D	1.3×10^{21}
As	0.049		D	2×10^{21}
Sb	0.039		D	7×10^{19}
Mn	0.53		D	3.5×10^{16}
Fe	0.55		D	--
Fe		0.40	D	3×10^{16}
S	0.18		D	3.6×10^{16}

TABLE 3.4. ENERGY LEVELS OF CHEMICAL IMPURITIES - GALLIUM ARSENIDE

Impurity	Ionization Energy[5], ev	Conductivity Type[5]	Maximum Solubility[6], atoms/cm^3
Si	0.005	D	4×10^{18}
Cu	0.42	A	--
Cu	0.14	A	--
Cd	0.021	A	$>10^{19}$
Zn	0.014	A	$>10^{20}$

Downloaded from https://www.everyspec.com

REFERENCES, TABLES 3.2, 3.3, AND 3.4

(1) Bube, R. H., Photoconductivity of Solids, John Wiley and Sons, Inc., New York (1960), p 136.

(2) Trumbore, F. A., "Solid Solubilities of Impurity Elements in Germanium and Silicon", Bell System Tech. J., 39, 205 (1960).

(3) Tables of Constants and Numerical Data Affiliated to the International Union of Pure and Applied Chemistry, 12. Selected Constants Relative to Semiconductors, Pergamon Press (1961), p 10.

(4) Bube, R. H., loc. cit., p 137.

(5) Neuberger, M., "Gallium Arsenide Data Sheets", EPIC Report No. DS-115 (November, 1962).

(6) Ibid., 30.

Downloaded from https://www.everyspec.com

TABLE 3.5. DIFFUSION CONSTANTS OF IMPURITIES – GERMANIUM[a]

Diffusant	D_o, cm^2/sec	$-\Delta H$, ev	D, cm^2/sec 300°K	500°K	700°K	References
B	1.8×10^{9}	4.55	1.0×10^{-67}	1.7×10^{-37}	3.6×10^{-24}	1
Al	--	--	--	--	--	--
Ga	20	3.04	3.2×10^{-50}	6.2×10^{-30}	3.2×10^{-21}	1
In	20	3.0	1.3×10^{-49}	4.5×10^{-30}	5.4×10^{-21}	1
P	12.	2.6	3.9×10^{-43}	6.0×10^{-26}	2.5×10^{-18}	1
As	2.12	2.4	1.5×10^{-40}	1.1×10^{-24}	1.2×10^{-17}	1
Sb	1.3	2.26	2.1×10^{-38}	1.7×10^{-23}	7.3×10^{-17}	1
Ta	--	--	--	--	--	--
Bi	4.7	2.4	3.4×10^{-40}	3.5×10^{-24}	2.6×10^{-17}	1
H	--	--	--	--	--	
Li	0.0025	0.51	7.5×10^{-12}	1.6×10^{-12}	5.5×10^{-7}	1
Cu	--	--	$D = 3.1 \times 10^{-5}$ (T = 1173°K)			1
Ag	0.044	1.0	7.9×10^{-19}	3.2×10^{-12}	3.0×10^{-9}	1
Au	12.6	2.25	3.0×10^{-37}	2.3×10^{-22}	8.4×10^{-16}	1
Be	--	--	--	--	--	
Zn	0.65	2.5	2.9×10^{-44}	3.5×10^{-26}	7.2×10^{-19}	1
Cd	1.75×10^{9}	4.4	4.6×10^{-64}	5.6×10^{-33}	4.2×10^{-23}	1
O	--	--	--	--	--	--
Cr	--	--	--	--	--	--
Fe	0.13	1.1	5.3×10^{-20}	9.4×10^{-13}	1.6×10^{-9}	1
Ni	0.8	0.91	4.9×10^{-16}	4.8×10^{-10}	2.4×10^{-7}	1

(a) -- indicates that the datum is not yet available.

Downloaded from https://www.everyspec.com

TABLE 3.6. DIFFUSION CONSTANTS OF IMPURITIES – SILICON[a]

Diffusant	D_o, cm^2/sec	$-\Delta H$, ev	D, cm^2/sec 300°K	500°K	700°K	References
B	10.5	3.6	6.2×10^{-58}	4.1×10^{-35}	1.4×10^{-24}	1
Al	8.0	3.4	1.0×10^{-56}	3.2×10^{-34}	2.9×10^{-24}	1
Ga	3.6	3.5	1.0×10^{-57}	1.4×10^{-35}	2.4×10^{-25}	1
In	16.5	4.0	2.1×10^{-64}	1.2×10^{-37}	3.0×10^{-27}	1
Tl	16.5	4.0	2.1×10^{-64}	1.2×10^{-37}	3.0×10^{-27}	1
P	10.5	3.6	6.2×10^{-58}	4.1×10^{-35}	1.4×10^{-24}	1
As	0.32	3.6	1.9×10^{-61}	1.2×10^{-38}	4.2×10^{-27}	1
Sb	5.6	3.9	3.2×10^{-63}	4.1×10^{-37}	6.7×10^{-29}	1
Ta	--	--	--	--	--	--
Bi	1030	4.7	2.4×10^{-81}	3.0×10^{-50}	1.6×10^{-37}	2
H	9.4×10^{-3}	--	--	--	--	3
Li	0.0023	0.66	2.1×10^{-14}	5.1×10^{-10}	4.1×10^{-8}	1
Cu	--	--	$D = 5.5 \times 10^{-5}$ (T = 1173°K)			1
Ag	--	--	--	--	--	--
Au	0.0011	1.1	4.95×10^{-22}	8.0×10^{-15}	1.3×10^{-11}	1
Be	--	--	--	--	--	--
Zn	--	--	$1.1 \times 10^{-6} - 1.1 \times 10^{-7}$ (1237°K < T < 1573°K)			1
Cd	--	--	--	--	--	--
C	--	2.8 ± 0.8	$D = 10^{-8}$ (T = 1640°K)			4
Cr	--	--	--	--	--	--
Fe	0.0062	0.87	1.7×10^{-17}	9.9×10^{-12}	3.3×10^{-9}	1
Ni	--	--	--	--	--	--

(a) -- indicates that the datum is not yet available.

Downloaded from https://www.everyspec.com

TABLE 3.7. DIFFUSION CONSTANTS OF IMPURITIES – GALLIUM ARSENIDE[a]

Diffusant	D_o	$-\Delta H$, ev	D, cm^2/sec 300°K	500°K	700°K	References
Zn	15	2.49	4.4×10^{-42}	1.0×10^{-24}	1.8×10^{-17}	5
Cd	0.05	2.43	1.1×10^{-42}	1.4×10^{-26}	1.7×10^{-19}	7
O	--	--	--	--	--	--
S	4×10^{3}	4.04	1.1×10^{-63}	5.2×10^{-38}	3.6×10^{-26}	5
Cr	--	--	--	--	--	--
Se	3×10^{3}	4.06	7.8×10^{-66}	2.6×10^{-39}	3.6×10^{-27}	5
Al	--	--	--	--	--	--
Ga	1×10^{7}	5.60	2.0×10^{-85}	2.4×10^{-50}	5.3×10^{-34}	5
In	--	--	--	--	--	--
As	4×10^{21}	10.2	9.6×10^{-147}	2.8×10^{-80}	1.9×10^{-51}	5
Ta	--	--	--	--	--	--
Si	--	--	--	--	--	--
Sn	6×10^{-4}	2.5	8.4×10^{-46}	3.2×10^{-29}	6.6×10^{-22}	7
H	--	--	--	--	--	--
Li	--	--	--	--	--	--
Cu	--	--	2.4×10^{-6} (T = 1373°K)			7
Au	$\sim 10^{-3}$	1.0	--	--	--	6
Fe	--	--	--	--	--	--
Ni	--	--	--	--	--	--

(a) -- indicates that the datum is not yet available.

Downloaded from https://www.everyspec.com

REFERENCES, TABLES 3.5, 3.6, AND 3.7

(1) Hannay, N. B., Semiconductors, Reinhold Publishing Corporation, (1959), Table 6.4, pp 245-246.

(2) Schulz, M., "Silicon: Semiconductor Properties", Infrared Phys., 4, 93 (1964).

(3) Hannay, N. B., loc. cit., p 235.

(4) Hannay, N. B., loc. cit., p 253.

(5) Hilsum, C., and Rose-Innes, A. C., Semiconducting III-V Compounds, Pergamon Press (1961), p 83.

(6) Sokolov, V. I., and Shilshiyanu, F. S., "Diffusion and Solubility of Gold in Gallium Arsenide", Soviet Phys., Solid State, 6, 265 (1964).

(7) Tables of Constants and Numerical Data Affiliated to the Union of Pure and Applied Chemistry, 12. Selected Constants Relative to Semiconductors, Pergamon Press (1961), p 30.

Downloaded from https://www.everyspec.com

3.4 Eutectic Melting Points

Table 3.8 contains eutectic points for various combinations of metals and semiconductors that might normally come in contact during use. A eutectic point is the concentration of one material in another at which the melting point of the composition is at a minimum.

TABLE 3.8. EUTECTIC MELTING POINTS[1]

Material A	Concentration, Material A, atomic per cent	Material B	Concentration, Material B, atomic per cent	Eutectic Temp, °C
Si	99	Ta	1	1385
Si	11.3	Al	88.7	577
Si	31	Au	69	370
Si	56.2	Ni	43.8	966
Ge	27	Au	73	356
Ge	30.3	Al	69.7	424
Al	99.3	Au	0.7	642

REFERENCE, TABLE 3.8

(1) Hansen, M., Constitution of Binary Alloys, McGraw-Hill Book Company, Inc., New York (1958).

Downloaded from https://www.everyspec.com

3.5 Vapor Pressure

Table 3.9 presents the vapor pressure at different temperature levels for numerous elements used in the construction of electronic devices.

TABLE 3.9. VAPOR PRESSURE[1]

Material	Vapor Pressure, mm Hg, at Indicated Temperature		
	300°K	500°K	700°K
Ge	$<10^{-11}$	$<10^{-11}$	$<10^{-11}$
Si	$<10^{-11}$	$<10^{-11}$	$<10^{-11}$
Ga	$<10^{-11}$	$<10^{-11}$	10^{-11}
As	$<10^{-11}$	6×10^{-4}	7
Cd	--	1.5×10^{-10}	2.5×10^{-5}
S	4×10^{-6}	5	450
Se	9×10^{-11}	4×10^{-3}	7
Al	$<10^{-11}$	$<10^{-11}$	$<10^{-11}$
Au	$<10^{-11}$	$<10^{-11}$	$<10^{-11}$
In	$<10^{-11}$	$<10^{-11}$	4×10^{-10}
Pb	$<10^{-11}$	$<10^{-11}$	10^{-6}

REFERENCE, TABLE 3.9

(1) "Vapor Pressure Curves for the More Common Elements", prepared by R. E. Honig, Radio Corporation of America, Copyright, 1962.

Downloaded from https://www.everyspec.com

3.6 Intrinsic Carrier Concentration

Figure 3.1 presents plots of intrinsic carrier concentration as a function of absolute temperature of silicon and germanium. The plots were made using the following expression to calculate the intrinsic concentration, n_i:

$$n_i = p_i = 2\left(\frac{2\pi kT}{h^2}\right)^{3/2} (m_e^* m_h^*)^{3/4} \exp \frac{E_g}{2kT} ,$$

where

h = Planck's constant, 6.624×10^{-34} joule/sec

k = Boltzmann's constant, 1.38×10^{-23} joule/deg

m_e^* = effective mass of electrons

m_h^* = effective mass cf holes

E_g = energy gap.

In the plotted values, the temperature dependence of the energy gap has been accounted for.

The intrinsic carrier concentration of a semiconductor is the number of electrons per unit volume that have been excited from the valence band to the conduction band. An equal number of holes is left in the valence band by the excited electrons. The properties of a semiconductor device are usually determined by the number of donor and acceptor impurities added during fabrication. At a temperature where the intrinsic concentration exceeds the donor or acceptor concentration, the distinctive properties of the device can be expected to disappear.

Downloaded from https://www.everyspec.com

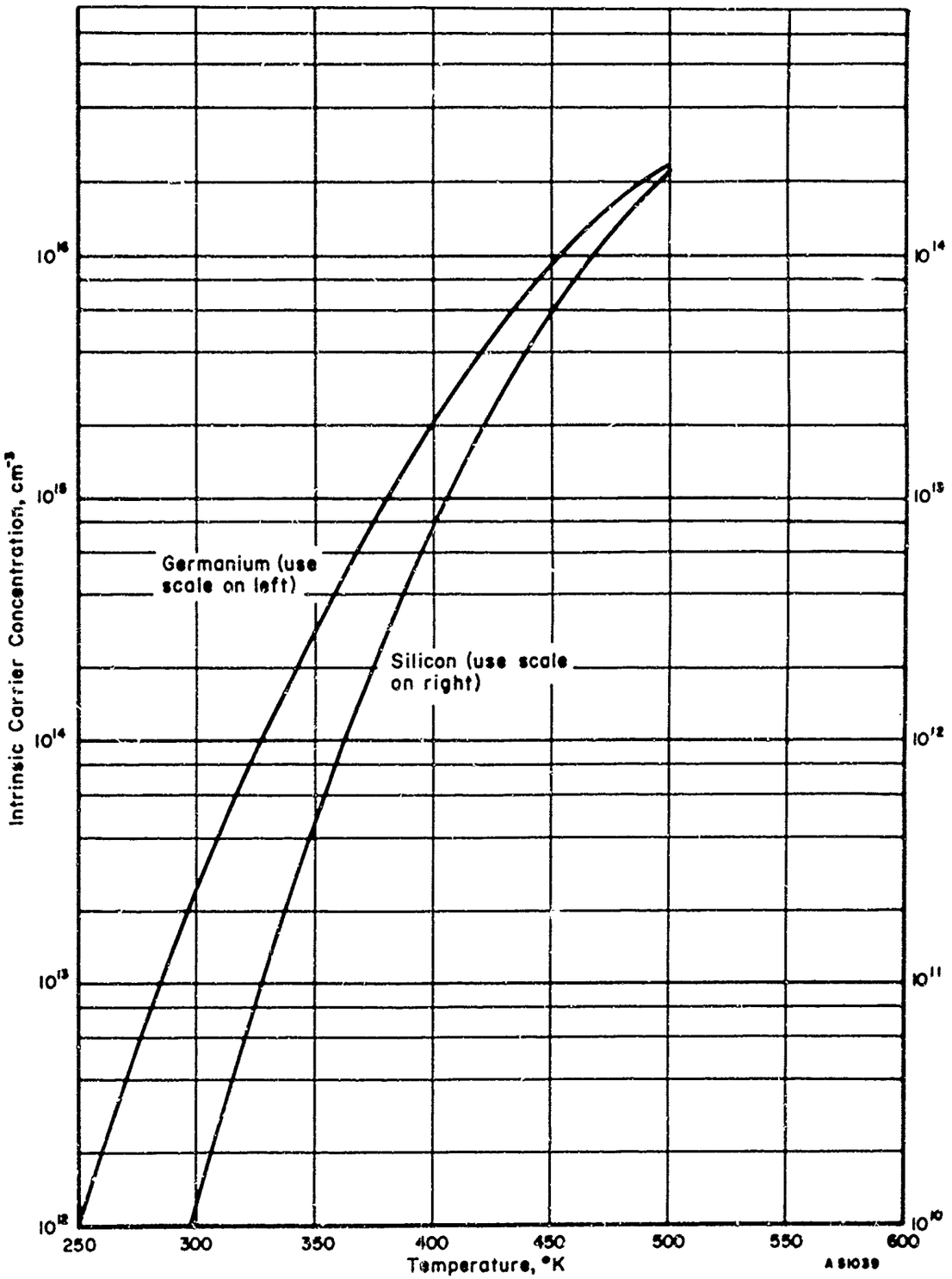


FIGURE 3.1. INTRINSIC CARRIER CONCENTRATION – GERMANIUM AND SILICON

Downloaded from https://www.everyspec.com

3.7 Thermal Conductivity

Table 3.10 lists thermal conductivity of many materials used in construction of electronic devices. Note that the units are watts/cm-deg.

TABLE 3.10. THERMAL CONDUCTIVITY

Material	Thermal Conductivity, w/cm-deg, at Indicated Temperature		
	300°K	500°K	700°K
Si	1.422[1]	0.692[1]	0.483[1]
Ge	0.521[2]	0.298[2]	0.209 (673°K)[2]
GaAs	0.44[3]	0.25[3]	0.14[3]
CdS	0.159[4]	--	--
SiO_2	0.0118 (fused silica at 40°C)[5]	--	--
Al_2O_3	0.2511[6]	0.039 (522°K)[6](a)	0.051 (657°K)[6](a)
Soda-lime glass	0.0092[8]	--	--
Borosilicate crown glass	0.011 (273°K)[9]	--	--
Aluminosilicate	0.013[10]	0.015 (573°K)[10]	--
Glazed alumina	0.352[11](a)	--	--

(a) Heat flow perpendicular to C-axis.

Downloaded from https://www.everyspec.com

REFERENCES, TABLE 3.10

(1) Shanks, H. R., Maycock, P. D., Sidles, P. H., and Danielson, G. C., "Thermal Conductivity of Si From 300 - 1400 K", Phys. Rev., 130, 1743 (1963).

(2) Stuckes, A. D., "The Thermal Conductivity of Germanium, Silicon, and Indium Arsenide From 40°C to 425°C", Phil. Mag., 5, 84 (1960).

(3) Timberlake, A. B., Davis, P. W., and Shilliday, T. S., "Thermal Diffusivity Measurements on Small Samples", Adv. Energy Conv., 2, 45 (1962).

(4) Ballard, S. S., McCarthy, K. A., and Wolfe, W. L., "IRIA, State-of-the-Art Report, Optical Materials for Infrared Instrumentation", University of Michigan, Ann Arbor, Michigan, 89 (1959).

(5) Ballard, S. S., loc. cit., 20.

(6) American Institute of Physics Handbook, McGraw-Hill Book Company, Inc., New York (1957), pp 4-73.

(7) Nixen, D., "Materials for Packaging Electronic Devices", Electron. Ind., 66 (January, 1965).

(8) Shand, E. B., Glass Engineering Handbook, Corning Glass Works, Corning, New York (1955), p 14.

(9) Morey, G. W., The Properties of Glass, Second Edition, Reinhold Publishing Corporation, New York (1954), p 218.

(10) Settzo, R. J., "Substrates for Deposited Film Passive Components", Electronics Reliability and Microminiaturization, 1, 347 (1962).

(11) Smith, P. C., and Genser, M., "An Investigation of Thin Film Resistor Failure", Physics of Failure in Electronics, 3, to be published.

Downloaded from https://www.everyspec.com

3.8 Emissivities of Various Materials

Emissivities of various materials are given in Table 3.11. Although the available data are sparce, it is felt that this information will be required to appropriately evaluate infrared scan pictures of electronic equipment, since the energy radiated by a body is a function of the emissivity of the body.

TABLE 3.11. EMISSIVITY

Material	Emissivity, ϵ, at Indicated Temperature[a]		
	300°K	500°K	700°K
Si	--	--	--
Ge	0.3[1]	--	--
GaAs	--	--	--
CdS	--	--	--
Au	0.02-0.03	0.02-0.03	0.02-0.03[2]
Al	0.03	0.04[2]	--
Nichrome (oxidized)	0.90 (λ = 0.65 μ)[3]	--	--
SiO	--	--	--
SiO_2	--	--	--

(a) -- indicates that the datum is not available.

REFERENCES, TABLE 3.11

(1) Abeles, B., Cody, G. D., and Beers, D. S., "Apparatus for Measurement of Thermal Diffusivity of Solids at High Temperatures", J. Appl. Phys., 31, 1585 (1960).

(2) American Institute of Physics Handbook, First Edition, McGraw-Hill Book Company, Inc., New York (1957), pp 6-68.

(3) Ibid., pp 6-74.

Downloaded from https://www.everyspec.com

3,9 Thermal Coefficient of Linear Expansion

The thermal coefficient of linear expansion of a material is used to calculate the change in length of a body that takes place upon heating or cooling. It is defined as:

$$\alpha = \frac{\Delta L}{L_o \Delta T} ,$$

where

ΔL is the change of length of the body

L_o is the initial length of the body

ΔT is the temperature difference over which the measurement was performed.

Thermal expansion coefficients are given for a number of materials in Table 3.12.

Downloaded from https://www.everyspec.com

TABLE 3.12. THERMAL COEFFICIENT OF LINEAR EXPANSION

Material	Thermal Coefficient of Linear Expansion, α, 10^{-6} °C^{-1}	Temperature, °K	Reference
	Metals		
Al	23.8	293-373	1
Al	25.7	293-573	1
Al	28.7	293-873	1
Au	14.3	290-373	1
In	41.7	313	1
2Pb-1Sn	25.08	273-373	1
Ni	12.79	313	1
Ni-Cr	13.5	300	6
Pt	8.99	313	1
Ag	18.8	293	1
Ta	6.7	293-373	1
	Glasses and Ceramics		
0211 Microsheet[a]	7.2	273-573	2, 3
Alkali	--	--	--
Borosilicate	3.25	273-573	2, 3
96% silicate (Vycor[a])	0.8	273-573	3
Pyrex	3.6	294-744	1
Soda-lime glass	8.5-9.2	273-573	4
Quartz	7.97 p[b]	273-353	1
Quartz	13.37 s[b]	273-353	1
94% alumina	6.7	300	2
Sapphire (Al_2O_3)	6.7 p	323	5
	5.0 s	--	--
SnO	4.6	--	6
MgO	13.8	293-1373	5
SiO	2	--	--
BeO (99.5%)	6.0	--	6
	Semiconductors		
Si	4.2	300	5
Ge	5.5-6.1	300	5
GaAs	5.7	--	5
CdS	4.2	300-343	5
CdTe	4.5	323	5

(a) Trademark, Corning Glass Works, Corning, New York.

(b) p indicates measurement made parallel to C-axis.
s indicates measurement made perpendicular to C-axis.

Downloaded from https://www.everyspec.com

REFERENCES, TABLE 3.12

(1) Handbook of Chemistry and Physics, 37th Edition (HCP), Chemical Rubber Publishing Company, Cleveland, Ohio (1955), pp 2059-2069.

(2) Settzo, R. J., "Substrates for Deposited Film Passive Components", Electronics Reliability and Microminiaturization, 1, 347 (1962).

(3) Corning Glass Works Bulletin: CEP 2/5.M/9-62.

(4) Shand, E. B., Glass Engineering Handbook, Corning Glass Works, Corning, New York (1955), p 9.

(5) Ballard, S. S., McCarthy, K., and Wolfe, W., "IRIA State-of-the-Art Report, Optical Materials for Infrared Instrumentation", University of Michigan, Ann Arbor, Michigan, 16 (1959).

(6) Hunter, J., "The Search for New Materials", Electronic Industries, 23 (12), 38 (December, 1964).

Downloaded from https://www.everyspec.com

3.10 Elastic Constants

The most commonly referred to elastic properties of materials are Young's modulus, E; bulk modulus, k; shear modulus or modulus of rigidity, n; and Poisson's ratio, σ. More basic than these, however, are the constants of proportionality in a crystal obeying Hooke's law between components of stress and components of strain, known as elastic-stiffness constants. In cubic crystals, there are three such constants in terms of which the three elastic moduli can be calculated. These are designated C_{11}, C_{12}, and C_{44}. The expressions relating elastic moduli to stiffness constants are[1]

$$E = \frac{(C_{11} + 2\,C_{12})\,(C_{11} - C_{12})}{(C_{11} + C_{12})} \quad ,$$

$$n = C_{44} \quad ,$$

$$k = \frac{C_{11} + 2\,C_{12}}{3} \quad ;$$

and for homogenous, isotropic bodies[2],

$$\sigma = \frac{E}{2n} - 1 \text{ and}$$

$$\sigma = \frac{3k - 2n}{6k + 2n} \quad .$$

The values of C_{11}, C_{12}, C_{44}, E, n, k, and σ for many electronic materials are given in Table 3.13. The values of the moduli are not calculated from stiffness coefficients.

Downloaded from https://www.everyspec.com

TABLE 3.13. ELASTIC COEFFICIENTS

Material	C_{11}, 10^{12} dynes/cm^2	C_{12}, 10^{12} dynes/cm^2	C_{44}, 10^{12} dynes/cm^2	Young's Modulus, E, 10^{12} dynes/cm^2	Shear Modulus, n, 10^{12} dynes/cm^2	Bulk Modulus, k, 10^{12} dynes/cm^2	σ
Si	1.67[3]	0.65[3]	0.80[3]	--	--	1.02[3]	--
Ge	1.29[4]	0.483[4]	0.671[4]	--	--	0.773[4]	--
GaAs	0.188[5]	0.594[5]	0.538[5]	--	--	--	--
CdS	0.843[6]	0.521[6]	0.149[6]	--	--	--	--
CdTe	0.5351[7]	0.3681[7]	0.1994[7]	--	--	--	--
Au	--	--	--	0.785[8]	--	--	--
Ni	--	--	--	2.001-2.136[8]	0.706-0.755[8]	--	--
Cr	--	--	--	2.48[8]	--	--	--
Al, rolled	--	--	--	0.68-0.70[8]	--	--	--
Ta	--	--	--	1.86[8]	--	--	--
SiO	--	--	--	--	--	--	--
SiO_2	--	--	--	7.39[11]	0.312[11]	3.71[11]	--
Al_2O_3 (sapphire)	--	--	--	3.45[9]	1.48[9]	0.0807[9]	--
MgO	2.90[10]	0.876[10]	1.55[10]	--	--	--	--
Soda-lime glass	--	--	--	0.725[12]	0.310[12]	--	0.21[12]
Borosilicate glass	--	--	--	0.503[12]	0.214[12]	--	0.22[12]
96% silica (Vycor[a])	--	--	--	0.662[12]	0.283[12]	--	0.18[12]
0211 Microsheet[a]	--	--	--	0.745[13]	0.302[13]	--	0.23[13]

(a) Trademark, Corning Glass Works, Corning, New York.

Downloaded from https://www.everyspec.com

REFERENCES, SECTION 3.10 AND TABLE 3.13

(1) Ballard, S., McCarthy, K., and Wolfe, W., "IRIA, State-of-the-Art Report, Optical Materials for Infrared Instrumentation", University of Michigan, Ann Arbor, Michigan, 8 (1959).

(2) Stephenson, R. J., Mechanics and Properties of Matter, John Wiley and Sons, Inc., New York (1952), p 241 ff.

(3) Ballard, et al., loc. cit., 50.

(4) Ballard, et al., loc. cit., 52.

(5) Tables of Constants and Numerical Data, 12. Selected Constants Relative to Semiconductors, Pergamon Press (1961), p 27.

(6) Ibid., p 22.

(7) Thomas, D. G., "Excitons and Band Splitting Produced by Uniaxial Stress in CdTe", J. Appl. Phys., Suppl. to 32 (10), 2298 (1961).

(8) Handbook of Chemistry and Physics, 37th Edition (HCP), Chemical Rubber Publishing Company, Cleveland, Ohio (1955), p 1985.

(9) Ballard, et al., loc. cit., 24.

(10) Ballard, et al., loc. cit., 26.

(11) Ballard, et al., loc. cit., 20.

(12) Shand, E. B., Glass Engineering Handbook, Corning Glass Works, Corning, New York (1955), p 18.

(13) Corning Glass Works, Corning, New York, Product Information, May 1, 1962.

Downloaded from https://www.everyspec.com

3.11 Dielectric Properties

Three properties of materials usually referred to as dielectric properties are dielectric constant, K; loss tangent, also known as dielectric dissipation factor; and dielectric strength. These properties are of importance in a material being considered for use as a capacitor dielectric, a thin-film circuit substrate, a printed-circuit board, a transistor or diode mount, or an incapsulant.

Dielectric constant is a dimensionless quantity, which, when multiplied by permittivity of free space, ϵ_o, gives the permittivity ϵ of the material.

Dielectric strength is the ability of a material to withstand electric stresses that cause either electronic or thermal breakdown. Electronic breakdown takes place when the voltage gradient in a material exceeds some limit that is an intrinsic property of the material. In such a case, electrons are accelerated to velocities that enable them to ionize atoms upon collision, and current avalanching occurs. Thermal breakdown occurs when electric losses in the material cause localized heating, which in turn causes the flow of current to increase further. Dielectric strength is the field strength a material can withstand before electronic breakdown takes place. To prevent excessive heating and the resultant thermal breakdown, this quantity must be measured while applying the electric field for times of 1 second or less.(1)

The loss tangent of a material provides a measure of the amount of a-c energy impressed on a dielectric material that is lost in absorption phenomena and conduction losses. The magnitude of such losses, W, is given for sinusoidal voltages by(1)

$$W = \frac{0.555\ \epsilon \tan \delta}{1 + \tan^2 \delta}\ fV^2 \times 10^{-6} \quad ,$$

where

W = watts lost per cm^3

ϵ = dielectric constant of the material

f = frequency in cps

V = voltage gradient in kv/cm, rms

δ = loss angle of the material.

The quantity tan δ is called the dielectric dissipation factor, and, for small values of δ, is equal to the dielectric power factor.(2)

Table 3.14 lists dielectric constant, breakdown strength, and loss tangent of many dielectrics and semiconductors.

Downloaded from https://www.everyspec.com

TABLE 3.14. DIELECTRIC PROPERTIES

Material	Dielectric Constant	Dielectric Strength, V/cm	Loss Tangent
Si	13(3)(a)	--	--
Ge	16.6(4)(b)	--	--
GaAs	11.6 ± 0.14(5)	--	--
CdS	(||)7.96(6)(c)	--	--
CdS	(⊥)8.82	--	--
SiO	6.8(7)	1-3 x 10^6(6)	--
SiO_2, fused silica	3.78(8)(d)	--	0.00025(1)
Al_2O_3	(||)10.55(8)(e)	1.6-6.4 x 10^5	--
Al_2O_3	(⊥)8.6(d)	--	--
Soda-lime glass	6.9(9)(c)	3.2 x 10^6(1)(g)	8.01(9)(c)
Soda-lime glass	9.3(9)(f)	7.5 x 10^2(9)(h)	0.17(9)(f)
0211 Microsheet(i)	6.6(9)(c)	3.4 x 10^6(1)(g)	0.0047(9)(c)
0211 Microsheet(i)	7.4(9)(f)	10^4(9)(h)	0.032(9)(f)
96% silica (Vycor(i))	3.9(9)(c)	3.5 x 10^6(1)(g)	0.0006(9)(c)
96% silica (Vycor(i))	3.9(9)(f)	7 x 10^3(9)(h)	0.001(9)(f)
98.5% BeO	6.9(10)(j)	--	0.0008(10)(j)

(a) Measured at 10^{10} cps.
(b) Measured at 9.4 x 10^9 cps.
(c) Measured at 10^6 cps.
(d) Measured from 10^2 to 10^{10} cps.
(e) Measured from 10^2 to 10^8 cps.
(f) Measured at 10^6 cps, 473°K.
(g) 60 cycle, rms; 300°K.
(h) 60 cycle, rms; 623°K, for 1 minute.
(i) Trademark, Corning Glass Works, Corning, New York.
(j) Measured at 4.45 to 4.04 x 10^9 cps.

Downloaded from https://www.everyspec.com

REFERENCES, SECTION 3.11 AND TABLE 3.14

(1) Morey, G. W., The Properties of Glass, Second Edition, Reinhold Publishing Corporation, New York (1954), p 529.

(2) Shand, E. B., Glass Engineering Handbook, Corning Glass Works, Corning, New York (1955), p 36.

(3) Ballard, S., McCarthy, K., and Wolfe, W., "IRIA, State-of-the-Art Report, Optical Materials for Infrared Instrumentation", University of Michigan, Ann Arbor, Michigan, 50 (1959).

(4) Ballard, et al., loc. cit., 52.

(5) Ballard, et al., loc. cit., 90.

(6) Neuberger, M., "Cadmium Sulfide Data Sheets", EPIC Report No. DS-124, 231 (April, 1963).

(7) Degenhart, H. J., and Pratt, I., "Preparation and Evaluation of Vacuum Deposited Thin Film Capacitors", USAERDL, Fort Monmouth, New Jersey, July, 1961.

(8) Ballard, et al., loc. cit., 13.

(9) Bulletin: CEP 2/5M/9-62, Corning Glass Works, Corning, New York.

(10) Massachusetts Institute of Technology, Laboratory for Insulation Research, Technical Documentary Report ML TDR 64-219, Technical Report 191, by A. von Hippel, et al., for AF Materials Laboratory, Research and Technology Division, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio, July, 1964, p 58.

Downloaded from https://www.everyspec.com

3.12 Properties of Solder

Table 3.15 presents electrical, thermal, and mechanical characteristics of numerous low-melting alloys used as solders.

TABLE 3.15. PROPERTIES OF SOLDERS

Composition	Electrical Conductivity, 10^5 (ohm-cm)$^{-1}$	Short-Time Tensile Strength, 10^8 dynes/cm^2	Plastic Temp, °K	Liquid Temp, °K	Reference
100 Tin	0.85	1.24	505	505	1
100 Lead	0.45	1.23	600	600	1
95 Tin-5 Antimony	0.73	4.07	505	515	1
95 Tin-5 Silver	0.79	9.66	486	494	1
90.5 Tin-9.5 Bismuth	0.57	--	--	--	1
70 Tin-30 Lead	0.68	4.24	456	459	1
63 Tin-37 Lead	0.62	4.32	456	456	1
60 Tin-40 Lead	0.62	4.34	456	461	1
97.5 Lead-1.5 Silver -1 Tin	0.42	2.48	577	581	1
97.5 Lead-2.5 Silver	0.46	2.48	577	580	1
100 Indium	1.36	--	--	156.4	1
50 Indium-50 Tin	0.68	0.354	--	429	1
90 Indium-10 Silver	1.24	1.42	--	503	1
50 Indium-50 Lead	0.34	2.11	--	489	1
55.5 Tin-3.4 Antimony -41.1 Lead	--	--	--	460	2
80 Gold-20 Tin	--	--	--	553	3
88 Gold-12 Germanium	--	--	--	629	3
94 Gold-6 Silicon	--	--	--	643	3

Downloaded from https://www.everyspec.com

REFERENCES, TABLE 3.15

(1) Haufman, A., "Soft Solder Conductivity", Electronics, 119 (October 23, 1959).

(2) McKeown, J., The Properties of Soft Solders and Soldered Joints, British Non-Ferrous Metals Research Association, London (1948), p 15.

(3) Nixen, D., "Materials for Packaging Microelectronic Devices", Electronic Industries, 66 (January, 1965).

Downloaded from https://www.everyspec.com

3.13 Properties of Encapsulants and Potting Components

The encapsulant of an electronic device provides electrical insulation and protection from water vapor and other contaminants, and helps to protect the device from damage caused by mechanical shocks and stresses. Therefore, the electrical properties of the encapsulating material – d-c resistance, dielectric strength, and dielectric constant and loss tangent over a wide range of frequencies – should be known. Furthermore, the resistance of the material to penetration by water vapor is important, not only because water vapor damages many electronic devices but also because it reduces the specific resistivity of many encapsulating materials.

Because the encapsulant of an electronic device will necessarily be subjected to the same thermal and mechanical stresses as the electronic device itself, properties such as thermal coefficient of expansion and ultimate tensile strength, are of importance.

Table 3.16, taken from Reference 1, lists the properties of many encapsulants. The physical meanings of most of the properties listed are assumed to be well known. Those that may not be well known are defined as follows:

(1) Arc-Tracking Resistance. This is a measure of the ability of the material to resist localized degradation when arcs and flashovers occur. Arcs and flashovers may cause a chemical reaction to take place in their paths, leaving a track of relatively low-resistance carbonaceous material of lower dielectric strength.

(2) Safe-Use Temperature. This is a characteristic rather loosely defined as the temperature a material will withstand for short periods of time without obvious degradation, such as charring, blistering, or gross distortion.

(3) Heat Distortion Temperature. This is the maximum temperature at which flow or deformation of the material under specified conditions of load and time will not exceed a specified value.

Table 3.17 presents solvents for uncured material and trade names and manufacturers of many encapsulants at the time of publication of Reference 1.

TABLE 3.16. TYPICAL CURED PROPERTIES OF BASIC ENCAPSULANTS

Polymer System	Electrical[a]						Thermal			Physical[a]		
	Dielectric Strength, volts/mil	Specific Resistivity, ohm-cm	Water Absorption[a], %	Dielectric Constant, 10^{10} cps	Loss Tangent, 10^{10} cps	Relative Arc-Track Resistance[c]	Heat Distortion Temperature, °C	Safe-Use Temperature, °C	Linear Expansion, x 10^{5}/°C	Ultimate Tensile Strength, psi	Ultimate Elongation, %	Hardness[d]
Thermoplastics												
Asphalt and tars[e]	300	10^{10}	0.08	3.5	0.04	5	54	71	11.2	600	5	SD60
Fluorocarbon[e]	450	10^{18}	0.00	2.1	0.0003	1	121	260	7.7	3000	200	SD60
Polyethylene[e]	500	10^{16}	<0.01	2.3	0.0005	3	--	116	13.3	4000	1000	SD65
Polystyrene[e]	550	10^{18}	0.04	2.5	0.0003	3	82	86	5.6	7000	1.5	M80
Polyvinyl chloride[e]	400	10^{15}	0.15	2.8	0.006	--	66	99	4.2	3000	100	SD80
Wax[e]	400	10^{17}	0.02	2.6	0.001	3	27	57	15.4	300	5	SD30
Thermosets												
Alkyd[e]	350	10^{14}	0.4	3.8	0.025	2	104	121	5.6	8000	--	SD90
Allyl ester[e]	400	10^{14}	0.7	--	--	3	>88	99	5.6	5500	--	M70
Butadienestyrene[e]	600	10^{16}	0.03	2.4	0.006	3	127	246	7.0	4000	4	SD80
Epoxide[e]	450	10^{14}	0.20	2.9	0.018	2	204	232	4.2	10000	<1	M90
Phenolaldehyde[e]	350	10^{12}	0.3	4.7	0.04	4	80	80	5.6	7000	1.5	M125
Polyester[e]	350	10^{13}	0.4	3.5	0.05	3	88	163	8.4	8000	<5	M100
Silicone[e]	350	10^{15}	0.03	2.8	0.002	2	38	260	9.8	2500	8	M60
Elastomers												
Buna-S rubber	500	10^{14}	--	2.5	0.01	4	--	121	8.4	300	400	SA50
Chloro-rubber	400	10^{12}	--	2.7	0.05	3	--	--	12.6	2500	500	SA70
Natural rubber	500	10^{16}	--	2.1	0.03	4	--	66	5.6	3000	700	SA50
Silicone rubber[e]	600	10^{13}	--	3.0	0.05	2	--	260	--	650	100	SA60
Thioplast[e]	150	10^{11}	--	14	0.15	4	--	121	14	300	400	SA40
Urethane[e]	350	10^{11}	0.4	3.5	0.04	4	--	93	14	5000	400	SA60
Ceramics												
Cold set cements	70	10^{11}	30	4	0.006	1	704	1038	4.2	500	0	M80
Glass-bonded mica	400	10^{12}	0.1	4	0.003	1	649	704	.84	7000	0	M115

(a) At 70-90°F.
(b) In 24 hours, 1/8 inch thick.
(c) 1 = best, 5 = poorest.
(d) M = Rockwell M; SA = Shore Durometer A; and SD = Shore Durometer D.
(e) Unfilled.

Downloaded from https://www.everyspec.com

Downloaded from https://www.everyspec.com

TABLE 3.17. SOLVENTS, TRADE NAMES, AND MANUFACTURERS OF ENCAPSULATING MATERIALS

Polymer System	Solvents[a] Thinning	Solvents[a] Cleanup	Trade Name and Manufacturer[b]
		Thermosets	
Asphalts and tars	1, 2	2	
Fluorocarbon	--	--	Teflon (12); Kel-F (23)
Polyethylene	1	2	Alathon (12), (10), (22), (25); Stycast (13)
Polystyrene	1, 2	2	Styron (10); Lustron (25); Dylene (22); Stycast (13)
Polyvinyl chloride	--	2	Geon (19); Tygon (35); Opalon (25), (4), (6)
Wax	1, 2	2	
		Thermoplastics	
Alkyd	7	--	Glyptal (18); Paraplex (30), (1), (3), (21)
Allyl ester	--	--	Lexan (18); Merlon (24); Vibrin 135 (28); Laminac (1)
Butadiene-styrene	4	2	Buton (14); Butarey (15), (29); Stycast LOK (13)
Epoxide	1	2, 5	Araldite (7); Epon (33), (21), (11); ERL2744 (2); Devion (5); Ecco and Stycast (13)
Phenolaldehyde	--	--	Resinox (25)
Polyester	4, 5	--	Vibrin (28); Plaskon (3); Paraplex (30); Laminac (1); Dapon (16)
Silicone	4, 8	1, 2	(11), (18); Eccosil (13)
		Elastomers	
Buna-S rubber	1, 3, 4	2	(15), (20)
Chloro-rubber	2	2	Neoprene (23)
Natural rubber	1, 4	2	(15), (20)
Silicone rubber	6	2, 5	(11), (18), (27); Cohrlastic (8), (3); Eccosil (13)
Thioplast	2	--	Thiokol (34)
Urethane	1, 4	--	Scotchcast XR5010 (23); LD167 (12), (24); Ecco (13)
		Ceramics	
Cold-set cements	9	9	Sauereisen (32); Eccoceram (13); Rostone (31)
Glass-bonded mica	--	--	Mycalex (26), (18),

Footnotes appear on following page.

Downloaded from https://www.everyspec.com

References for Table 3.17:

(a) List of Solvents

(1) Hydrocarbon.
(2) Chlorinated hydrocarbon perchloroethylene, methylene chloride.
(3) Diphenyl or terphenyl.
(4) Unsaturated hydrocarbon.
(5) Ketone.
(6) Fluid silicones.
(7) Drying oils.
(8) Aromatic hydrocarbon.
(9) Water.

(b) List of Manufacturers

(1) American Cyanamid Co., 30 Rockefeller Plaza, New York 20, N. Y.
(2) Bakelite Co., Div. of Union Carbide and Carbon Corp., 30 East 42nd St., New York 17, N. Y.
(3) Barrett Div. of Allied Chemical Corp., 40 Rector St., New York 6, N. Y.
(4) Borden Chemical Company, 350 Madison Avenue, New York 17, N. Y.
(5) Chemical Development Corp., Endicott St., Danvers, Mass.
(6) Chemical Products Co., King Philip Road, E. Providence, R. I.
(7) Ciba Co., Inc., Kimberton, Pa.
(8) Connecticut Hard Rubber Co., New and East Sts., New Haven, Conn.
(9) Consolidated Electrodynamics, 300 N. Sierra Madre Villa, Pasadena, Calif.
(10) Dow Chemical Company, Midland, Mich.
(11) Dow Corning Corporation, Midland, Mich.
(12) E. I. du Pont de Nemours and Co., Wilmington 98, Del.
(13) Emerson and Cuming, Inc., Canton, Mass.
(14) Enjay-Humble Oil Co., 15 West 51st St., New York 19, N. Y.
(15) Firestone Plastics Co., Box 690, Pottstown, Pa.
(16) FMC Organic Chemicals Div. of Food Machinery and Chemical Corp., 161 East 42nd Street, New York 17, N. Y.
(17) General Aniline and Film Corp., 435 Hudson St., New York 14, N. Y.
(18) General Electric Co., One Plastics Ave., Pittsfield, Mass.
(19) Goodrich Chemical Co., Cleveland 15, Ohio.
(20) Goodyear Tire and Rubber Co., 1144 East Market St., Akron 16, Ohio.
(21) Jones-Dabney Div. of Devoe Reynolds, 148 South 11th St., Louisville, Ky.
(22) Koppers Company, Pittsburgh 19, Pa.
(23) Minnesota Mining and Manufacturing Co., St. Paul, Minn.
(24) Mobay Products Company, Pittsburgh 5, Pa.
(25) Monsanto Chemical Company, Springfield 2, Mass.
(26) Mycalex Corporation of America, Clifton Blvd., Clifton, N. J.
(27) Narmco Resins and Coatings Co., 600 Victoria St., Costa Mesa, Calif.
(28) Naugatuck Chemical Div. of U. S. Rubber Co., 203 Elm St., Naugatuck, Conn.
(29) Phillips Chemical Co., Bartlesville, Okla.
(30) Rohm and Haas Co., Washington Square, Philadelphia 5, Pa.
(31) Rostone Corporation, Lafayette, Ind.
(32) Sauereisen Company, Pittsburgh 15, Pa.
(33) Shell Chemical Corporation, 380 Madison Ave., New York 17, N. Y.
(34) Thiokol Chemical Corporation, Trenton 7, N. J.
(35) U. S. Stoneware Company, Tallmadge Ave., Akron 9, Ohio.

Downloaded from https://www.everyspec.com

REFERENCE, TABLE 3.17

(1) Volk, M. C., Lefforge, J. W., and Stetson, R., Electrical Encapsulation, Reinhold Publishing Corp ration, New York (1962), Chapters 3, 4, and 5.

Downloaded from https://www.everyspec.com

3.14 Resistivity and Temperature Coefficient of Resistance of Metals and Alloys

The resistivity and temperature coefficient of resistance of some metals and alloys often used in electronic devices are given in Table 3.18. All values of each quantity were measured at 20°C.

Temperature coefficient of resistance, α, is calculated from the formula:

$$R_T = R_o (1 + \alpha \Delta T)$$

where

R_T = resistance at temperature T

R_o = resistance at 20°C

ΔT = T (°C) - 20°C.

TABLE 3.18. RESISTIVITY AND TCR OF METALS AND ALLOYS[a]

Material	Resistivity, 10^{-6} ohm-cm	TCR $(°C)^{-1}$	References
Gold	2.44	.0034	1
Platinum	10	.003	1
Aluminum	2.824	.0039	1
Copper	1.771	.00382	1
Silver	1.59	.0038	1
Kovar	49[b]	.0037[b]	4
Nichrome	100	.0004	1
Manganin	44	.00001	1
Tantalum	15.5	.0031	1
Titanium	55	-.0001 to +.00012[c]	2,3

(a) The values given in this table are for the materials in bulk form except where noted. The same materials may exhibit somewhat different values of these properties when in thin-film form.
(b) Measured at 25°C.
(c) Measured on thin films.

Downloaded from https://www.everyspec.com

REFERENCES, TABLE 3.18

(1) Handbook of Chemistry and Physics, 36th Edition (HCP), Chemical Rubber Publishing Company, Cleveland, Ohio (1955), p 2353.

(2) American Institute of Physics Handbook, McGraw-Hill Book Company, Inc., New York (1957), pp 5-204.

(3) Fuller, W. D., "Titanium Thin Film Circuits", Proc. Natl. Electron. Conf., 17, 32 (1961)*

(4) Carborundum Company, Electronics Division, Latrobe, Pennsylvania, Product Literature.

Downloaded from https://www.everyspec.com

SECTION 4. ACCELERATED TESTING

Downloaded from https://www.everyspec.com

TABLE OF CONTENTS

	Page
4. Accelerated Testing.	4-1
4.1 Introduction.	4-1
4.2 Methods of Generating Accelerated Test Data	4-2
4.2.1 Constant-Stress Testing	4-2
4.2.1.1 Experimental Procedure	4-2
4.2.1.2 Assumptions in Constant-Stress Accelerated Testing. . .	4-5
4.2.1.3 Example of a Constant-Stress Accelerated Test Design	4-6
4.2.2 Step-Stress Accelerated Testing.	4-9
4.2.2.1 Experimental Procedure in Step-Stress Testing	4-10
4.2.2.2 Assumptions Involved in Step-Stress Testing	4-12
4.2.2.3 Relation of Time in Step-Stress Tests to Time in Constant-Stress Tests	4-13
4.2.2.4 Example of a Step-Stress Test Design.	4-14
4.2.3 Progressive-Stress Testing	4-14
4.2.3.1 Experimental Procedure in Progressive-Stress Testing	4-16
4.2.3.2 Assumptions Involved in Progressive-Stress Testing. . .	4-18
4.2.3.3 Relation of Time in Constant-Stress Testing to Time in Progressive-Stress Testing	4-20
4.2.3.4 Example of a Progressive-Stress Test Design	4-21
4.3 Theoretical Models for Analyzing Test Data	4-22
4.3.1 The Arrhenius Model as a Basis for Accelerated Tests	4-23
4.3.1.1 Mathematical Description of the Arrhenius Model. . . .	4-23
4.3.1.2 Stepwise Procedures in Applying the Arrhenius Model	4-25

Downloaded from https://www.everyspec.com

TABLE OF CONTENTS
(Continued)

Page

4.3.1.2.1 Stepwise Analysis Procedures for Constant-Stress Data 4-25

4.3.1.2.2 Stepwise Procedures for Analysis of Step-Stress Data 4-28

4.3.1.2.3 Stepwise Procedure for Analyzing Progressive-Stress Data 4-29

4.3.1.3 Extension of the Arrhenius Model to Include Nonthermal Stress 4-32

4.3.2 The Eyring Model as a Basis for Accelerated Testing 4-33

4.3.2.1 Physical Concept of the Eyring Model. 4-34

4.3.2.2 Generalized Eyring Model to Include Nonthermal Stresses 4-35

4.3.2.3 A Further Modification of the Eyring Model. 4-37

4.3.2.4 Stepwise Procedures for the Analysis of Accelerated Test Data Using the Modified Eyring Model. 4-37

4.3.2.4.1 Stepwise Procedures for the Analysis of Constant-Stress Data 4-38

4.3.2.4.2 Stepwise Procedures for Analyzing Step-Stress Data. 4-41

4.3.2.3.3 Stepwise Analysis Procedures for Progressive-Stress Data 4-43

4.4 Comparison of Accelerated Testing Procedures. 4-48

4.4.1 Evaluation of Methods of Generating Accelerated Test Data. . . 4-48

4.4.2 Comparison of the Arrhenius and Eyring Models 4-49

REFERENCES. 4-51

Downloaded from https://www.everyspec.com

LIST OF FIGURES

		Page
Figure 4.1.	Illustrative Transistor Accelerated Test Design	4-7
Figure 4.2.	Latin-Square Design Layout Within Each of the Five Junction Temperatures	4-8
Figure 4.3.	Stress-Time Curve Relating Time Rate of Parameter Degradation to Applied Stress	4-10
Figure 4.4.	Illustrative Step-Stress Component Aging Specification.	4-15
Figure 4.5.	Illustrative Experimental Layout for Step-Stress Test at One Time Increment, Δt_{ss}	4-15
Figure 4.6.	Schematic Representation of Progressive Stress Testing	4-17
Figure 4.7.	Progressive-Stress Test Curves	4-19
Figure 4.8.	Progressive-Stress (Voltage) Test Design Layout	4-22
Figure 4.9.	Graphical Description of the Analysis of Constant-Stress Accelerated Test Data Using the Arrhenius Model	4-27
Figure 4.10.	Graphic Description of the Analysis of Step-Stress Accelerated Test Data Using the Arrhenius Model	4-30
Figure 4.11.	Graphic Description of the Analysis of Progressive-Stress Test Data Using the Arrhenius Model	4-31
Figure 4.12.	Potential-Energy Diagram for the Aging of an Electronic Component Part.	4-34
Figure 4.13.	Graphical Description of the Analysis of Constant-Stress Data Using the Eyring Model	4-39
Figure 4.14.	Graphical Description of the Analysis of Constant-Stress Data Using the Eyring Model	4-40
Figure 4.15.	Graphical Description of the Analysis of Constant-Stress Data Using the Eyring Model	4-40
Figure 4.16.	Linear Cumulative Damage Plots as a Function of Nonthermal Stress for Various Temperature Levels in the Analysis of Step-Stress Data Using the Eyring Model	4-44

Downloaded from https://www.everyspec.com

LIST OF FIGURES
(Continued)

Page

Figure 4.17. Graphical Description of the Analysis of Step-Stress Data Using the Eyring Model 4-45

Figure 4.18. Graphical Description of the Analysis of Step-Stress Data Using the Eyring Model 4-45

Figure 4.19. Linear Cumulative Damage Plots Over Time as a Function of the Rate of Increase of the Progressive Nonthermal Stress Parameter for Various Constant Temperature Levels in the Analysis of Progressive-Stress Test Data Using the Eyring Model 4-47

Downloaded from https://www.everyspec.com

4. Accelerated Testing

4.1 Introduction

The purpose of a component life test is to generate significant performance data on the basis of which reliability can be predicted for a component operating in a specified environment. These data may be simply time to failure, or they may involve measurement of parameter degradation. In many applications of components, however, the parts are so reliable under normal operating environments that years of testing are required to generate significant performance data. Thus a paradox exists in that the greater a component's reliability, the more difficult it is to determine its reliability.

A possible solution to this problem is to design accelerated life tests. In an accelerated test, components are run at higher stress levels than encountered in normal operation. Stress comprises factors such as ambient temperature, junction temperature, power dissipation, bias voltage, etc. The problem in the analysis of component performance data generated at high stress levels is how to predict the component's performance at a normal stress level. This prediction is generally based on an extrapolation to normal stress levels of performance data generated in several accelerated tests, each run at a different level of increased stress. It is important to recognize that, in practice, reliability prediction at normal stress level is obtained through extrapolation. This extrapolation leads to the most difficult and important problem associated with the analysis of accelerated test data – namely, whether or not "true" acceleration is achieved. For example, at increased stress levels, the mode of failure may change. If several aging processes are acting simultaneously in a component, the "dominant" process may change with increasing stress, or with increasing time. Further, if several stress parameters are involved, such as temperature and voltage, interaction problems may arise between the different stress parameters.[1] Clearly, before valid extrapolations to normal stress levels can be made, the functional relationships between the aging processes acting in a component and the environmental stress parameters must be determined.[2]

Hence, the major weakness in accelerated testing is that the aging processes at high stress levels may be different from those at normal stress levels. Ideally, the application of high stress levels would accelerate aging in such a way that a component degrades in the same manner as at normal stress levels; thus, the only variable that would be accelerated is time. That is, 1 hour of test time at high stress levels should produce identical component performance as T test hours at normal stress levels. If this is the case, "true" acceleration is said to be achieved. In practice, it is not always possible to accelerate only the time variable. Where this is the case, true acceleration cannot be achieved.

In summary, it is seen that accelerated testing consists of

(1) Designing stress tests that yield significant performance degradation during a reasonable time period

(2) Determining, on the basis of component parameter measurements at increasingly higher stresses whether true acceleration has been achieved, or on a physiochemical analysis of the internal processes of a subsample of test devices

Downloaded from https://www.everyspec.com

(3) Validly extrapolating performance at higher stresses to normal stress levels, given that true acceleration has been achieved

(4) Deriving formulas for the calculation of acceleration factors that relate 1 hour of test time at a specified high stress level to an equivalent number of test hours at a reference (normal) stress level.

These problems are treated in detail in the following sections. In Section 4.2, various methods are described for generating accelerated test data. In Section 4.3, two mathematical models characterizing rate processes are described in theory and in application to the analysis of accelerated life-test data. Finally, in Section 4.4, the various accelerated test procedures are compared.

4.2 Methods of Generating Accelerated Test Data

In this section, three test procedures for generating accelerated test data are described. These are:

(1) Constant-stress testing

(2) Step-stress testing

(3) Progressive-stress testing.

For each method of testing, the steps in the experimental procedure are described and the assumptions contingent to the procedure are presented. Further, step-stress and progressive-stress testing are related to constant-stress testing. That is, the equivalent number of hours of constant stress test time is determined for 1 hour of step-stress or progressive-stress test time. Theoretical models for the analysis of the performance data generated by these procedures are presented in Section 4.3.

4.2.1 Constant-Stress Testing

In a constant-stress accelerated test, a statistically significant number of components are placed on life test in each of several levels of environmental stress, which are held constant for the duration of the test. Measurements are taken periodically during the test on those parameters descriptive of component performance over time. At the end of the test, the data obtained are analyzed according to specified mathematical models under appropriate assumptions. The output of this analysis is designed to yield component reliability as a function of the stress acting on the component. Moreover, the analysis should determine the range of stress conditions over which valid reliability predictions can be made.

4.2.1.1 Experimental Procedure

In designing a constant-stress experiment, several detailed questions must be resolved. These are:

Downloaded from https://www.everyspec.com

(1) What parameter or parameters should be measured as descriptors of component performance (the dependent variable)?

(2) What are the thermal and nonthermal stress parameters that make up the stress environment (the independent variables)?

(3) What levels of each of the stress parameters should be combined to make up the various stress environments, what stress ranges should the stress parameters cover, and what should be the spacing between the various levels of each stress parameter?

(4) What should be the sample size within each stress environment and how should the samples be drawn such that statistically significant data are generated?

(5) How long should the test run within each stress level?

It is apparent that these questions cannot be considered independently. The answer to any one question will depend on or partially determine the answers to the remaining questions.

The choice of parameter to be used to describe component performance is, perhaps, the governing issue in designing the experiment. The performance parameter may be (1) failures times or per cent failing as a function of time, (2) time history of a measure such as resistance for a resistor, or reverse leakage current for a diode, or (3) time history of a fundamental aging process such as impurity diffusion or changing surface morphology. The selection of the performance parameter to be used will generally depend on the state of knowledge of the dynamic physical properties of the component type and the nature of the results of other experiments run on the component. Further, in addition to selecting a performance parameter, the question of the appropriate measure of that parameter must be answered. For example, for a resistor, the most appropriate measure may be per cent deviation of resistance from its initial value.

Parameter measures for resistors, capacitors, diodes, and transistors that are generally useful for characterizing component aging under accelerated test conditions are described in Section 5.6, "Physical Methods of Screening". In that section, parameter measures for each component type are tabulated and their utility as precursors of failure described. It should be recognized that those parameters presented are described in a screening context. As such, not all of the parameter measures are suitable for the general accelerated test problem. Parameter measures such as resistance change due to short-term voltage overload for resistors obviously fall into this category.

Specification of the stress parameters for the experiment should be determined from the stress environments expected to be encountered in application of the component. That is, the life-test conditions should simulate in terms of the stress parameters the actual use environment. Acceleration then, is achieved by increasing the magnitude of the stress parameters. For example, in the case of transistors, junction temperature may be an appropriate stress parameter. Acceleration may be achieved in this case by increasing the ambient temperature alone, increasing the applied voltage alone, or by increasing both parameters in combination. It should be noted, however, that a given stress test may accelerate a particular mechanism, or mechanisms, but not others. It is one of the

Downloaded from https://www.everyspec.com

objectives of reliability physics to determine the stress-aging mechanism relationships, thus permitting more exacting test designs.

A further consideration in determining the test stress environment is the possible need for pre-stressing the components prior to actual initiation of the life test. Since accelerated testing is primarily concerned with failure mechanisms that are stress-time dependent, it may be necessary to submit the components to a pre-test stress environment to screen out early failures. Such a pre-test may consist of temperature cycling, thermal shock, vibration, etc.

It is axiomatic that the stress levels chosen for an accelerated test should not be so severe as to activate failure mechanisms that would not be activated under normal conditions. In addition to observing this precaution, it is necessary to consider reversible changes that can take place in device properties at time zero when the device is stressed. In particular, those reversible changes which can lead to irreversible changes and immediate failure should be identified. (A reversible change is here meant to be one which has a finite value when the stress is applied and is zero after the stress is removed.) Some of the more common reversible changes are described briefly in the following paragraphs. Most of these reversible changes are temperature induced since many accelerated tests call for the use of elevated temperatures.

One of the most important factors to be considered in setting a stress temperature level for all electronic devices is the melting point of solder used in fabricating the device. Table 3.15 of this Notebook lists the melting points of many types of solder. Other materials, such as insulation and encapsulation, used in fabrication of devices may soften, melt, or deform under elevated-temperature conditions.

The temperature coefficient of resistance of a resistor and temperature coefficient of capacitance of a capacitor will usually be known and accounted for in designing an accelerated test. The differences in thermal coefficients of expansion between, for example, a resistive thin film and its substrate, or a dielectric film and its electrode, should also be considered since differences in expansion of contiguous materials may result in film rupture. Table 3.12 lists thermal coefficients of linear expansion of many materials used in construction of electronic devices.

The electric breakdown strength of both capacitor dielectrics and encapsulants is temperature dependent. Above a certain "critical temperature", which is different for different materials, the dielectric strength, F, is often described by a function of the form $F = a \exp b/T$, where a and b are constants for a given material. Whitehead[3] gives curves of dielectric strength plotted against temperature for a number of materials.

The leakage current, I_L, of an electrolytic capacitor at a fixed voltage rises exponentially with temperature. The exact functional relationship between I_L and T is different for different types of capacitors, but in general I_L can be expected to rise sharply as T is increased. The leakage resistance of most dielectrics decreases rapidly with increasing temperature. Thus, if a capacitor is to be placed in a voltage-temperature stress, care should be taken that the combined effects of these stresses do not result in high leakage currents and subsequent excessive heating, which may be destructive.

Downloaded from https://www.everyspec.com

Many of the basic properties of semiconductors are temperature sensitive, including energy gap, carrier concentration, electron and hole mobility, minority-carrier lifetime, reverse current across a p-n junction, and p-n junction reverse breakdown voltage. These variable properties may influence the choice of stress levels for an accelerated test. The two properties whose temperature dependences are of most interest are reverse leakage current and reverse breakdown voltage. Reverse leakage current is a steeply varying exponential function of temperature, whereas reverse breakdown voltage increases approximately linearly with temperature.

Once the stress parameters have been defined, the next step consists of determining the particular levels of stress for the test. A desirable set of stress levels would include the lowest stress level that is still sufficiently severe to yield parameter degradation during the test period, the highest stress level at which the same aging or failure mechanisms are dominant, and a sufficient number of levels between these extremes to permit valid extrapolations to use conditions. For the models for analyzing accelerated test data presented in Section 4.3, it is desirable to have at least five stress levels to determine the functional relation between degradation and stress.*

Once the questions involving the physical factors of the accelerated test have been resolved, the number of components that are to be tested within each stress level must be determined. This is basically a statistical question of what constitutes a significant sample size. The number of components within each cell should be sufficiently large (1) to permit precise estimates of the parameter degradation rates and (2) to detect with statistical significance differences between degradation rates at the various stress levels. Sample sizes necessary to yield "precision" and "statistical significance" will depend on the variance of parameter measurements within the various stress levels and on the variance of estimates of the difference in degradation rates between stress levels, respectively. In general, higher stress levels may be expected to yield greater variations in parameter measurement, and hence, require larger sample sizes.

The length of the test is closely related to the problem of sample size. In this case, the problem is to run the test within each stress level for a period of time that is sufficiently long (1) to yield significant component degradation in terms of the measured parameters and (2) to evaluate the time dependency of dominant aging mechanisms acting in a component. In other words, the test should run long enough to discriminate between "true" aging effects and random disturbances. These disturbances include (1) experimental error due to variations in the physical properties of nominally identical devices, (2) error due to variations in control of the experimental stress levels, and (3) measurement error due to instrumentation, or inaccurate recording.

4.2.1.2 Assumptions in Constant-Stress Accelerated Testing

The most critical assumption in constant-stress accelerated testing is that the relative effects of the aging mechanisms on component degradation are invariant with respect to stress over the range of stress levels for which acceleration factors are to be calculated. The assumption implies that new failure mechanisms are not induced in a

*At least five stress levels will be required in order to detect, in the presence of experimental error, whether a shift in the dominant aging mechanism occurs at some stress level. The choice of five is somewhat arbitrary, but does represent a reasonable minimum.

Downloaded from https://www.everyspec.com

component at high stress levels. A procedure for testing this assumption is postulated in Section 4.3, "Theoretical Models for Analyzing Test Data". Basically, the postulate states that if a new (or different) aging mechanism becomes "dominant" at some increased stress level, the activation energies (estimated by the analysis) for the lower and higher stress levels will be different.

In addition to the critical physical assumption, various statistical assumptions must be made for a rigorous analysis of the test data. These assumptions involve (1) the distribution of statistical variations of parameter measurements among the set of components under test, (2) the distribution of errors due to instrumentation in regulating the values of the stress parameters, and (3) the combination of various error distributions in constructing the estimates and hypothesis test of quantities such as degradation rates, acceleration factors, and activation energies. To date, well-defined statistical procedures for analyzing these data using the Arrhenius and Eyring models are generally nonexistent. However, a procedure for treating the statistical aspects of the data analyzed according to the Arrhenius model is presented in Section 4.3.1. Further, several aspects of this problem are treated in Reference (4).

4.2.1.3 Example of a Constant-Stress Accelerated Test Design

To illustrate the experimental procedure described in Section 4.2.1.1, a transistor accelerated test design is described in the following paragraphs. The purpose of this test design is to relate transistor parameter degradation to stress, where the Arrhenius reaction-rate equation is hypothesized as the appropriate analytical model. The experimental design is illustrated in Figure 4.1.

The test design shown in Figure 4.1 is based on the premise that junction temperature is the appropriate stress. The design includes explicitly several stress parameters on which junction temperature depends. These stress parameters in the design are seen to be power dissipation, voltage-current ratio, and ambient temperature. The incremental increases in ambient temperature are chosen such that $\Delta T = \theta \Delta P$, where θ is the coefficient of thermal impedance for the transistor type under test. Thus, the test cells along each of the five diagonals of the test layout are at constant junction temperatures. It is seen that five different junction temperatures are achieved in this way, thus providing a reasonable basis for extrapolation of acceleration factors to stress levels of lesser magnitude.

This design has several desirable attributes in addition to simply providing test data for relating junction temperature (stress environment) to transistor-parameter degradation rates. These are:

(1) The use of collector voltage and collector current ratio permits the evaluation of transistor degradation with respect to varying bias conditions.

(2) The effects of transistor degradation at constant junction temperature with respect to varying combinations of power dissipation and ambient temperatures can be evaluated.

(3) Two test cells at zero power dissipation permit a test of whether different aging mechanisms are dominant under powered and unpowered stress conditions.

Downloaded from https://www.everyspec.com

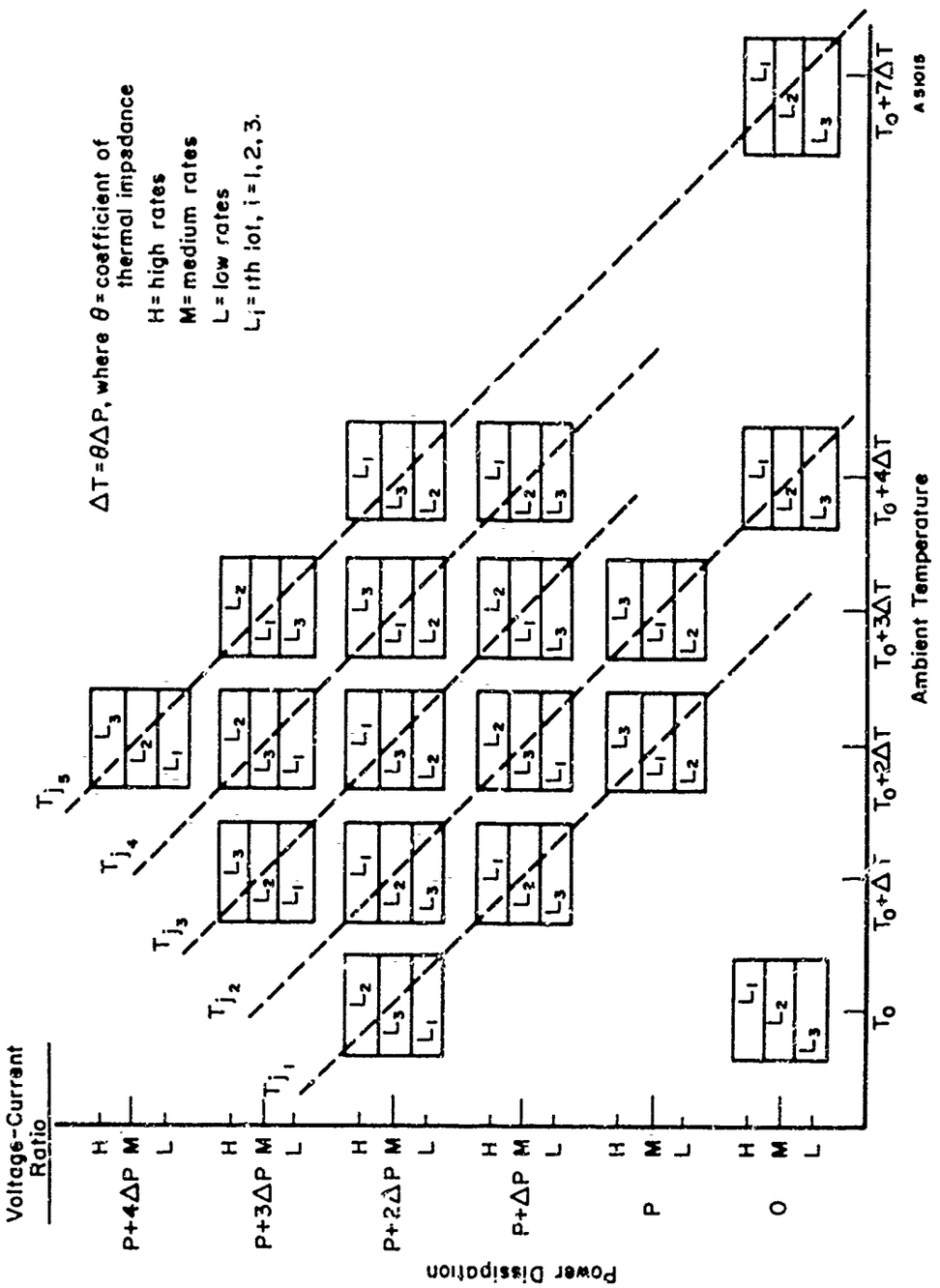


FIGURE 4.1. ILLUSTRATIVE TRANSISTOR ACCELERATED TEST DESIGN

Downloaded from https://www.everyspec.com

(4) A control test cell provides a basis for validating extrapolation of acceleration factors to lower stress levels.

(5) Variations in degradation rates among product lots (or possibly production processes) can be evaluated.

The description of the test design brings out the salient attributes of the design with respect to physical properties of the problem. From a statistical point of view, the design consists of five 3 x 3 Latin squares plus three unpowered stress cells. Each Latin square is nested within a constant junction temperature. The Latin-square layout is shown in Figure 4.2. The same three voltage-current ratios and the same three lots appear in each square. However, there are 15 levels of power-temperature combinations.

Power-Temperature Combination ($T_\alpha + \theta P = T_j$)	Voltage-Current Ratio: High	Medium	Low
(P_L, T_H)	L_2	L_1	L_3
(P_M, T_M)	L_3	L_2	L_1
(P_H, T_L)	L_1	L_3	L_2

FIGURE 4.2. LATIN-SQUARE DESIGN LAYOUT WITHIN EACH OF THE FIVE JUNCTION TEMPERATURES

The analysis of variance model for the design (not including unpowered test cells) is given by

$$Y_{ijkl} = \mu + (T_j)_i + (V/I)_{(i)j} + (P,T)_{(i)k} + L_{(i)l} + e_{ijkl} \quad , \tag{4.1}$$

where

Y_{ijkl} = the observed transistor parameter

μ = overall mean

$(T_j)_i$ = ith junction temperature (a fixed factor)

$(V/I)_{(i)j}$ = the jth voltage-current ratio nested within the ith junction temperature (a fixed factor)

$(P,T)_{(i)k}$ = the kth power-temperature combination nested within the ith junction temperature (a fixed factor)

Downloaded from https://www.everyspec.com

$L_{(i)l}$ = the lth lot nested within the ith junction temperature (a random factor)

e_{ijkl} = the unexplained experimental error assumed to be normally and independently distributed with mean zero and variance σ^2.

Analysis of the effects of the various factors in the model, of course, require compliance with the statistical assumption associated with the analysis of variance. If, for example, change in I_{CBO} is the measured transistor parameter, the model may fail since there is considerable evidence showing that these factors are multiplicative, and not additive as used in the model. However, satisfactory compliance with the model may be achieved by taking $y = \log \Delta I_{CBO}$, thus obtaining the additive effect as required. If the tests in the analysis show the effects of one or more of the factors (voltage-current ratio, power-temperature combination, or production lot) to be not statistically significant, the data within each junction temperature may be "pooled", thus permitting a more precise estimate of degradation rates as a function of junction temperature.

4.2.2 Step-Stress Accelerated Testing

One problem in constant-stress accelerated test programs is the excessively long time periods that may be required at the lower stress levels to yield significant parameter degradation. An alternative approach, which resolves the problem, is step-stress testing. In a step-stress test, a set of components are tested for fixed time intervals at successively higher stress levels until sufficient degradation data (or component failures) are obtained. That is, the components are initially placed on test for a time increment, Δt, at a low stress level, S_1. After the Δt hours have elapsed, the stress level is increased to S_2, and the test is run for Δt more hours. The step-stress sequence is repeated until a sufficient quantity of parameter degradation data is generated. This procedure is repeated for several different time increments, $\Delta t_1, \ldots, \Delta t_p$, and the stress step, $S_i - S_{i-1}$, is the same for all tests.[5,6,7]

Step-stress testing in the above manner depends on the assumption that the probability of accumulating some fixed amount of parameter degradation (or probability of failure) at a given point in the stress-time domain is independent of the particular combinations of stress and time used to arrive at that point.[8] Thus, in a sense the roles of stress and time are interchanged in step-stress testing, in contrast to constant-stress testing. This point is illustrated in Figure 4.3. This figure shows that if the stress level is fixed at some level S_i, as in constant-stress testing, a distribution, $f_{S_i}(t)$, is obtained where t denotes time to reach a specified degradation in some parameter (e.g., 100 per cent change from the initial value of I_{CBO}), or t may denote time to failure. Analogously, if time is fixed at some period t_i, as in step-stress testing, a distribution, $g_{t_i}(S)$, is obtained where S is the stress level at which a specified degradation was accumulated. In either case, under the assumptions of the testing procedures, the same stress-time function relating the time rate of parameter degradation to applied stress is obtained.

The weakness of step-stress testing as a method for obtaining "true" acceleration is essentially the same as in constant-stress testing. The relative effects of the aging mechanisms acting in a component may change as the stress level is increased. Therefore, it must be known, a priori, over what range of stress levels the relative effects of

Downloaded from https://www.everyspec.com

the aging mechanisms are invariant or, more desirably, the methods for analyzing step-stress data must include a procedure for detecting such changes. Such a detection procedure is given for the analysis models described in Section 4.3.

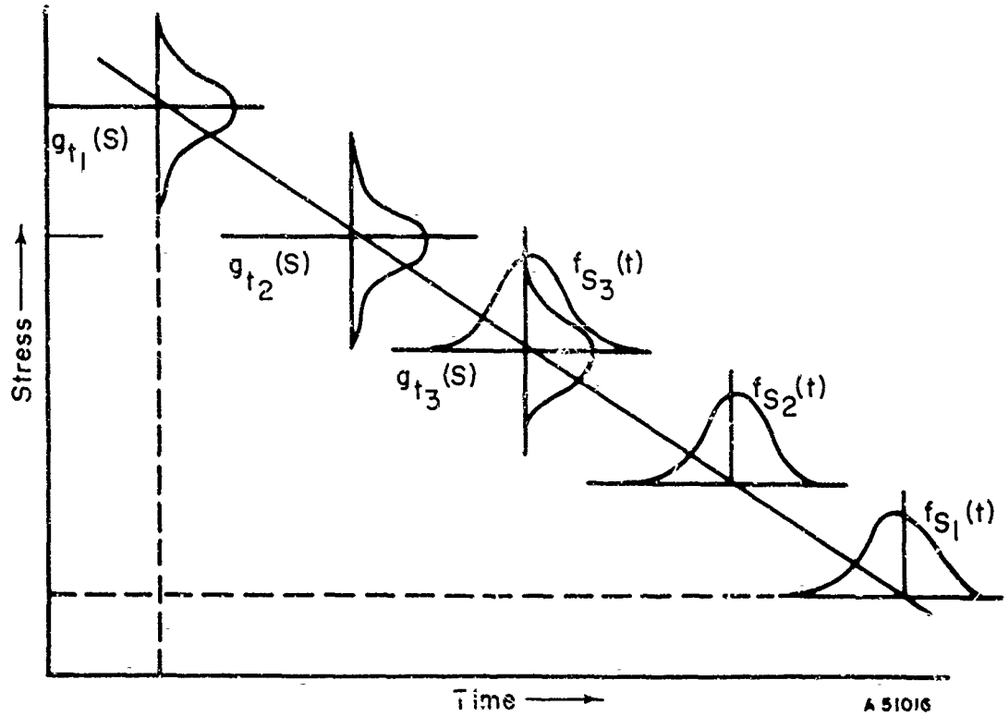


FIGURE 4.3. STRESS-TIME CURVE RELATING TIME RATE OF PARAMETER DEGRADATION TO APPLIED STRESS

Scales may be in some transformed units of stress and/or time such as log time.

4.2.2.1 Experimental Procedure in Step-Stress Testing

The design of a step-stress experiment requires that several detailed questions be resolved that are analogous to those in designing a constant-stress test. These are:

(1) What parameter or parameters should be measured as descriptors of component performance (the dependent variable)?

(2) What are the thermal and nonthermal stress parameters (the independent variables)?

(3) At what stress level should the components be tested in the first time increment? What size stress step should be used?

Downloaded from https://www.everyspec.com

(4) How many tests using different time increments should be run? How long should the time increments be for each test?

(5) What should the sample size be for each of the step stress tests?

The answers to these questions will depend primarily on what is known, a priori, about component performance in the stress-time domain. Where the a priori knowledge is substantial, answers to the above questions can be determined that may result in a more efficient test design in terms of factors such as sample size and number and spacing of time increments.

The question of what parameters should be measured is the same for step-stress as for constant-stress testing. Not only must those parameters that are most descriptive of component performance be identified, but the appropriate numerical measures of those parameters must also be determined. The type of measure may be dictated in part by the mathematical model to be used in analyzing the data. For example, in application of the Arrhenius model, log $[I_{CP}$ (at time t)$/I_{CBO}$(initial)] has been shown to be an appropriate measure for degradation of some types of transistors. Similarly, identification of the interactions between the thermal and nonthermal stress parameters is the same in step-stress as in constant-stress testing.

In resolving the third question, specification of the sequence of stress steps, the experimental procedure departs from that for constant-stress testing. In this case, it is necessary to determine (1) what scale of measurement units should be used, (2) how the sequence of stress steps should be spaced, and (3) how large should the stress step be. The scale of measurement should correspond to the form in which the stress parameter appears in the theoretical analysis model. For example, thermal stress appears in the Arrhenius and Eyring models as $1/T$, where T is in degrees Kelvin. Therefore, the appropriate scale for thermal stress should be reciprocal temperature, expressed in degrees Kelvin. Secondly, it is desirable that stress steps be equally spaced in the measurement scale previously defined. Thus, for the thermal-stress case above, the stress steps would be spaced such that $(1/T_i - 1/T_{i-1}) =$ constant for $i = 1, \ldots n$. Inspection of this relation reveals that equal spacing in units of reciprocal temperature yields successively larger stress steps in units of temperature. For example, a reciprocal temperature step of 0.1×10^{-3} yields a step of 8 degrees at 25°C and a step of 18 degrees at 150°C. Thirdly, the magnitude of stress step must be determined. Here, a balance should be achieved between specifying a stress step that is sufficiently large to permit discrimination among the degradation rates through statistical analysis of the test data, and, on the other hand, that is sufficiently small to yield a statistically significant number of data points over a range of stress levels in which a given failure mechanism is dominant.

The governing considerations on the limits of the stress conditions for accelerated tests on the various component types are essentially the same as those described in Section 4.2.1.1 for constant-stress testing. The additional problem in this case is to select the magnitude of the stress step such that significant parameter degradation will result before the stress limits are exceeded.

The fourth question, how many tests using different time increments should be run, is completely analogous to the question of how many stress levels should be used in

Downloaded from https://www.everyspec.com

constant-stress testing. This stems from the fact that the role of stress and time are essentially reversed in step-stress testing. For analysis of the data on the basis of the Arrhenius or Eyring models, it follows from the analogy that there should be at least five time increments to permit an adequate test of the linearity requirement of the models.

Finally, the number of samples that should be tested in each step stress test is a statistical question identical with that in constant-stress testing. The sample sizes should be sufficiently large to permit reasonably precise estimates of component parameter degradation within test conditions, and to permit discrimination of degradation effects among the test conditions.

4.2.2.2 Assumptions Involved in Step-Stress Testing

Two major assumptions are critical to the analysis of step-stress test data. These are:

(1) Component damage (time-dependent degradation measured in terms of one or more parameters) is strictly cumulative. That is, damage accumulated in a given stress step is given by the expression $\Delta D_i = R(S_i)\,\Delta t$, where $R(S_i)$ is the degradation rate at the ith stress step and Δt is the increment in the step. Thus, the total accumulated damage through the kth stress step is given by

$$D = \sum_{i=1}^{k} \Delta D_i = \sum_{i=1}^{k} R(S_i)\,\Delta t \quad . \tag{4.2}$$

(2) The degradation rates, $R(S_i)$, are independent of each other. That is, the rate at which damage accumulates in the kth stress step, $R(S_k)$, does not depend on the accumulated damage through (k-1) preceding steps or on the rates, $R(S_i)$, i = 1, . . ., (k-1), in those steps.

These assumptions essentially state that the aging process acting in a device at a given point in time depends only on the intrinsic physical properties of the device and the applied stress at the given time, and is independent of its prior time history.

In addition, step-stress testing involves those assumptions associated with constant-stress testing. Most important of these assumptions is that the relative effects of the aging mechanisms on component degradation are invariant with respect to stress over the range of stress levels for which acceleration factors are to be calculated. Further assumptions are involved in the statistical analysis of such data. These include (1) the distribution of statistical variations of parameter measurements among the set of components under test, (2) the distribution of errors due to instrumentation in regulating the values of stress parameters, and (3) the combination of various error distributions in constructing estimates and hypothesis tests.

Downloaded from https://www.everyspec.com

4.2.2.3 Relation of Time in Step-Stress Tests to Time in Constant-Stress Tests

If, in addition to the first two assumptions cited above, it is assumed that the state of a device at a specified damage (degradation) level is independent of how that level was obtained, operating time under step stress can be related to operating time under constant stress. An equivalent statement of this assumption is that failure of homogeneous components occurs when a specified damage is accumulated irrespective of the rate (or rates) at which damage occurs. The required relationship is obtained by equating the functions of cumulative damage for constant and step stress.[5,6]

For constant-stress testing at stress level S_c, the cumulative damage at time t_c is given by

$$D(t) = R(S_c)\, t_c \quad , \tag{4.3}$$

where $R(S_c)$ is the degradation rate at stress level S_c. For step-stress testing, cumulative damage is given by

$$D(t) = \Delta t_{ss} \sum_{n=1}^{N} R(S_n) = \frac{t_{ss}}{N} \sum_{n=1}^{N} R(S_n) \quad , \tag{4.4}$$

where $R(S_n)$ is the degradation rate in the nth step and Δt_{ss} is the operating time in each step. The total operating time in the step-stress test is, therefore, seen to be $t_{ss} = N\Delta t_{ss}$. Equating (4.3) and (4.4) and solving for t_c as a function of t_{ss} yields

$$R(S_c)\, t_c = \frac{t_{ss}}{N} \sum_{n=1}^{N} R(S_n) \tag{4.5}$$

$$t_c = \frac{t_{ss}}{NR(S_c)} \sum_{n=1}^{N} R(S_n) \quad . \tag{4.6}$$

Note that t_c is the constant-stress time and t_{ss} is the total step-stress time to obtain equal component damage.

If the degradation rate for a particular T is given by the Arrhenius equation

$$R(T) = e^{A-B/T} \quad , \tag{4.7}$$

where T is temperature (operating temperature in the case of resistors and junction temperature for semiconductors), the required relation is seen to be

$$t_c = \frac{t_{ss}}{N} \sum_{n=1}^{N} e^{-B/T_n} \Big/ e^{-B/T_c} \quad . \tag{4.8}$$

Downloaded from https://www.everyspec.com

For dielectric materials, it is generally assumed that, for fixed temperature, stress is given by the logarithm of applied voltage, V. If it is further assumed the degradation rate follows an Eyring rate equation,

$$R(T,S) = ATe^{-B/kT}\, e^{S(C+D/kT)} = R_o e^{\phi S} \quad , \tag{4.9}$$

when R_o is the degradation rate in the absence of nonthermal stress and $e^{\phi s}$ represents the adjustment of R(T,S) due to thermal stress, then the degradation rate for fixed temperature is given by

$$F(T_o,S) = R_o e^{\phi \ln V} = R_o V^{\phi} \quad . \tag{4.10}$$

substituting (4.10) into (4.6) yields the relation of t_c to t_{ss}:

$$t_c = \frac{t_{ss}}{N} \sum_{n=1}^{N} V_n^{\phi} \Big/ V_c^{\phi} \quad . \tag{4.11}$$

4.2.2.4 Example of a Step-Stress Test Design

To illustrate the experimental procedures described in Section 4.2.2.1, a transistor step-stress test design is outlined in the following paragraphs. The purpose of the test is to generate acceleration factors based on the Arrhenius rate equation for a silicon mesa-type transistor. The experimental design is illustrated in Figures 4.4 and 4.5.

The test design is based on the premise that junction temperature is the appropriate stress parameter and that collector reverse leakage current is an appropriate performance parameter. Figure 4.4 shows that the design consists of three separate step-stress tests, each based on a different stress-step time increment, Δt_{ss}. Stress is increased in equal increments of reciprocal junction temperature, $1/T_j$, for all tests. Figure 4.5 describes the design within each of the three time increments. It is seen that ambient temperature is held constant throughout the test; thus, junction temperature is increased by step increases in power dissipation, ΔP. Further, three voltage-current ratios are used within each stress step to permit a test of I_{CBO} degradation with respect to current and/or voltage as well as junction temperature. Thus, the total sample of components under test is divided into 3 3 = 9 subsets, S_1, S_1, ... , S_9.

Each of the three tests is terminated in accordance with a predetermined stopping rule. For example, testing may be terminated when at least 90 per cent of the components in each subset have accumulated a specified amount of damage in terms of the performance parameter I_{CBO}, or when that stress step is reached in which at least 50 per cent of the components have failed. In the second case, failure may be defined in terms of such parameters as I_{CBO}, BV_{CES}, I_{EBO}, and V_{CE} (sat.).

4.2.3 Progressive-Stress Testing

In progressive-stress testing, a sample of components is placed on test at a stress level that increases at a constant time rate, i.e., $S = \alpha + \beta t$. To obtain the functional

Downloaded from https://www.everyspec.com

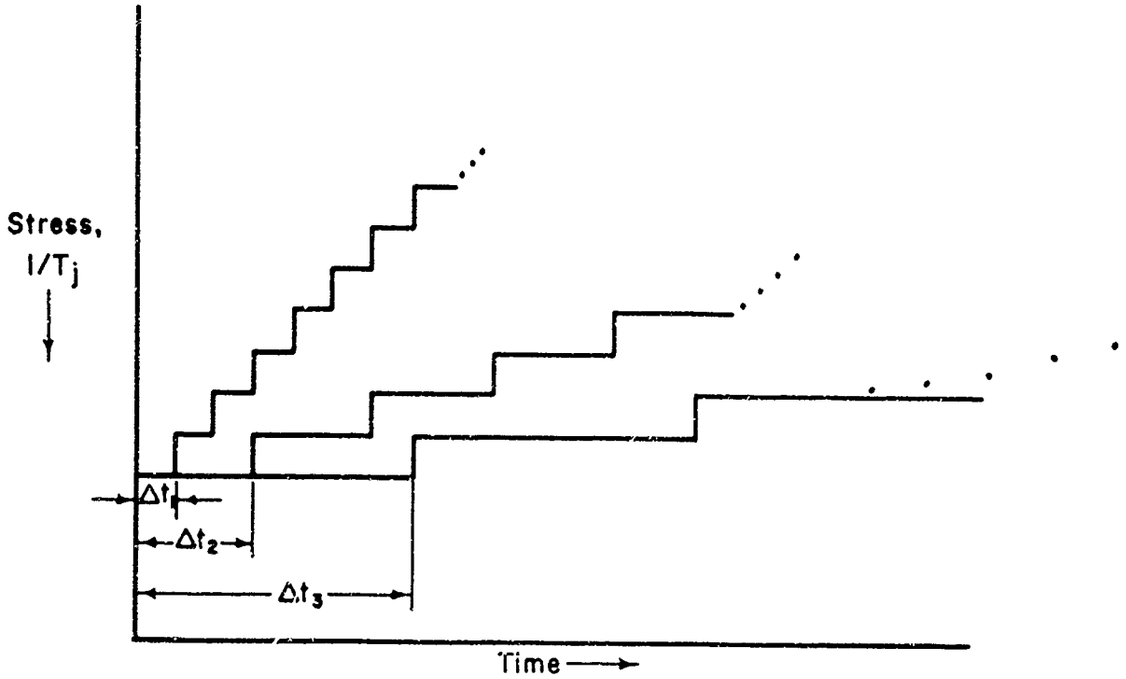


FIGURE 4.4. ILLUSTRATIVE STEP-STRESS COMPONENT AGING SPECIFICATION

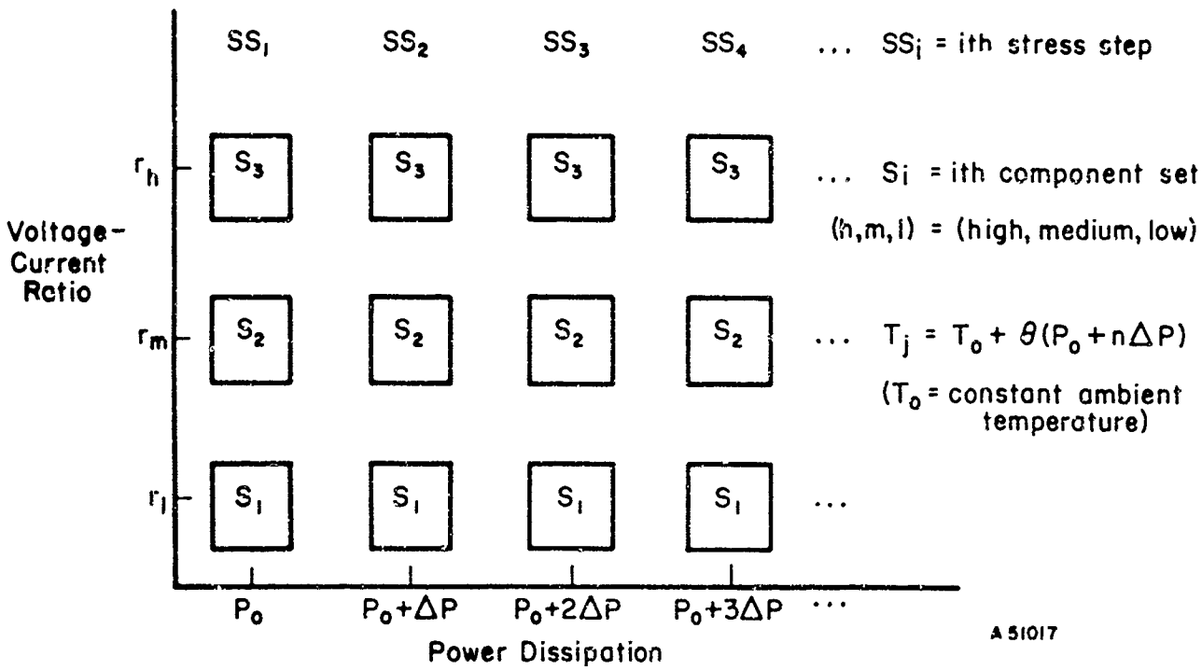


FIGURE 4.5. ILLUSTRATIVE EXPERIMENTAL LAYOUT FOR STEP STRESS TEST AT ONE TIME INCREMENT, Δt_{ss}

The design is replicated for each time increment in Figure 4.

Downloaded from https://www.everyspec.com

dependence of component degradation in the stress-time domain, the procedure is repeated at several different time rates of change of stress, $\beta_1, \ldots, \beta_p$.[9] These tests are terminated according to the same stopping rules used in constant- or step-stress testing. The tests may be terminated when a specified amount of degradation (damage) is accumulated, or when a specified per cent of the sample has failed. This procedure is illustrated schematically in Figure 4.6. Thus, progressive-stress testing consists of taking step-stress testing to the limit. That is, a step-stress test becomes a progressive test when the time increment, Δt_{ss}, is zero. Therefore, it follows that progressive stress depends on the same assumption as step stress. That is, it is assumed that the probability of accumulating some fixed amount of parameter degradation at a given point in the stress-time domain is independent of how that point is arrived at. The implications of the assumption are completely analogous to those described in introduction to Section 4.2.2 and illustrated in Figure 4.3.

Moreover, the threat to progressive-stress testing as a method of generating valid accelerated test data is the same as in constant- and step-stress testing. That is, unless it is known, a priori, over what range of stress the relative effects of the aging mechanisms acting in a component are invariant, or a procedure is available for testing for invariance, invalid acceleration functions may be obtained. Also, in progressive-stress testing linear programming of temperature is not easy to achieve without the use of expensive control equipment.

4.2.3.1 Experimental Procedure in Progressive-Stress Testing

The design of a progressive-stress test involves the consideration of several detailed factors. These are:

(1) What parameter or parameters should be measured as descriptors of component degradation (the dependent variables)?

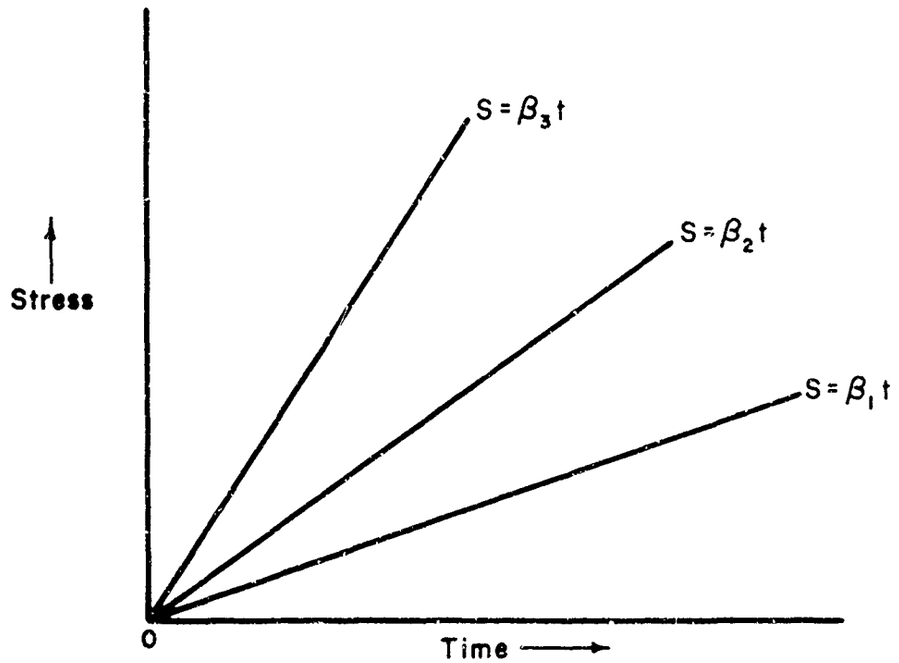
(2) What are the thermal and nonthermal stress parameters (the independent variables)?

(3) How many different rates of applying stress should be used? At what stress value should the test start? What spacing should be used among the various stress rates?

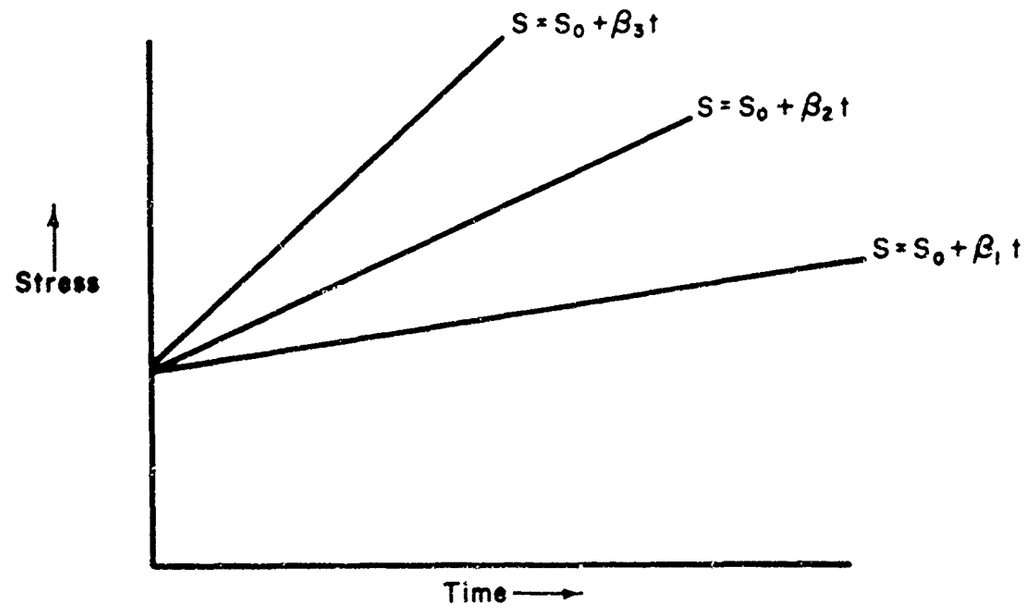
(4) What should be the sample size for each progressive-stress test?

The question of what parameters should be measured involves the same considerations for progressive-stress testing as for constant- and step-stress testing. The second question, what thermal and nonthermal stress parameters should be used, involves problems peculiar to progressive-stress testing. Continuously increasing operating temperature at a constant rate is difficult to regulate precisely with standard test equipment. Therefore, continuous stress increase may most easily be attained in terms of power stresses with temperature held constant for a given test. The effect of temperature stress can be estimated by running several progressive-stress tests at different, but constant, temperature levels and increasing the power stress at the same rate in all tests.

Downloaded from https://www.everyspec.com



(a) Stress increased linearly from S = 0 at time t = 0



(b) Stress increased linearly from S = S_0 at time t = 0

A 51018

FIGURE 4.6. SCHEMATIC REPRESENTATION OF PROGRESSIVE STRESS TESTING

Downloaded from https://www.everyspec.com

The number of progressive-stress tests (each test at a different stress rate) depends on the number of factors to be investigated in the stress-time domain. In general, the question in this case is how many points are needed in the stress-time curve (the Arrhenius plot, for example) to provide "valid" estimates of acceleration factors. If the problem is to determine acceleration factors as a function of applied voltage for a fixed temperature level (for example, for capacitor aging based on the Eyring model), then at least five different rates of applying voltage should be used in order to obtain meaningful estimates of the two parameters in the model associated with power stress. If, in addition, temperature is considered as a variable, then several tests at the same rate of applying voltage must be run, each at a different temperature level.

Determination of the initial stress level at $t = 0$ requires that the experimenter decide whether the power stress should start at the same value for all tests, or whether the combined temperature-power stresses should have the same initial value. For example, consider the progressive-stress test design for transistors shown in Figure 4.7. It is seen in Figure 4.7(a) that, if each test is to be started at the same level of power dissipation, P_o, then the initial level of junction temperature, T_j, will be different in each case. Alternatively, Figure 4.7(b) shows that, if T_j is the same for each test, then P_o will be different.

Finally, determination of how far apart the stress rates should be spaced is analogous to the stress-step and time-increment problem in step-stress testing. Specifically, the rate of applying stress should be sufficiently high to produce significant parameter degradation in short periods of time, sufficiently low so as not to stress the components beyond that stress range over which the relative effects of the aging mechanisms acting in a device are invariant, and spaced between these extremes to permit a desirable spread of points on the stress-time curve in analyzing the test data.

The fourth question, what the sample size for each progressive stress test should be, is a statistical question involving the same considerations as in constant- and step-stress testing. The sample size should be sufficiently large to permit precise estimates of parameter degradation measures for each rate of applied stress, and to permit accurate discrimination of effects on degradation measures among the various rates of applied stress.

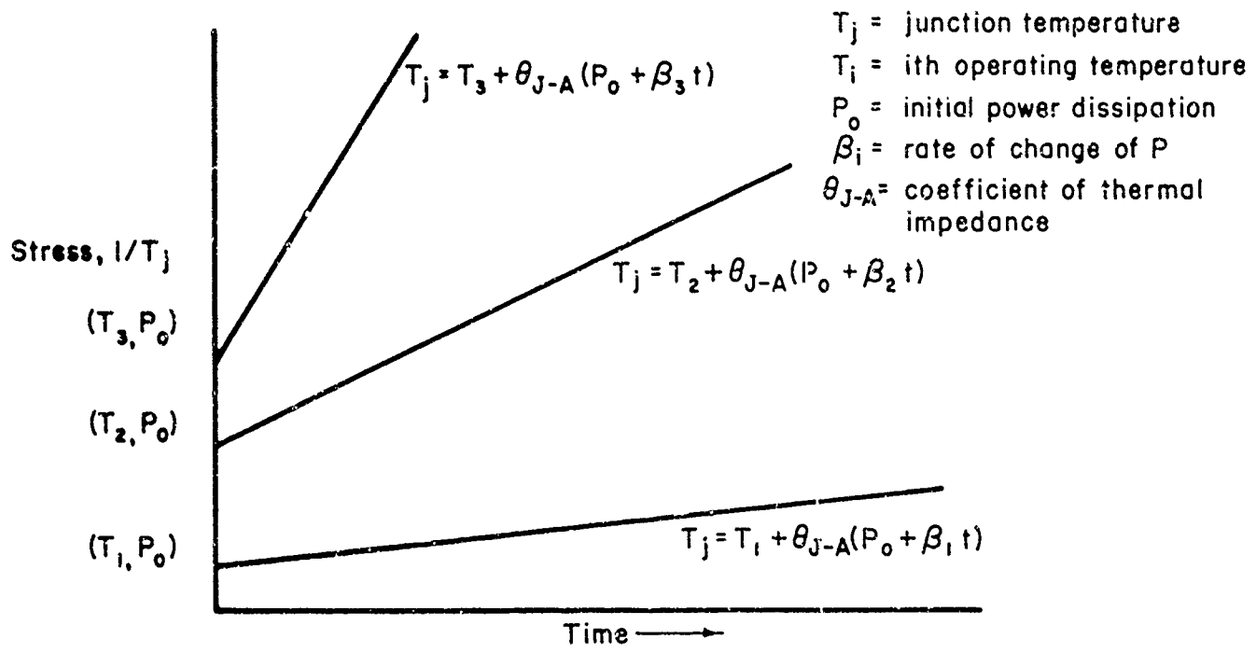
4.2.3.2 Assumptions Involved in Progressive-Stress Testing

The assumptions on which progressive-stress testing depend are analogous to those in step-stress testing. These are:

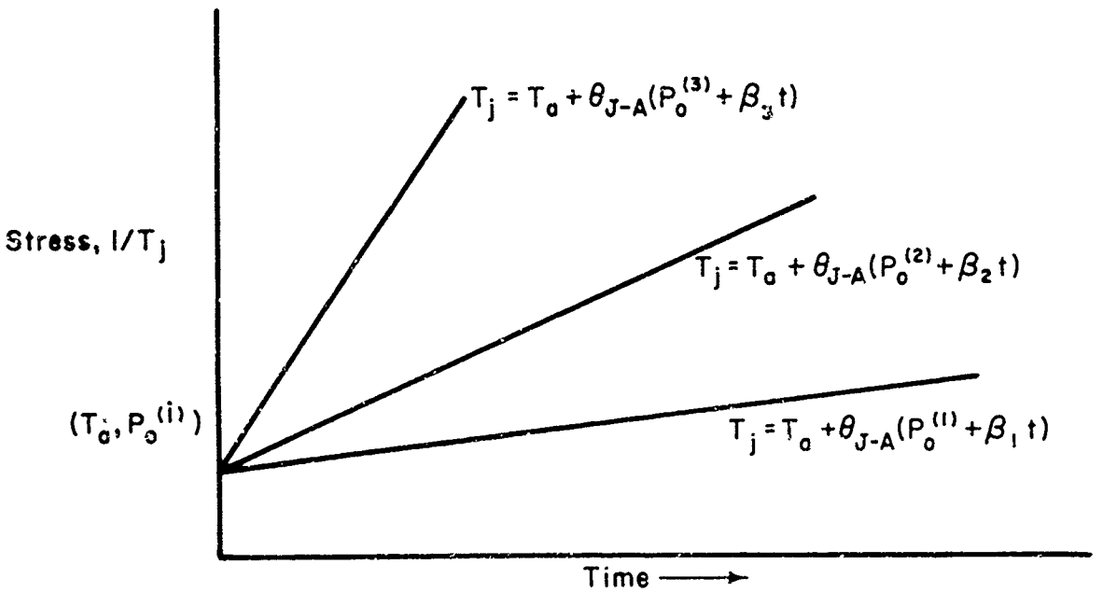
(1) Component damage (stress-time dependent degradation measured in terms of one or more parameters) is strictly cumulative. That is, damage accumulated in an operating period, t, is given by

$$D(t) = \int_o^t R(S)\,dt = \int_o^t R[S(t)]\,dt \quad , \tag{4.12}$$

Downloaded from https://www.everyspec.com



(a) Various constant operating temperature levels and identical initial values of power dissipation



(b) Same initial value of junction temperature and various initial values of power dissipation

A 51019

FIGURE 4.7. PROGRESSIVE-STRESS TEST CURVES

Downloaded from https://www.everyspec.com

where R(s) denotes parameter degradation rate at stress level s. Since s 's a linearly increasing function of time, Equation (4.12) can be written

$$D(t) = \int_o^t R(S_o + \beta t)\, dt \quad , \tag{4.13}$$

where S_o denotes the applied stress of $t = 0$, and denotes the time rate of increase in stress.

(2) When components are tested to failure, failure occurs after a specified amount of degradation accumulates irrespective of degradation rate, R, and, therefore, is independent of the rate of stress increase.

(3) The relative effects of the aging mechanisms acting in a device are invariant over the range of stress conditions used in testing. For progressive-stress testing, it is seen that this stress range will depend on the product of rate of stress increase, β, and the test time.

In addition to these physical assumptions relating stress and degradation, there are several statistical assumptions required for data analysis. These assumptions are concerned with distinguishing true effects in the presence of random variations in the data, and have been previously discussed in Sections 4.2.1.2 (constant stress) and 4.2.2.2 (step stress).

4.2.3.3 Relation of Time in Constant-Stress Testing to Time in Progressive-Stress Testing

For constant-stress testing at a stress level S_c, the cumulative damage at time t_c is given by

$$D(t_c) = R(S_c)\, t_c \quad , \tag{4.14}$$

where $R(S_c)$ is the rate of measured parameter degradation at stress level S_c. Alternatively, for progressive-stress testing, cumulative damage at time t_p is given by

$$D(t_p) = \int_o^{t_p} R(S_o + \beta t)\, dt \quad , \tag{4.15}$$

where S_o is the stress level at $t = 0$. Equating (4.14) and (4.15), constant-stress time, t_c, as a function of progressive-stress time is given by

$$t_c = \int_o^{t_p} R(S_o + \beta t)\, dt \Big/ R(S_c) \quad . \tag{4.16}$$

Downloaded from https://www.everyspec.com

If the stress parameter is temperature, and the Arrhenius law applies, then

$$R(T) = e^{A-B/T} .$$

In this case, Equation (4.16) takes the form

$$t_c = \int_o^{t_p} e^{A - B/(T_o + \beta t)}\, dt \Big/ e^{A - B/T_c}$$

$$= e^{B/T_c} \int_o^{t_p} e^{-B/(T_o + \beta t)}\, dt \quad . \tag{4.17}$$

Equation (4.17), however, does not have a closed form solution. Therefore, to obtain the desired relation in a particular case, it is necessary to resort to approximate numerical solutions of the integral.

Alternatively, if the Eyring equation is taken to represent degradation, and stress is increased in units of voltage for a fixed temperature[5,6], R(T,S) is given by Equations (4.9) and (4.10) in Section 4.2.2.3. Substituting Equation (4.10) into (4.16) yields

$$t_c = \int_o^{t_p} R_o V^{\phi}\, dt \Big/ R_o V_c^{\phi}$$

$$= \int_o^{t_p} (V_o + \beta t)^{\phi}\, dt \Big/ V_c^{\phi}$$

$$= \left[(V_o + \beta t_p)^{\phi+1} - V_o^{\phi+1} \right] \Big/ \beta(\phi+1)\, V_c^{\phi} \quad . \tag{4.18}$$

If the voltage stress is zero at t = o, (4.18) reduces to

$$t_c = (\beta t_p)^{\phi+1} \Big/ \beta(\phi+1)\, V_c^{\phi}$$

$$= t_p\, V_p^{\phi} \Big/ (\phi+1)\, V_c^{\phi} \quad . \tag{4.19}$$

where $V_p = \beta t_p$ is the progressive power stress.

4.2.3.4 Example of a Progressive-Stress Test Design

To illustrate the experimental procedures for progressive-stress testing, a capacitor life test design is described in the following paragraphs. The objective of the test is to determine the applicability of the Eyring rate equation in describing capacitor degradation processes.

Downloaded from https://www.everyspec.com

The test design involves two stress parameters: temperature and voltage. Voltage is taken as the progressive-stress parameter, and the various temperature levels are held fixed for each test in the total design as shown in Figure 4.8. As shown in this figure, the design consists of 5 different time rates of voltage increase, β_i, . . .i = 1,2,. . .,5, and 4 temperature levels, T_k, k = 1,. . .,4. Each voltage rate, β_i, appears once with each temperature level; thus, there are 20 different voltage rate-temperature combinations. Finally, it is seen that voltage rate is increased proportionally. That is, the ith voltage rate is given by

$$V_i = \beta_i t = \left(2^{i-1}\beta_1\right) t \ . \tag{4.20}$$

$$\ln V_i = (i-1) \ln 2\beta_1 + \ln t = (i-1)c + \ln t \ . \tag{4.21}$$

Therefore, the experiment involves life testing a statistically significant number of capacitors at each of 20 voltage rate-temperature combinations. Capacitor degradation is measured in terms of performance parameters such as capacitance, dissipation factor, or leakage current. The various tests are terminated according to a predetermined stopping rule. For example, each test may be terminated when 50 per cent of the capacitors on test have failed.

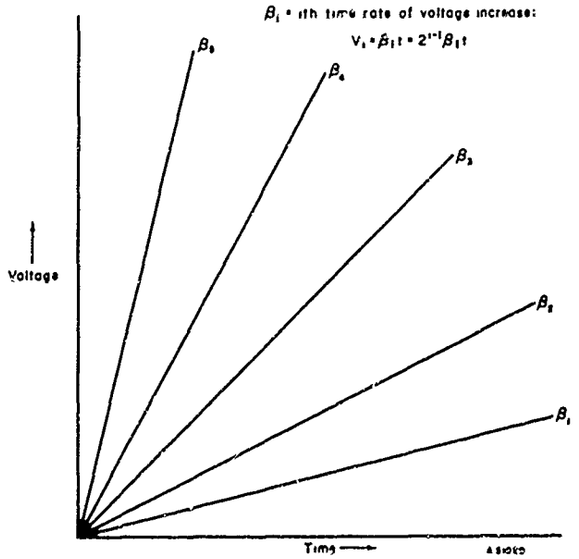


FIGURE 4.8. PROGRESSIVE-STRESS (VOLTAGE) TEST DESIGN LAYOUT

The design is repeated for each of four temperature levels.

4.3 Theoretical Models for Analyzing Test Data

In the history of life testing, various procedures have been derived for calculating acceleration factors. Most of these have been statistical in nature, in that acceleration factors were calculated by taking the ratios of failure rates, per cent surviving, cumulative

Downloaded from https://www.everyspec.com

or percentage drift in some parameters, etc., observed under "normal" and some high stress level. All of these procedures have several apparent shortcomings. First, they do not provide a basis for answering the question: Has "true" acceleration been achieved? That is, are the relative effects of the failure mechanisms acting in a device invariant over the range of stress levels involved? Second, they provide no basis for explaining the observed component behavior in terms of the physical structure of the component and its stress environment. The importance of these shortcomings is that purely predictive statistical data are of limited value to design improvement.

In the following paragraphs, two theoretical models for analyzing accelerated test data are described which show promise of alleviating the problems associated with the statistical procedures cited above. The description in each case includes a physical interpretation of the model including its advantages and limitations and specification of the procedures involved in applying the model to the analysis of accelerated test data.

4.3.1 The Arrhenius Model as a Basis for Accelerated Tests

The Arrhenius equation relates the time rate of change of a process to the temperature at which the process is taking place. Originally, this function was empirically derived by Arrhenius in the 1880's from data on reaction rates of chemical processes. More recently, in reliability the applicability of the Arrhenius law in describing component parameter degradation has been suggested by "rules of thumb" such as the observation that component life is halved for every 10°C increase in temperature. Further, for relatively simple devices such as resistors, degradation can be easily envisioned as being due to a chemical change in the device. For more complex devices, the physical nature of the degradation processes is not so obvious. However, the basic atomic processes such as diffusion that lead to degradation in many devices generally satisfy the conditions of the Arrhenius equation.

Attempts to derive the Arrhenius equation theoretically have not been generally successful. This follows from the fact that modern developments in quantum theory lead to a different (and more theoretically sound) description of reaction-rate kinetics. The Eyring reaction rate model described in Section 4.3.2 is a result of these developments. However, over the range of stress conditions normally used in accelerated testing, the Arrhenius model serves as an excellent first-order approximation to the Eyring model. In fact, within this range of stress conditions, the estimation error due to random variations in the test data will generally exceed the approximation error by a substantial amount. Moreover, even though the Eyring model has significantly more value in obtaining a theoretical understanding of aging processes, the simplicity of the Arrhenius model makes it highly useful in analyzing accelerated test data.

4.3.1.1 Mathematical Description of the Arrhenius Model[2,10]

Let Q denote a component performance parameter such as reverse-leakage current in a transistor, and let D = f(Q) denote a function of the performance parameter, Q, such as per cent change in reverse-leakage current in terms of which degradation is to be measured. Application of the Arrhenius model in calculating acceleration factors requires two assumptions:

Downloaded from https://www.everyspec.com

(1) Degradation in the performance measure, $D = f(Q)$ is a linear function of time, the rate of degradation depending only on the stress level.

(2) The logarithms of the degradation rates yield a linear function of reciprocal absolute temperature.

In theory, the first assumption can always be satisfied. That is, for any given set of experimental data, some transformation, $D = f(Q)$, of the performance parameter Q can be found that yields linear degradation over time. The practical problem is to find a transformation that will satisfy this assumption for all sets of experimental data on a given component type. The second assumption that is characterized graphically in the Arrhenius plot is the crux of the analysis. If the Arrhenius plot is linear, acceleration is said to be "true"; otherwise, the observed acceleration is not "true". This postulate follows from the expectation that changes in the relative effects of the failure mechanisms acting in a device at some increased stress level will yield a nonlinear Arrhenius plot. When this occurs, linear extrapolations to normal operating temperatures are not valid and, therefore, acceleration factors cannot be unambiguously associated with various temperature levels.

More specifically, the assumption of a linear degradation of performance over time can be written as follows:

$$f(Q) = R(T)\,t \quad , \tag{4.22}$$

where t denotes operative time under a specified temperature level T, and R(T) denotes a constant degradation rate that depends on the temperature level, T. This equation shows that R(T) is the slope of the degradation line obtained by plotting f(Q) versus t. A different slope R(T) is expected for each thermal-stress level, T, used in generating the accelerated test data.

The Arrhenius equation may be written as follows:

$$R(T) = e^{A-B/T} \quad , \tag{4.23}$$

where A and B denote empirical constants. If the logarithm of R(T) is plotted against (1/T) on an Arrhenius plot, then Constants A and B denote the "intercept" and "slope" of the resulting straight line, as shown by the expression

$$\ell n\, R(T) = A - B(1/T) \quad . \tag{4.24}$$

Suppose that the Arrhenius plot is linear. Then least-squares analysis may be used to fit Equation (4.24) to the data and obtain numerical estimates for Constants A and B. From these estimates the relation between accelerated time and normal time can be obtained as follows. The expression for R(T) given in Equation (4.23) is substituted into Equation (4.22) to yield

$$f(Q) = \left\{ e^{A-B/T} \right\} t \quad . \tag{4.25}$$

Downloaded from https://www.everyspec.com

We now denote by primes those quantities associated with accelerated stress levels (the quantities without primes refer to reference stress levels). Suppose the test times have been such that the same degradation of quality has occurred at the two stress levels. This means that $f'(Q) = f(Q)$; from Equation (4.25), it follows that

$$(e^{A-B/T'})\,t' = (e^{A-B/T})\,t \quad . \tag{4.26}$$

This equation may be solved for t to obtain

$$t = e^{-B(1/T' - 1/T)}\,t' \quad . \tag{4.27}$$

On the other hand, the acceleration factor is defined by the relation

$$t = \tau t' \quad , \tag{4.28}$$

so that Equations (4.27) and (4.28) yield

$$\tau = e^{-B(1/T' - 1/T)} \quad . \tag{4.29}$$

This is the mathematical formula for the acceleration factor obtained from the Arrhenius model. Thus, if a true accelerated test is run at an elevated thermal stress for t' hours, then this is equivalent to t hours of operation at normal stress, where $t = \tau t'$ and τ is given by Equation (4.29). It is of interest to note that the acceleration factor involves only the constant B. This means that the acceleration factor depends only on the slope of the Arrhenius plot and not on its "intercept".

4.3.1.2 Stepwise Procedures in Applying the Arrhenius Model

In the following paragraphs, step-by-step procedures are described for the analysis of accelerated test data using the Arrhenius model. Procedures are described separately for (1) constant-stress data, (2) step-stress data, and (3) progressive-stress data. Although separate procedures are described for step-stress and progressive-stress data, it should be recognized that these data can be transformed into equivalent constant-stress data according to the transformation functions given in Sections 4.2.2.3 and 4.2.3.3. Thus, in theory, accelerated test data generated by any of the three methods described in Section 4.2 can be analyzed according to the step-by-step procedures for constant-stress data.

4.3.1.2.1 Stepwise Analysis Procedures for Constant-Stress Data

Step 1. Obtain a Measure Yielding Linear Degradation Over Time.

A suitable measure, f(Q), may be known, a priori, from experience, or alternatively, may be determined from time-history plots of the performance parameter, Q. Candidate transformation functions, f(Q), include

Downloaded from https://www.everyspec.com

Transformation Function	D = f (Q)
The performance parameter	Q
Logarithms	log Q
Ratios to some reference value	Q/Q_o
Incremental values	ΔQ
Relative incremental values	$\Delta Q/Q_o$

Further, a summary measure, f(Q), is made on the basis of which degradation rate will be calculated. Here, the mean value, median value, etc., of the performance measure may be used [Figure 4.9(a)].

Step 2. Estimate the Slope (Rate) of the Degradation Lines.

Estimate the slope, $R(T_i)$, for each of the degradation lines that correspond to various temperature levels, T_i. The lines can be fitted using least-squares analysis.

Step 3. Construct the Arrhenius Plot.

For each temperature stress level, T_i, plot the logarithm of $R(T_i)$ versus reciprocal absolute temperature, $(1/T_i)$. [Figure 4.9(b)].

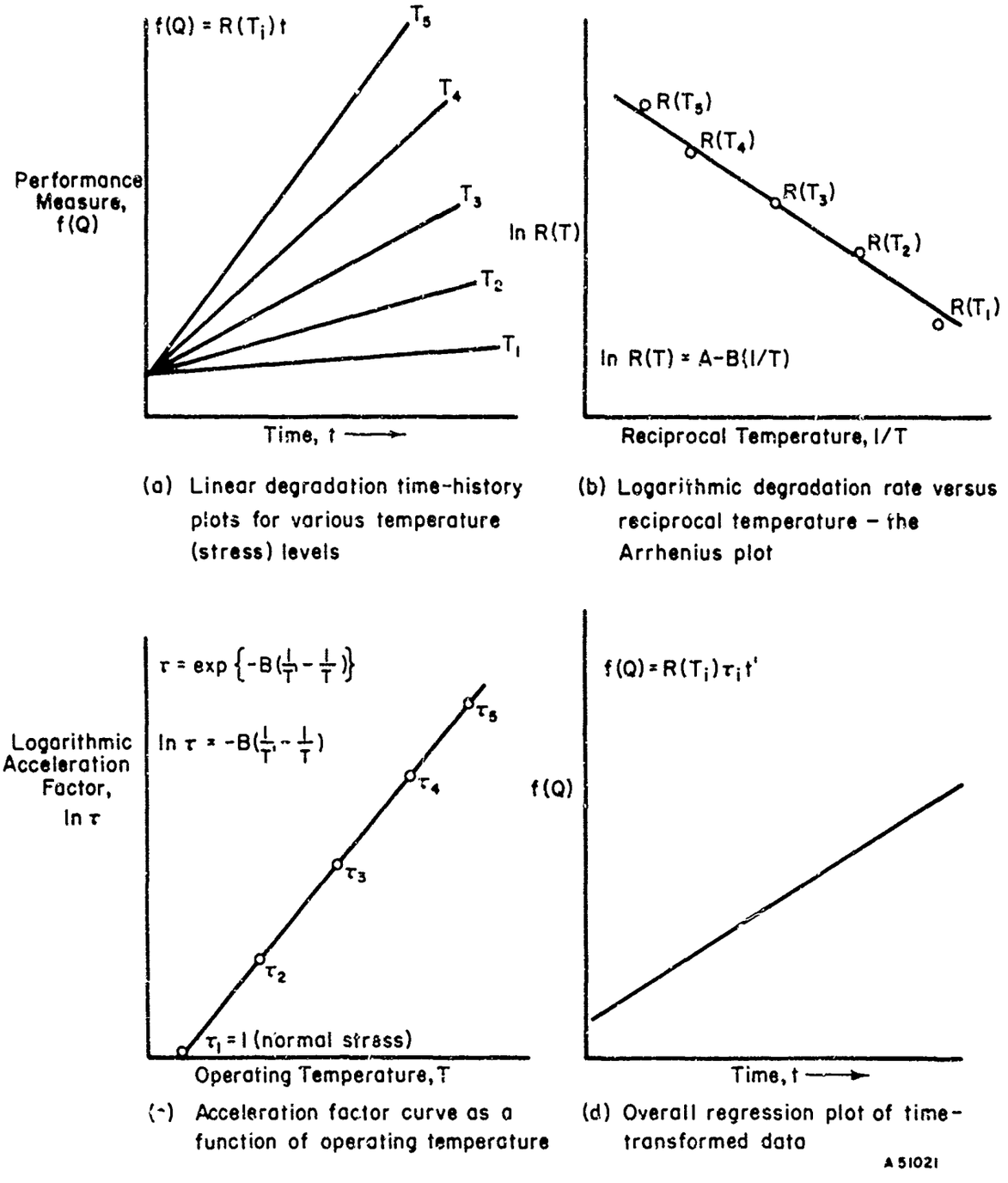
Step 4. Determine Whether "True" Acceleration Exists.

If the points plotted in Step 3 "line up" on the Arrhenius plot, then it is concluded that "true" acceleration exists among these stress levels. If only certain subsets of the points line up, then only these subsets are associated with true accelerations. The remaining points may be associated with changes in aging mechanisms, or other difficulties that threaten the validity of acceleration factors. In some instances it may be desirable to consider different transformations, f(Q), in order to obtain Arrhenius plots that are linear. Two different linear segments on the Arrhenius plot may indicate the simultaneous occurrence of two aging mechanisms. In those cases which do not yield Arrhenius plots that are linear, it would generally be concluded that true accelerations do not exist among the selected stress levels. It would then appear reasonable to suppose that the range of stress covered in the accelerated tests is too large. If new experimental programs are run with a sufficiently restricted range of thermal stress, the range over which true acceleration could be achieved would gradually be determined.

Step 5. Compute Acceleration Factors.

If the points line up on the Arrhenius plot, then a least-squares fit will yield the slope of the best fitting line. The numerical value of

Downloaded from https://www.everyspec.com



(a) Linear degradation time-history plots for various temperature (stress) levels

(b) Logarithmic degradation rate versus reciprocal temperature – the Arrhenius plot

(c) Acceleration factor curve as a function of operating temperature

(d) Overall regression plot of time-transformed data

A 51021

FIGURE 4.9. GRAPHICAL DESCRIPTION OF THE ANALYSIS OF CONSTANT-STRESS ACCELERATED TEST DATA USING THE ARRHENIUS MODEL

Downloaded from https://www.everyspec.com

this slope is equal to the constant B given in Equation (4.29). Thus, for any thermal stress T' and any reference stress T, Equation (4.29) permits the computation of the associated acceleration factor. In particular, if T denotes the normal operating temperature and T' denotes an increased level of temperature, then Equation (4.29) may be used to extrapolate from the test condition to the normal operation. The acceleration factor obtained in this way supposes that the linear Arrhenius plot remains linear when extrapolated to normal operating temperatures. The validity of this extrapolation must be carefully considered in each case. It must be noted that the extrapolation cannot be avoided except by testing component parts at normal stress levels. This is precisely the test condition that it is hoped to avoid by means of the accelerated test.

Step 6. Make a Final Correlation Plot.

Suppose the Arrhenius plot is linear and the acceleration factors have been computed for each of the stress conditions used in the accelerated test. Let τ_i denote the acceleration factor that corresponds to temperature level T_i. Then the linear degradations found in Step 1 may be replotted using a transformed time scale. The new time scale is obtained by multiplying the actual scale by τ_i. In this way all of the data may be plotted using equivalent hours of operation. If the Arrhenius model is valid, all of the data for all of the temperature levels will plot on a single straight line. This final correlation plot often serves to give a view of the strengths, or weaknesses, of the over-all analysis.

4.3.1.2.2 Stepwise Procedures for Analysis of Step-Stress Data

In Section 4.2.2, it was noted that, in step-stress testing, the roles of stress and time are in a sense reversed. Thus, the basic data generated in step-stress testing gives the stress level at which components fail (or at which some amount of damage is accumulated) for various stress-step time increments. Therefore, the step-by-step procedure in analyzing these data will differ from that for constant-stress data. Under the assumptions associated with step-stress testing and the Arrhenius model, the same acceleration factor should be obtained as in a corresponding constant-stress test.

Step 1. Obtain a Linear Damage Function Over Reciprocal Temperature

A measure of degradation, g(D), of the components within each step-stress time increment must be determined such that the plot of g(D) versus reciprocal temperature will be linear. For example, g(D) might be taken as the logarithm of the median cumulative damage, D, at the kth stress step. Alternatively, for failure data, g(D) would take the form of cumulative per cent of the sample units failing at the kth stress step measured on that probability scale that yields the required linear relationship.

Downloaded from https://www.everyspec.com

Step 2. Construct a plot of g(D) versus reciprocal temperature, 1/T, for each set of data corresponding to a given time increment, Δt_{ss} [Figure 4.10(a)].

Step 3. Determine Whether "True" Acceleration Exists

Calculate the slopes of each of the lines in the plot obtained in Step 2. If the calculated values are equal in a statistical sense, "true" acceleration is said to occur. This follows from the expectation that, if the relative effects of the aging processes change at some increased stress level, the slopes of the damage lines will change also.

Step 4. Plot Stress Versus Incremental Time

Plot the logarithm of the time interval of the stress step, Δt_{ss}, versus reciprocal temperature. This can be done for various values of constant damage (e.g., 1 per cent, 10 per cent, 50 per cent failure). These lines will be parallel (have a constant slope, B) if true acceleration exists [Figure 4.10(b)].

Step 5. Calculate Acceleration Factors

If true acceleration exists, acceleration factors can be calculated using Equation (4.29). The parameter B in Equation (4.29) is the value of the common slope of the lines obtained in Step 4. The explanation of the acceleration factors here is identical to that given in Step 5 in Section 4.3.1.2.1, "Stepwise Analysis Procedures for Data" [Figure 4.10(c)].

4.3.1.2.3 Stepwise Procedure for Analyzing Progressive-Stress Data

In progressive-stress testing, the stress level is a continuously increasing function of time. Accelerated test data is generated by running tests at several different time rates of stress increase. Thus, as in the case of step-stress testing, the step-by-step procedures for calculating acceleration factors from progressive-stress test data will differ from those for constant-stress data.

Step 1. Obtain Linear Degradation Over Time

Obtain a transformation of the damage function, g(D), that yields a linear plot of g(D) versus time for each of the time rates, β, of progressive stress. For example, if D denotes cumulative per cent failing over time, g(D) will be that probability scale that yields a linear time plot [Figure 4.11(a)]. Various standard probability papers (Weibull, lognormal, etc.) can be used for this purpose.

Downloaded from https://www.everyspec.com

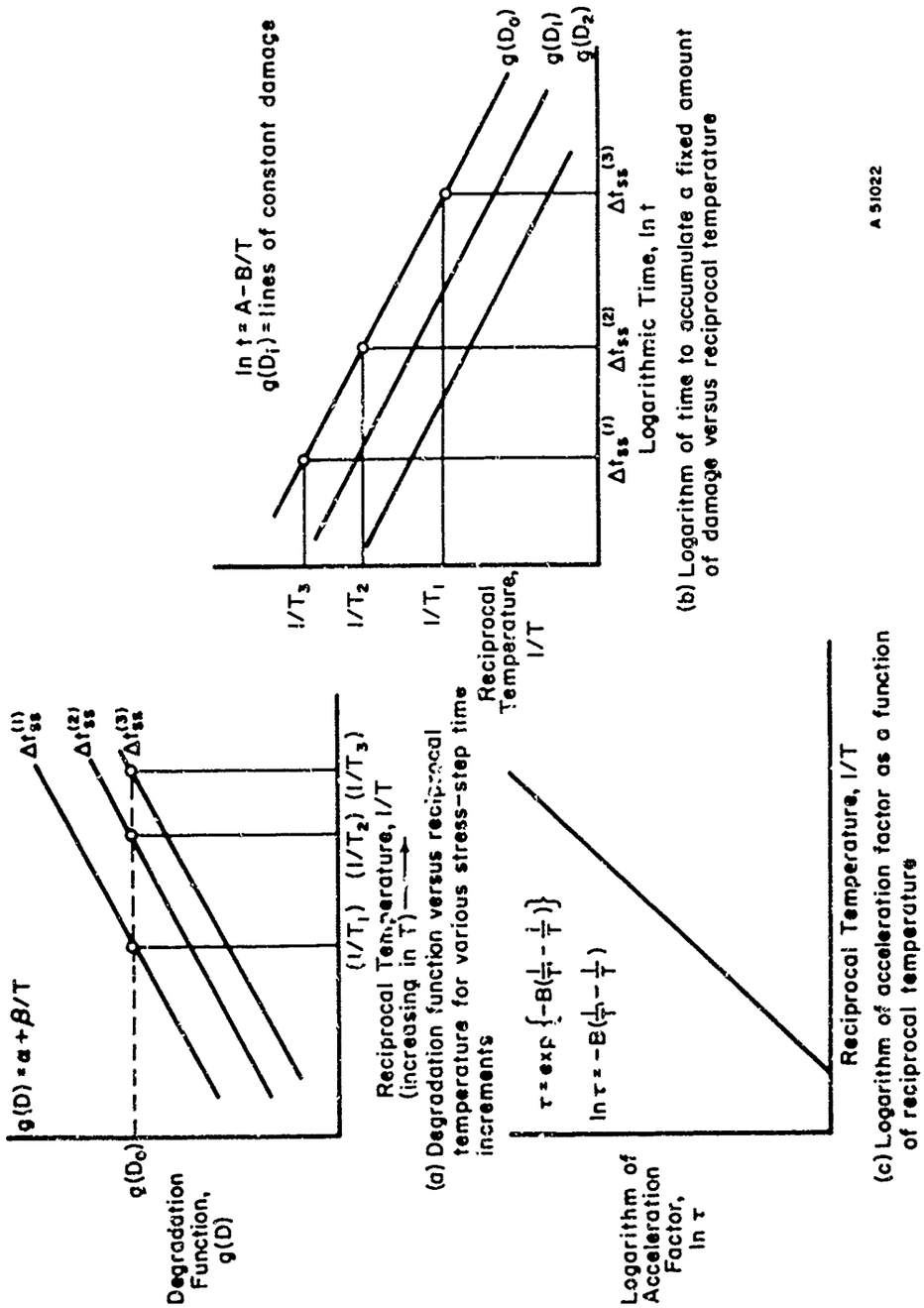


FIGURE 4.10. GRAPHIC DESCRIPTION OF THE ANALYSIS OF STEP-STRESS ACCELERATED TEST DATA USING THE ARRHENIUS MODEL

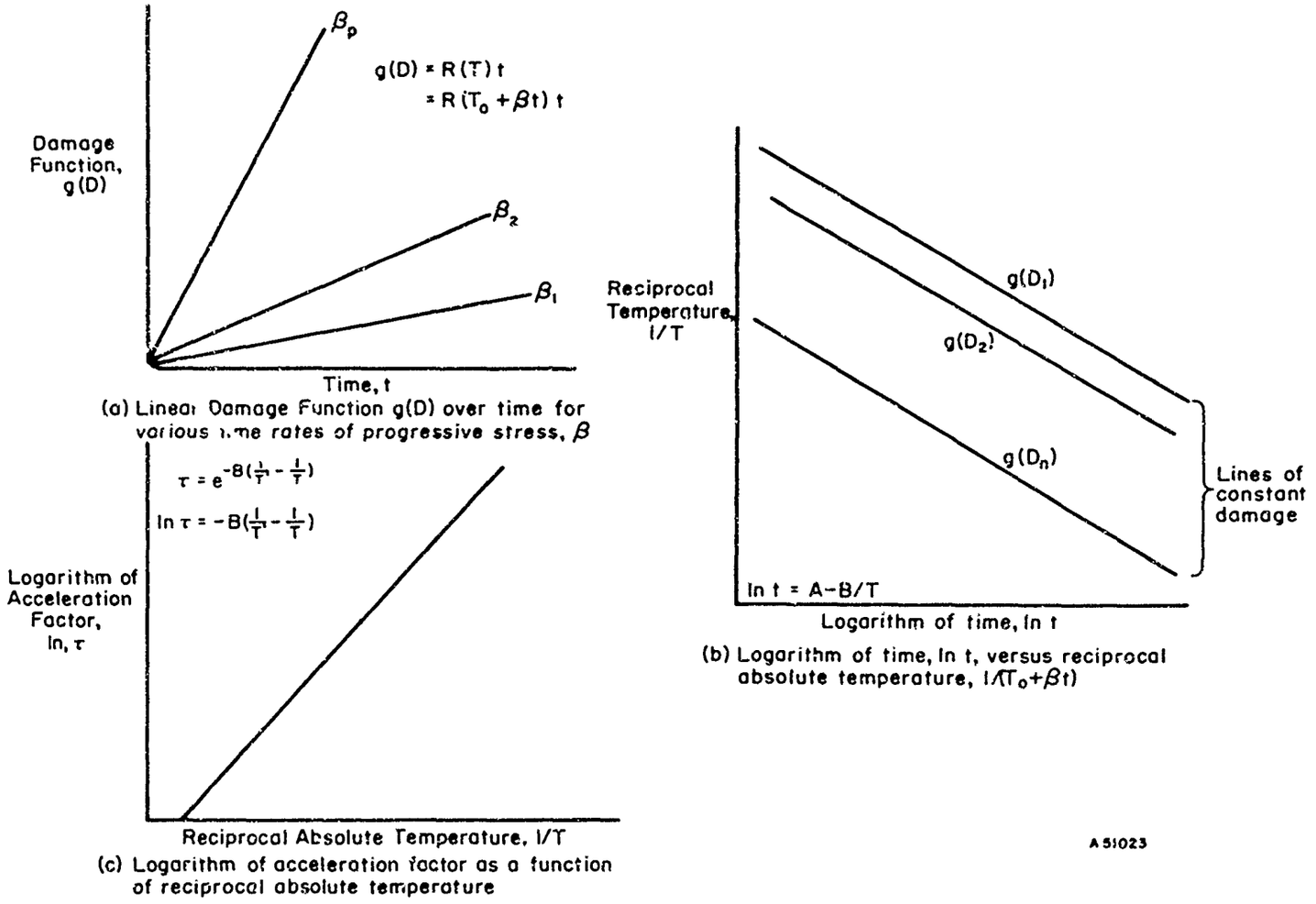


FIGURE 4.11. GRAPHIC DESCRIPTION OF THE ANALYSIS OF PROGRESSIVE-STRESS TEST DATA USING THE ARRHENIUS MODEL

Downloaded from https://www.everyspec.com

Downloaded from https://www.everyspec.com

Step 2. Construct Plots of the Time-Stress Relationship

For several fixed values of the damage function, $g(D_1), \ldots, g(D_n)$, plot reciprocal absolute temperature, $1/T = 1/(T_o + \beta t)$, versus the logarithm of time [Figure 4.11(b)].

Step 3. Determine Whether "True" Acceleration Exists

Calculate the slopes of each of the lines obtained in Step 2 above. If the slopes of these lines are equal in a statistical sense, true acceleration is said to be achieved. However, if at some value of the damage function, say $g(D_k)$, the slopes of the lines for $g(D) > g(D_k)$ change, it is concluded that the dominant aging mechanism has also changed. An explanation of the effect of such an event on the acceleration factor problem is analogous to that given in Step 4 for constant-stress test data analysis.

Step 4. Calculate Acceleration Factors

If true acceleration exists, acceleration factors can be calculated using Equation (4.29). The constant B in Equation (4.29) is the value of the common slope of the lines obtained in Step 3. Thus, for any thermal stress, T', and a reference stress, T (where T has the form $T_o + \beta t$), Equation (4.29) permits the calculation of the associated acceleration factor. In particular, if T denotes the normal operating temperature and T' denotes an increased level of temperature, then Equation (4.29) may be used to extrapolate between the test condition and normal operation. The acceleration factor obtained in this way supposes that the linear Arrhenius plot remains linear when extrapolated to normal operating temperatures. The validity of this extrapolation must be carefully considered in each case. It must be noted that the extrapolation cannot be avoided except by testing component parts at normal stress levels. This is precisely the test condition which it is hoped to avoid by means of the accelerated test.

4.3.1.3 Extension of the Arrhenius Model to Include Nonthermal Stress

In the preceding paragraphs, procedures were given for analyzing accelerated test data where the only stress parameter was temperature. These analysis procedures were based on the basic Arrhenius model given by

$$R(T) = e^{A-B/T} \quad . \tag{4.30}$$

This model can be modified to include nonthermal stress by multiplying Equation (4.29) by an exponential term including the nonthermal stress parameter. The resultant modified Arrhenius model is given by

$$R(T,S) = e^{A-B/T}\, e^{S(C+D/T)} \quad , \tag{4.31}$$

Downloaded from https://www.everyspec.com

where S denotes the nonthermal stress parameter (e.g., voltage) and C and D denote constants analogous to A and B to be determined from the analysis of accelerated test data.

Acceleration factors can be obtained from the modified Arrhenius model in the same way as was shown in Section 4.3.1.1. If a performance measure, f(Q), can be found such that

$$f(Q) = R(T,S)\,t \tag{4.32}$$

(i.e., f(Q) is a linear function of time where the time rate of degradation depends on T and S), then the acceleration resulting from testing at some stress level, T' S', is obtained by equating

$$f'(Q) = R(T',S')\,t' \tag{4.33}$$

and Equation (4.32), where f(Q) and f(Q') denote an equal degradation level. Since the acceleration is defined by the relation $t = \tau t'$, from Equation (4.32) and (4.33), τ is seen to be

$$\tau = R(T',S')/R(T,S) \quad . \tag{4.34}$$

Substituting Equation (4.31) into (4.34) yields

$$\tau = \exp\left\{-B\left[\frac{1}{T'} - \frac{1}{T}\right] + C(S'-S) + D\left[\frac{S'}{T'} - \frac{S}{T}\right]\right\}. \tag{4.35}$$

Equation (4.35) shows that the acceleration factor is a function of reciprocal temperature, nonthermal stress, and an interaction between reciprocal temperature and nonthermal stress.

Stepwise analysis procedures of accelerated test data in this case are analogous to those given previously, although somewhat more complex. Moreover, the procedures in this case are almost identical to those associated with the Eyring model described in Section 4.3.2 and, therefore, are not given in this section.

4.3.2 The Eyring Model as a Basis for Accelerated Testing

The fundamental problem in accelerated testing is that rates of aging under various stresses must be understood in terms of basic material properties and physical processes. Otherwise, extrapolations from accelerated stress conditions to normal operating conditions cannot be completely validated. The value of the Arrhenius model described in Section 4.3.1 stems from the fact that it "worked" in practice in spite of its lack of theoretical validity. Because of its empirical basis, reasons for the success of the Arrhenius model in correlating accelerated tests for electronic devices is also not well understood.

The Eyring reaction-rate equation, because of its theoretical basis, shows considerable promise in yielding the desired understanding cited above. This model is derived from

Downloaded from https://www.everyspec.com

the principles of quantum mechanics and has been successfully applied to a wide variety of physical and chemical rate processes. A modified version of the Eyring model, which includes nonthermal stress, is described in the following paragraphs.

4.3.2.1 Physical Concept of the Eyring Model

Consider the following simplified model of an electronic component. It is assumed that the performance of the component can be characterized by a parameter ξ. For a given range of ξ, the component is considered "good"; for a different range of ξ, the component is considered "bad" or to have failed. It is further assumed that a potential energy, $V(\xi)$, exists which depends on the value of the parameter ξ. Because in actual operation, component performance generally goes from "good" to "bad", it is assumed that the potential energy associated with a "good" component $V(\xi_{good})$ is greater than that associated with a "bad" component $V(\xi_{bad})$. This is shown in Figure 4.12.

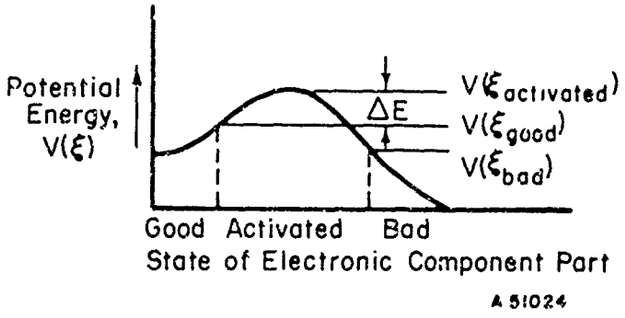


FIGURE 4.12. POTENTIAL-ENERGY DIAGRAM FOR THE AGING OF AN ELECTRONIC COMPONENT PART

Because the transistion from good to bad usually requires an interval of time, it is assumed that the good components are in a "metastable state". This means that the potential energy curve has a "hump" between the good and bad states. The energy of this "hump", shown in Figure 4.12, corresponds to the energy of the activated state. The energy that must be supplied in order for the component to go from "good" to "bad" is called the "activation energy", and is denoted by ΔE in the sketch.

In a solid at temperature where $kT < \Delta E$, thermal fluctuations may result in localized values of $kT > \Delta E$ in small incremental volumes (k is Boltzmann's constant and T is the absolute temperature). The probability that the activation energy is spontaneously supplied by thermal fluctuations is proportional to $\exp(-\Delta E/kT)$. Whenever this occurs, the associated incremental volume may make the transition to "bad". After sufficient time, the entire component will be composed of "bad" incremental volumes and will itself be "bad".

The Eyring reaction-rate theory shows that the time rate, R_o, for the transition from "good" to "bad" stated for the entire component is given by the expression

$$R_o = a(kT/h)e^{-b/T} \quad , \tag{4.36}$$

Downloaded from https://www.everyspec.com

in which h is Planck's constant, a depends on the partition functions of the "good" and "activated" states, and b denotes the ratio of the activation energy to Boltzmann's constant, $b = \Delta E/k$. A theoretical evaluation of a and b is difficult. However, an empirical determination can be made by taking the logarithm of the Eyring equation, to obtain

$$\log(hR_o/kT) = \log a - b(1/T) \quad . \tag{4.37}$$

It is seen that a plot of the logarithm of (hR_o/kT) versus $(1/T)$ will give a straight line with "intercept" log a and scope b.

A detailed application of this theory presents many difficulties. First, a suitable parameter ξ has to be determined. Then the potential energy levels $V(\xi)$ have to be determined, and the range of values for ξ_{bad}, $\xi_{activated}$, and ξ_{good} have to be defined. The activation energy is required to evaluate the constant b, and a theoretical determination of a presents many difficulties because of the partition functions involved. This theoretical picture is further complicated by the fact that there may be several "paths" by which the potential energy may degrade from a "good" state to a "bad" state. These "paths" may be associated with various "modes of failure" for the component part.

Because of these difficulties, it appears necessary to determine by experimental methods those physical processes involved in the aging of electronic-component parts. These experiments may use constant-stress, step-stress, or progressive-stress accelerated tests. However, constant-stress testing appears to be best suited to the task. As the experiment proceeds over time, the task consists of detecting changes in measurement parameters and subjecting these changes to graphical analysis using the Eyring model. This procedure will yield estimates of activation energies and thereby aid in identifying the physical mechanisms associated with the aging process.

4.3.2.2 Generalized Eyring Model to Include Nonthermal Stresses[2]

The theoretical basis of the Eyring model suggests that nonthermal stresses may be included in the model as follows. Let R_o denote the Eyring rate in the presence of thermal stresses only, so that R_o is given by

$$R_o = ATe^{-B/kT} \quad . \tag{4.38}$$

In the presence of a nonthermal stress S, this rate of degradation is multiplied by two factors, f_1 and f_2, to give

$$R = R_o f_1 f_2$$

for the modified rate equation. The factor f_1 is applied to adjust the energy distributions for the presence of nonthermal stress, and the factor f_2 is applied to adjust the activation energy for the presence of nonthermal stress. Mathematically, these factors are written as

$$f_1 = e^{CS}$$

Downloaded from https://www.everyspec.com

and

$$f_2 = e^{DS/kT} ,$$

where C and D denotes constants. It now follows that

$$R(T,S) = [ATe^{-B/kT}]\ e^{(C+D/kT)S} = R_o e^{\phi S} , \tag{4.39}$$

where

$$R(T,O) = R_o = ATe^{-B/kT} , \tag{4.40}$$

and

$$\phi = C + D/kT . \tag{4.41}$$

The analysis of accelerated test data using the Eyring model is similar to that for the Arrhenius model. In particular, it is supposed that linear degradations have been obtained (by a transformation, if necessary) of a measured parameter of interest. Let primes denote conditions at high stress and write

$$f(Q) = R(T,S)t$$

and

$$f'(Q) = R(T',S')t'$$

for the degradation measure, Q, under normal and accelerated test conditions, respectively. The "slope" R(T,S) is shown to be a function of the operating temperature T and the non-thermal stress S. Now suppose the stress times are such that the same degradation exists under the normal and accelerated test conditions. Then f(Q) = f'(Q) and it follows that

$$R(T,S)t = R(T',S')t'$$

so that

$$t = [R(T',S')/R(T,S)]\ t' .$$

The acceleration factor τ is defined by the relation $t = \tau t'$ so that

$$\tau = R(T',S')/R(T,S) . \tag{4.42}$$

From Equation (4.39), it follows that

$$\tau = \left[\frac{T'}{T}\right] \exp\left\{-\frac{B}{k}\left[\frac{1}{T'} - \frac{1}{T}\right] + C(S'-S) + \frac{D}{k}\left[\frac{S'}{T'} - \frac{S}{T}\right]\right\} \tag{4.43}$$

This is the formula for the acceleration factor obtained from the Eyring model.

Downloaded from https://www.everyspec.com

Two special cases of acceleration can be obtained from Equation (4.43). First, if acceleration is achieved by testing at some increased nonthermal stress level, S', and normal operating temperature is used, i.e., T = T', then Equation (4.43) reduces to

$$\tau_S = \exp\left\{\left[C + \frac{D}{kT}\right](S' - S)\right\} = e^{\Phi(S' - S)} \quad . \tag{4.44}$$

Alternatively, if acceleration is achieved by testing at some increased temperature level, T', and S' = S, Equation (4.43) reduces to

$$\tau_T = \frac{T'}{T} \exp\left\{-\frac{1}{k}(B - DS)\left[\frac{1}{T'} - \frac{1}{T}\right]\right\} \quad . \tag{4.45}$$

4.3.2.3 A Further Modification of the Eyring Model

In the preceding section, the Eyring model was modified to include a nonthermal stress parameter, S. In general, however, R may not be a simple exponential function of the parameters. Therefore, it may be more appropriate to replace S with some function, f(S), of the nonthermal stress parameter. Thus, modifying the Eyring model given by Equation (4.40) to include this generalization yields

$$R(T,S) = ATe^{-B/kt}\, e^{f(S)(C + D/kt)} = R_o e^{\Phi f(S)} \quad . \tag{4.46}$$

For example, test results on dielectric materials and components indicate that f(S) can be expressed as the logarithm of voltage at a constant temperature. Thus, substituting f(S) = ln V into Equation (4.46) yields

$$R(T,S) = R_o e^{\Phi \ln V} = R_o V^{\Phi} \quad . \tag{4.47}$$

The corresponding formulas for acceleration factors are identical to Equations (4.43) to (4.45) with f(S) substituted for S. For example, acceleration due to increased voltage alone is given by Equation (4.44):

$$\tau = e^{\Phi(\ln V' - \ln V)} = (V'/V)^{\Phi} \quad . \tag{4.48}$$

4.3.2.4 Stepwise Procedures for the Analysis of Accelerated Test Data Using the Modified Eyring Model

In the following sections, step-by-step procedures are given for analyzing accelerated test data generated by constant-stress, step-stress, and progressive-stress testing. The analysis procedures are based on the modified Eyring model given by Equation (4.39). In those cases where some functions, f(S), of the nonthermal stress parameter is used, the procedures are the same with f(S) substituted for S.

Downloaded from https://www.everyspec.com

4.3.2.4.1 Stepwise Procedures for the Analysis of Constant-Stress Data

Step 1. Obtain Linear Degradations Over Time

This procedure is carried out in the same manner as indicated in Section 4.3.1.2.1 for constant-stress testing [Figure 4.13(a)].

Step 2. Determine Degradation Rate for Thermal Stresses

From Equation (4.39), it is seen that

$$\ln R(T,S) = \ln R(T,O) + \phi S \quad , \tag{4.49}$$

so that a plot of ln R(T,S) versus S yields a straight line with "intercept function" ln R(T,O) and "slope function" ϕ. More specifically, for a given temperature, T, plot the logarithm of the observed degradation rate as a function of S. This will yield a straight line with intercept ln $R(T_1,O)$. Repeat this procedure for T_2, T_3, etc., and obtain the intercept function, ln R(T,O), as a function of temperature [Figure 4.13(b)]. The temperature used in this procedure may be ambient temperature, junction temperature, or whatever temperature is associated with the physical processes of the electronic part.

Step 3. Determine the Constants A and B

With R(T,O) obtained in Step 2, plot ln R(T,O)/T versus (1/T) to get ln R(T,O)/T = ln A-(B/k) (1/T) as shown by Equation (4.40) [Figure 4.14(a)]. From this plot, the intercept, ln A, and slope, B/k, can be numerically determined so that A and B can be evaluated.

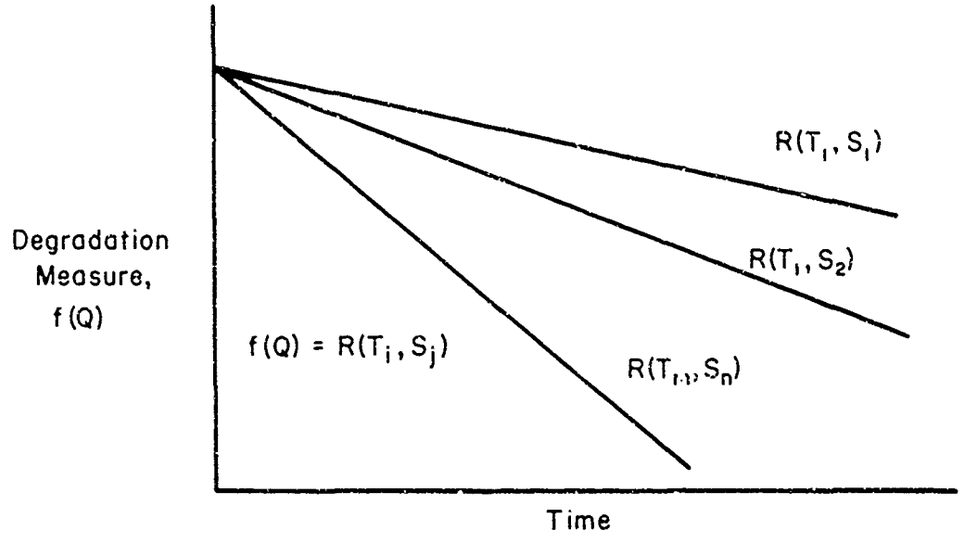
Step 4. Determine the Constants C and D

In Step 2, the values of ϕ were determined for various values of T. Equation (4.41) shows that a plot of ϕ versus (1/T) will have intercept C and slope D/k so that C and D can be numerically evaluated from such a plot [Figure 4.14(b)]. This step completes the determination of the four constants, A, B, C, and D, which occur in the Eyring model given by Equation (4.39). The model then gives the rate of aging for any temperature, T, and nonthermal stress, S, over the ranges of stress with which a given aging mechanism is dominant.

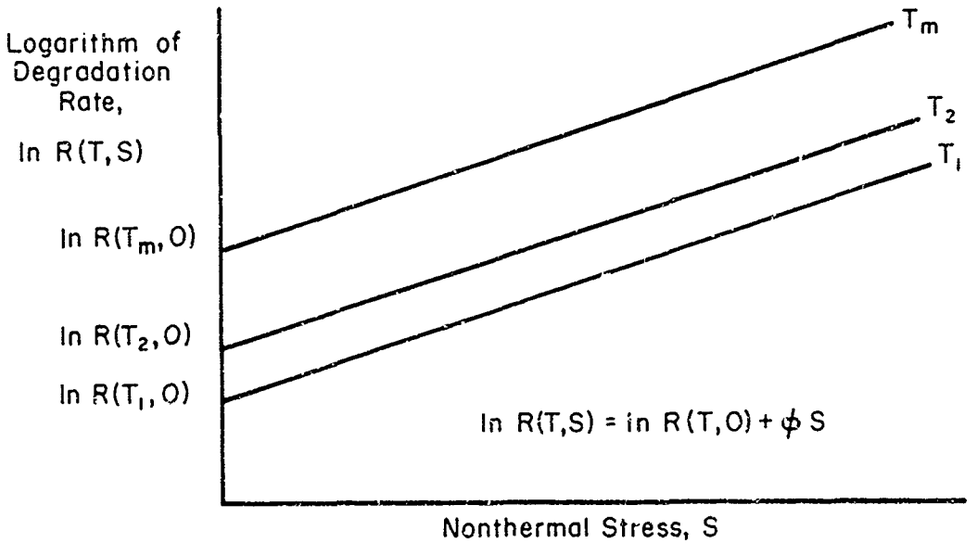
Step 5. Compute Acceleration Factors

Acceleration factors may be computed for those sets of test conditions that are found to be collinear in Step 3. For a specified accelerated test condition characterized by a temperature T' and a nonthermal stress S', and a reference condition characterized by T and S, the acceleration factor relative to the reference condition is given by Equation (4.43) [Figure 4.15(a)]. Because the acceleration

Downloaded from https://www.everyspec.com

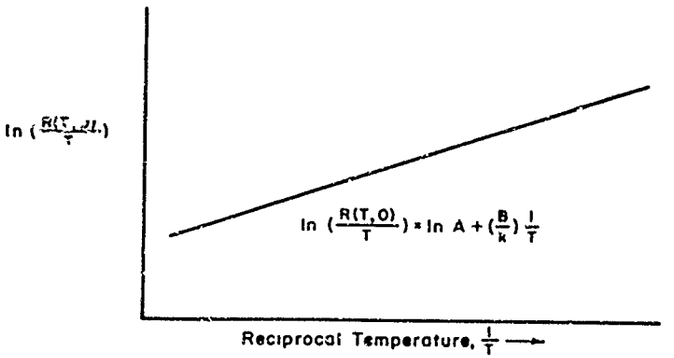


(a) Obtain linear degradation over time

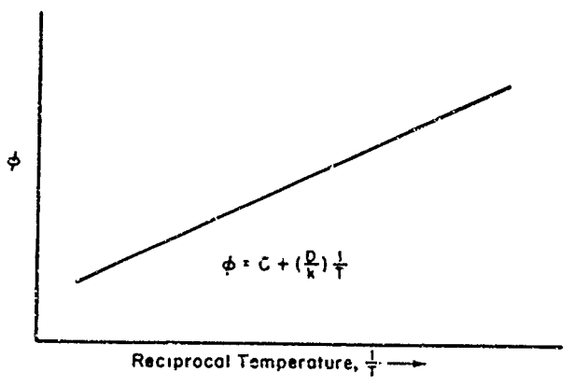


(b) Logarithm degradation rate as a function of nonthermal stress for various temperature levels - the Eyring plot A51025

FIGURE 4.13. GRAPHICAL DESCRIPTION OF THE ANALYSIS OF CONSTANT-STRESS DATA USING THE EYRING MODEL

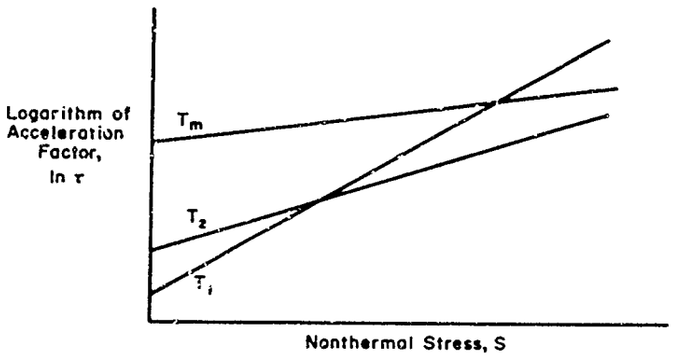


(a) Determination of the parameters A and B of the Eyring model

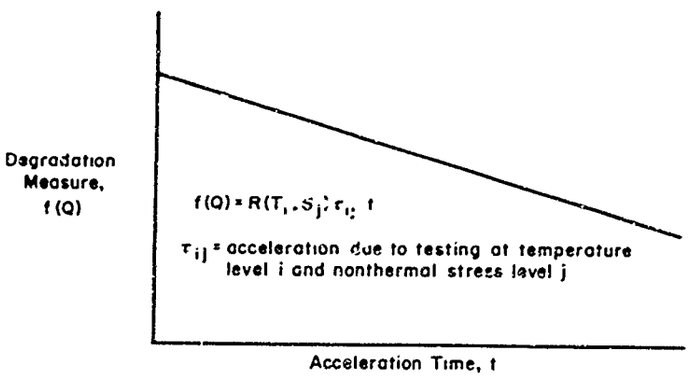


(b) Determination of the parameters C and D of the Eyring model

FIGURE 4.14. GRAPHICAL DESCRIPTION OF THE ANALYSIS OF CONSTANT-STRESS DATA USING THE EYRING MODEL



(a) Plot of logarithm of the acceleration factor, τ_i as a function of nonthermal stress for various temperature levels



(b) Final correlation plot of degradation plots adjusted to a common scale

A 51026

FIGURE 4.15. GRAPHICAL DESCRIPTION OF THE ANALYSIS OF CONSTANT-STRESS DATA USING THE EYRING MODEL

Downloaded from https://www.everyspec.com

Downloaded from https://www.everyspec.com

factor is a function of both temperature and nonthermal stress, a plot of $\tau'(T,S)$ as a function of temperature for various levels of nonthermal stress would yield a convenient set of curves for the acceleration factor.

Step 6. A Final Correlation

The acceleration factors obtained in Step 5 may be used to transform actual operating times at accelerated conditions to equivalent operating times at the reference condition. The linear degradations obtained in Step 1 should then be plotted against equivalent time [Figure 4.15(b)]. The scatter of the data about an extrapolated straight line representing the reference conditions will yield an over-all measure of the success of the analysis.

4.3.2.4.2 Stepwise Procedures for Analyzing Step-Stress Data

Accelerated step-stress tests can be run using (1) temperature as the step-stress parameter and repeating the test design for several fixed levels of the nonthermal stress parameter or (2) the nonthermal parameter may be stepped and the test design repeated for several fixed temperature levels. The most appropriate method will depend on the device under investigation. Therefore, step-by-step procedures are given for both of these testing methods.

In the first case, where temperature is the step-stress parameter, graphical analysis of the data does not yield estimates of the parameters A and B of the Eyring model. This follows from the fact that the linear function

$$\ln t = a + b/T \tag{4.50}$$

is obtained corresponding to a line of constant damage, $g(D_o)$, across the basic data plots of $g(D)$ versus $(1/T)$. For acceleration due to temperature defined by $t = \tau t'$, and for a fixed nonthermal stress level, Equation (4.50) yields

$$\tau_T = e^{-b\left[\frac{1}{T'} - \frac{1}{T}\right]} \quad . \tag{4.51}$$

This result, however, lacks the pre-exponential term (T'/T) given in Equation (4.45) for appropriate acceleration factor due to temperature based on the Eyring model.

In the second case, where nonthermal stress is the step stress parameter and temperature is fixed at a constant but different level for each test, estimates for all four Eyring model parameters can be obtained. For fixed temperature, T, the acceleration attained by increasing the applied stress to S' is obtained from Equation (4.43) as

$$\tau = e^{(C + D/kt)(S' - S)} = e^{\phi(S' - S)} \quad , \tag{4.52}$$

Downloaded from https://www.everyspec.com

or equivalently

$$\ln \tau = \phi(S' - S) \quad . \tag{4.53}$$

Step 2 of the analysis procedure requires a plot of the logarithm of applied stress versus the logarithm of step-stress time, yielding a linear equation

$$\ln t = a - b\,S \quad . \tag{4.54}$$

For the acceleration factor τ defined by $t = \tau t'$, Equation (4.53) yields

$$\ln \tau = \ln t - \ln t' = b(S' - S) \quad . \tag{4.55}$$

Thus, the slope b of the ln t versus S plot is equal to ϕ. Therefore, the plot yields the acceleration due to the nonthermal stress parameter, S (for fixed temperature).

To determine the parameters A and B of the modified Eyring model, observe that, in the absence of applied voltage, acceleration is obtained from Equation (4.45) as

$$\tau = \frac{T'}{T} \exp\left\{ - \frac{B}{K}\left[\frac{1}{T'} - \frac{1}{T}\right] \right\} \quad , \tag{4.56}$$

or equivalently

$$\ln \tau = \ln\left(\frac{T'}{T}\right) - \frac{B}{k}\left(\frac{1}{T'} - \frac{1}{T}\right) \quad . \tag{4.57}$$

The constant B in Equation (4.57) is obtained in the analysis of step-stress data by plotting (a + ln T) versus reciprocal temperature, (1/T), yielding an equation of the form

$$a + \ln T = \alpha + \beta/T \quad . \tag{4.58}$$

The parameter a is the constant term in Equation (4.54) and will be different for each temperature level at which a nonthermal step-stress test is run. For S = O,

$$\ln \tau = \ln t - \ln t' = a - a' \tag{4.59}$$

$$= \ln\left(\frac{T'}{T}\right) - \beta\left(\frac{1}{T'} - \frac{1}{T}\right) \quad ; \tag{4.60}$$

hence, $\beta = B/k$. Further, since

$$\tau = R(T',O) \Big/ R(T,O) = AT'e^{-B/kt} \Big/ ATe^{-B/kt} \quad , \tag{4.61}$$

it is seen that $\alpha = \ln A$ since Equation (4.54) also yields

$$\tau = t/t' = a/a' = e^{\alpha} T' e^{-\beta/kT'} \Big/ e^{\alpha} Te^{-\beta/kt} \quad . \tag{4.62}$$

Hence, the step-by-step analysis procedure generates a sequence of linear functions whose slopes and intercepts determine the constants A, B, C, and D of the modified Eyring model.

Downloaded from https://www.everyspec.com

Step 1. Obtain a Linear Damage Function Over Nonthermal Stress

For a constant temperature, T_1, plot damage, D, versus nonthermal stress, S, for each stress time increment, Δt_{ss}. Find a transformation, g(D), of the measure of damage that makes the plots linear. For example, if D denotes cumulative per cent failing at each successive stress step, g(D) is that probability scale yielding a linear relationship against S. This procedure is repeated for several constant temperature levels, $T_1, \ldots, T_m$ (Figure 4.16).

Step 2. Determine Relationship of Applied Stress to Step-Stress Time

For each of the constant temperature levels, $T_1, \ldots, T_m$, and for a specified value of g(D) (e.g., 50 per cent of test units failed), plot applied stress, S, versus the logarithm of step-stress time, Δt_{ss} [Figure 4.16(a)]. The equation for these lines is of the form $\ln t = a - \phi S$. Constants a and ϕ of each line (each temperature level) can be determined using least-squares analysis.

Step 3. Determine the Constants C and D

Plot the values of ϕ obtained in Step 2 versus reciprocal temperature [Figure 4.17(b)]. The resultant equation is given by $\phi = C + D/kT$. Hence, C is the intercept and D/k is the slope of the line. These values can be determined using least-squares analysis.

Step 4. Determine the Constants A and B

Plot $h(a,T) = a + \ln T$ versus reciprocal temperature [Figure 4.18(a)]. The resultant line is given by

$$a + \ln T = \ln A + B/kT$$

Thus, the constant A is determined by the slope and the constant B by the intercept of the line.

Step 5. Calculate Accelerated Factors

Acceleration factors can be calculated over the range of temperature and voltage stress levels for which the relative effects of the aging processes are invariant (Steps 1 and 2). Thus, for specified stress levels T' and S' and a pair of reference stress levels, T and S, the acceleration factor relative to the reference condition is given by Equation (4.43) [Figure 4.18(b)]. Plots of acceleration factors can be constructed identical to those for constant stress data (Figure 4.18).

4.3.2.3.3 Stepwise Analysis Procedures for Progressive-Stress Data

Procedures for calculating acceleration factors based on progressive-stress accelerated test data are similar to those for step-stress data. Further, if temperature is

Downloaded from https://www.everyspec.com

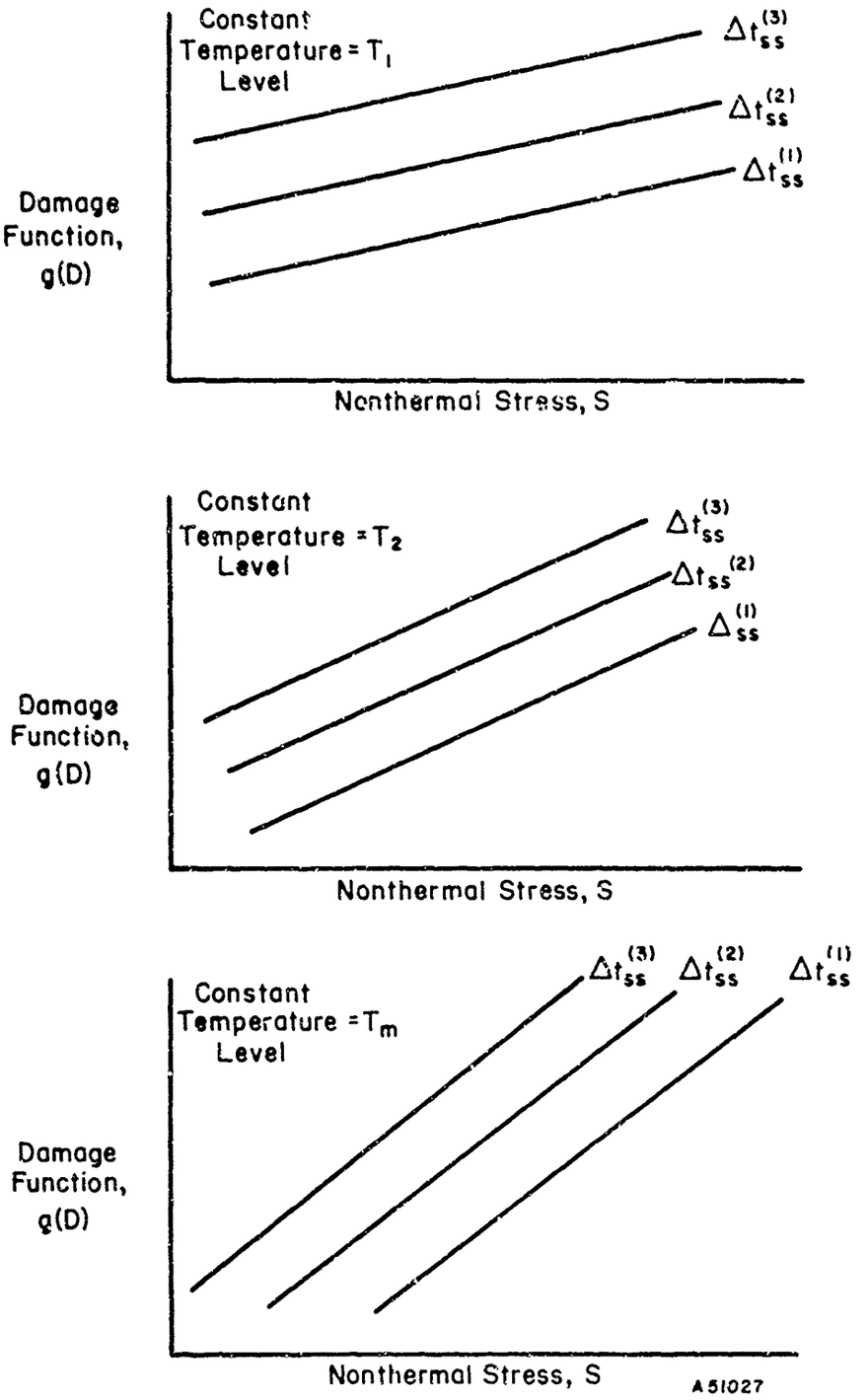
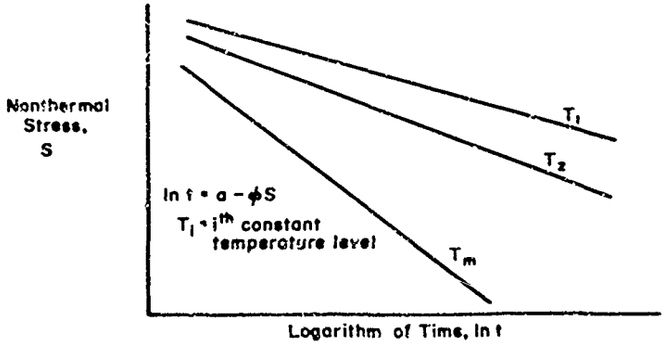
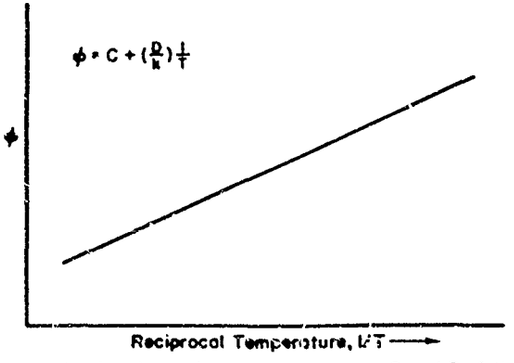


FIGURE 4.16. LINEAR CUMULATIVE DAMAGE PLOTS AS A FUNCTION OF NONTHERMAL STRESS FOR VARIOUS TEMPERATURE LEVELS IN THE ANALYSIS OF STEP STRESS DATA USING THE EYRING MODEL

(Δt_{ss} = incremental time in each stress step)

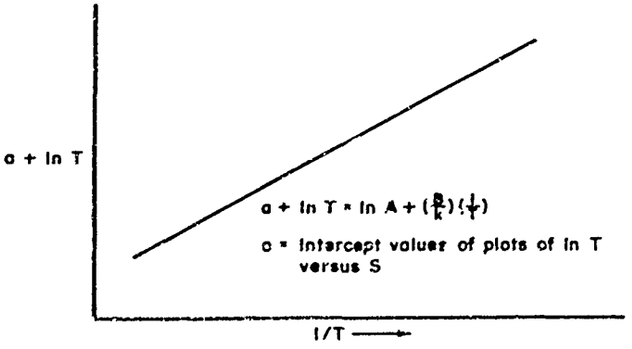


(a) Constant damage lines as a function of nonthermal stress and logarithm of time for various constant temperature levels

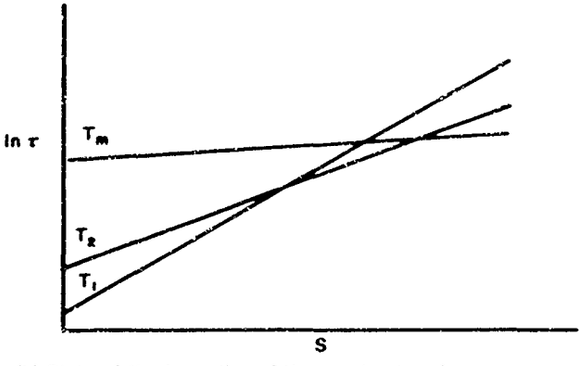


(b) Linear plot for determining the parameters C and D of the Eyring Model

FIGURE 4.17. GRAPHICAL DESCRIPTION OF THE ANALYSIS OF STEP-STRESS DATA USING THE EYRING MODEL



(a) Linear plot for determining the constants A and B of the Eyring Model



(b) Plots of the logarithm of the acceleration factor, τ, as a function of nonthermal stress for various temperature levels

A S1028

FIGURE 4.18. GRAPHICAL DESCRIPTION OF THE ANALYSIS OF STEP-STRESS DATA USING THE EYRING MODEL

Downloaded from https://www.everyspec.com

Downloaded from https://www.everyspec.com

the progressive-stress parameter, acceleration factors based on the Eyring model cannot be determined from the logarithm of time versus stress plot. Therefore, the procedure given below assumes that nonthermal stress is the progressive-stress parameter and that temperature is fixed for each progressive-stress test. The coefficients of the Eyring model dependent on temperature can be determined by replication of the progressive-stress test (using the same time rates $\beta_1, \ldots, \beta_p$) at several temperature levels.

Step 1. Obtain Linear Degradation Over Time

Obtain a transformation of the damage function, g(D), that yields a linear plot of g(D) versus time for each of the time rates, $\beta_1, \ldots, \beta_p$, of the nonthermal progressive-stress parameter. This step is identical to Step 1 based on the Arrhenius model given in Section 4.3.1.2.3 and illustrated in Figure 4.11(a). A set of the above plots should be generated for each of the constant temperature levels, $T_1, \ldots, T_m$ (Figure 4.19).

Step 2. Determine Relationship of Stress to Progressive-Stress Time

For each of the constant temperature levels, and for a specified value of the damage function, $g(D_0)$, plot applied stress, $S = \beta t$, versus the logarithm of the time at which $g(D_0)$ occurred. The equations for the m resulting plots are of the form $\ln t = a - \phi(\beta t)$ [Figure 4.17(a)]. The constants a and ϕ in each case can be determined using least-squares analysis.

Step 3. Determine the Constants C and D

Plot the values of ϕ obtained in Step 2 versus reciprocal absolute temperature [Figure 4.17(b)]. The resultant equation is given by $\phi = C + \left(\frac{D}{k}\right) t$. Hence, C is the intercept and D/k is the slope of the line.

Step 4. Determine the Constants A and B

Plot $h(a,T) = a + \ln T$ versus reciprocal absolute temperature [Figure 4.18(a)]. The equation of the resultant line is $a + \ln T = \ln A + \left(\frac{B}{k}\right) T$. Thus, the constant A is determined by the slope and the constant B by the intercept of the line.

Step 5. Calculate Acceleration Factors

Acceleration factors can be calculated for the range of temperature and nonthermal stress levels over which the relative effects of the aging processes are invariant. Thus, for given accelerated stress levels, T' and S', and for specified normal stress levels, T and S, the acceleration, τ, relative to normal operating conditions is given by Equation (4.43) [Figure 4.18(b)].

Downloaded from https://www.everyspec.com

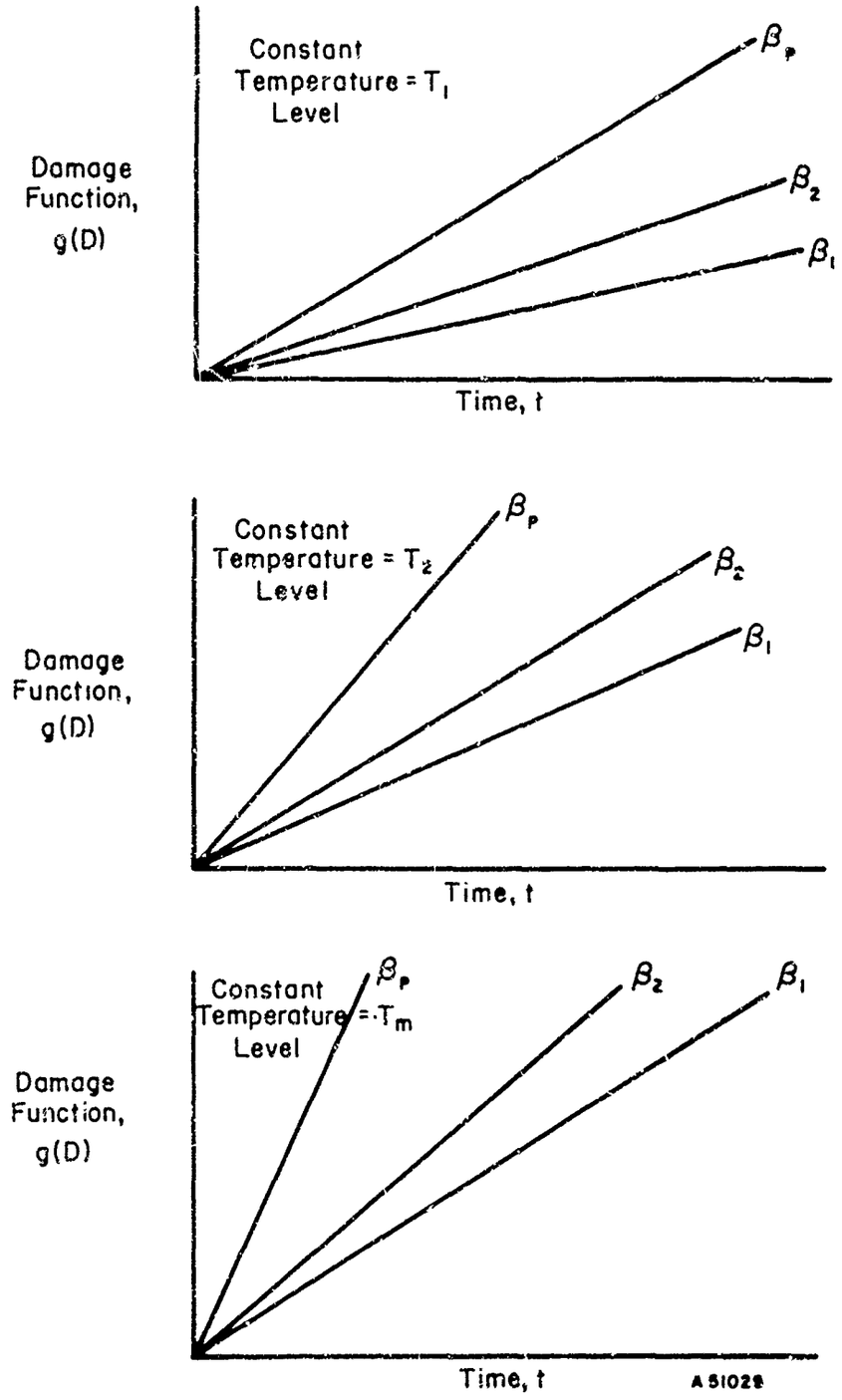


FIGURE 4.19. LINEAR CUMULATIVE DAMAGE PLOTS OVER TIME AS A FUNCTION OF THE RATE OF INCREASE OF THE PROGRESSIVE NONTHERMAL STRESS PARAMETER FOR VARIOUS CONSTANT TEMPERATURE LEVELS IN THE ANALYSIS OF PROGRESSIVE-STRESS TEST DATA USING THE EYRING MODEL

Downloaded from https://www.everyspec.com

4.4 Comparison of Accelerated Testing Procedures

In Section 4.2, three experimental methods for generating accelerated test data are described. The objective of each of these methods is to generate definitive performance data on important device parameters as functions of time and stress. The attributes of the various methods are compared in the following paragraphs. This comparison is made in terms of (1) assumptions, (2) the test time required to generate definitive data, and (3) the simplicity of analysis.

Further, in Section 4.2, the Arrhenius and Eyring models as bases for analyzing accelerated test data are compared. The comparison is made in terms of (1) the assumptions associated with each model, (2) the theoretical basis for each model, and (3) the ease of application of each model in analyzing accelerated test data.

4.4.1 Evaluation of Methods of Generating Accelerated Test Data

Constant stress, step stress, and progressive stress have the common (critical) assumption that the dominant aging mechanism does not change over the range of stress levels and the operating times involved in the tests. This assumption is most easily checked in constant-stress testing. In this case, a change in the dominant aging mechanism at some stress level will be reflected by a change in the slope of the Arrhenius or Eyring plot. However, for step-stress or progressive-stress testing, a change in the dominant aging mechanism is reflected by a change in the slopes of the various constant-damage lines at some value of the damage function on the logarithm of time versus stress plot. Since the lines of constant damage correspond to a discrete set of values of the damage function (e.g., 1 per cent, 10 per cent, 50 per cent, 90 per cent cumulative failures), additional plots will generally be required to determine the value of the damage function at which the slope of the lines changes.

Further, it is assumed that, in all three methods of testing, the rate of degradation (rate of damage accumulation) in the time increment $(t + \Delta t)$ depends only on the stress levels at that point in time and is independent of the rate or rates of degradation at any time prior to t or of the cumulative degradation up to time t. In other words, the degradation rate in the time increment $(t + \Delta t)$ is independent of the time history of the component. This assumption is, in some sense, more severe in step-stress and progressive-stress testing than in constant-stress testing. In step-stress testing, several discrete stress levels (and, therefore, degradation rates) are involved; in progressive-stress testing, stress (and, hence, degradation rate) is an increasing function of time. The net effect of this situation is that, if a stress-time interaction exists, it will be difficult to extract the interaction effect from the test data.

The test time required to generate definitive data is different for different accelerated testing procedures. In constant-stress testing, relatively long test periods are required in order to obtain significant degradation at the various stress levels in terms of the measured device parameters. On the other hand, step-stress and progressive-stress testing require substantially shorter test periods to obtain an equivalent amount of degradation. Moreover, in the latter cases, the experimenter can control to a greater degree the rates at which degradation is accumulated. In step-stress testing, this can be done by changing the size of the stress step. In progressive-stress testing, the rate at which degradation accumulates can be regulated by appropriate selection of the time rates of increases of the progressive-stress parameter.

Downloaded from https://www.everyspec.com

Also the relative simplicity of analysis of the data is different for the different accelerated testing procedures. On this basis, constant-stress data are best suited to analysis using the Arrhenius or Eyring models. This stems principally from the fact that the step-by-step procedure for analyzing constant-stress test data follows directly from the mathematical form of the analysis model. That is, for example, the Arrhenius plot required in Step 2 of the procedure is a straightforward empirical plot of the logarithm of the Arrhenius reaction-rate model; thus, the constants A and B of the model are obtained directly from this plot. However, for step-stress and progressive-stress procedures, the relation of the steps in the graphical analysis procedures are not so obviously related to the functional form of the Arrhenius or Eyring models. Thus, a greater amount of mathematical manipulation is required to determine the constants of the selected model.

4.4.2 Comparison of the Arrhenius and Eyring Models

In the introduction to Section 4.3.2, it was observed that the Eyring model is superior to the Arrhenius model in several respects: (1) the Eyring equation can be derived from fundamental physical principles of quantum mechanics in contrast to the empirical origins of the Arrhenius equation; (2) the "slopes" and "intercepts" of the Arrhenius plots are replaced in the Eyring model by the more definitive concepts of "activation energies" and "partition functions" associated with the underlying physical mechanisms of aging; (3) the theoretical basis of the Eyring model supports the validity of the graphical method of analysis. In practice, the processing of accelerated test data can be analyzed graphically by either model. If the temperature range is small, the Arrhenius equation is a good approximation to the Eyring equation.

Because of the theoretical base of the Eyring model, it is expected that each aging process in an electronic component part will be described by its own Eyring model. That is, the constants A, B, C, and D will completely characterize each aging mechanism. Experimentally determined values of these constants may prove useful in identifying the particular aging mechanisms that occur in an electronic part. The Eyring model also appears to have sufficient flexibility to deal with various failure modes. Differing modes of failure may be associated with a "potential map" which permits differing "paths" from a good state to a bad state. Alternatively, different modes of failure may be associated with the simultaneous occurrence of several aging mechanisms, each of which is characterized by its own set of Eyring constants. The failure mode that occurs is that mode which "dominates" the other possible modes of failure under the operating conditions of the device. The simultaneous occurrence of several aging mechanisms may be expected to occur in complex electronic components. It may be possible to treat aging processes as occurring sequentially in "series" or simultaneously in "parallel", and combine the associated Eyring constants in a manner analogous to that of circuit analysis.

It should also be noted that, for simple electronic parts, it may be possible to measure the Eyring constants directly by laboratory methods. If this were accomplished for a specific device, then the aging behavior of that device for its lifetime would be completely characterized by the Eyring model with the Eyring constants peculiar to the specific device. If possible, such a procedure would eliminate much of the reliance on statistical methods and mass testing, which are currently required to obtain reliability data. The estimates obtained from such reliability data do not apply to any particular part, but only to the entire "population" of "nominally identical" parts. It would clearly

Downloaded from https://www.everyspec.com

be desirable to characterize the aging behavior of each part by its own set of Eyring constants. Applications to problems of screening would be immediate. Applications would also include the improvement in reliability obtained from better design and manufacturing techniques made possible by a better understanding of the physical mechanisms of aging.

Further evidence of the basic role of the Eyring model has been shown in analyzing capacitor data. It has been shown that the empirically established "power law" for the lifetimes of capacitors can be derived from the Eyring model.(11)

Since the Arrhenius model is an excellent approximation to the Eyring model over the range of stress levels normally encountered in accelerated testing, it has substantial utility for analyzing accelerated test data. From a practical viewpoint, the significance of this close approximation is that the Arrhenius model is easier to apply in calculating acceleration factors. The quantitative nature of the Arrhenius approximation to the Eyring model can be seen from the following:

Let $\tau = 1/T$ and write the Eyring model in the form

$$R = a(k/h)(1/\tau)\, e^{-b\tau} \quad . \tag{4.63}$$

Differentiation with respect to τ yields

$$\begin{aligned} dR/d\tau &= -a(k/h)(1/\tau^2)\, e^{-b\tau} + a(k/h)(1/\tau)(-b)\, e^{-b\tau} \\ &= -\left[a(k/h)(1/\tau)\, e^{-b\tau}\right]\left[(1/\tau) + b\right] \\ &= -R\left[(1/\tau) + b\right] \quad . \end{aligned} \tag{4.64}$$

Rearranging (4.64) yields

$$dR/d\tau = -(R/\tau)\left[1 + \tau b\right] \quad . \tag{4.65}$$

It has been observed that for chemical processes, b has an order of magnitude of approximately 5×10^3 K and, in the cases where T is not greater than $t \times 10^2$ K,

$$\tau b \geq b/T = 5 \times 10^3/5 \times 10^2 \geq 10.$$

This shows that (τb) is large relative to 1. Therefore, if the factor $[1 + \tau b]$ is replaced by the approximation (τb), Equation (4.65) becomes

$$dR/d\tau = -(R/\tau)(\tau b) = -Rb \, . \tag{4.66}$$

Integration of (4.66) then yields

$$\ln (R/R_o) = -b(\tau - \tau_o) \quad . \tag{4.67}$$

Substituting $1/T$ for τ then gives

$$R = R_o e^{-b\left[(1/T) - (1/T_o)\right]} \, . \tag{4.68}$$

Downloaded from https://www.everyspec.com

This equation is identical to the Arrhenius model.

Thus, it is seen from the above analysis that the Arrhenius model is a good approximation to the Eyring model wherever

$$\tau b > > 1 \quad . \tag{4.69}$$

Since $\tau = 1/T$ and $b = E/k$, Equation (4.69) may be expressed as

$$E/(kT) > > 1 \ .$$

Thus, whenever the activation energy E exceeds kT by a factor of 10 or more, the Arrhenius model will be a good approximation to the Eyring model.

REFERENCES

(1) Sikora, G. C., and Miller, L. E., "Application of Power Stress Techniques to Transistor Life Predictions", Physics of Failure in Electronics, 3, to be published.

(2) Thomas, R. E., and Gorton, H. C., "Research Toward a Physics of Aging of Electronic Component Parts", Physics of Failure in Electronics, 2, edited by Goldberg and Vaccaro, Cato Show Printing Company, 25-60 (1964).

(3) Whitehead, S., Dielectric Breakdown of Solids, Oxford University Press, London (1951), Chapter II.

(4) "Accelerated Life Testing of Guidance Components", AL-TDR-64-235, Autonetics Division of North American Aviation (September 30, 1964).

(5) Walsh, T. M., et al., "Accelerated Testing of High-Reliability Parts", General Electric Company, Spacecraft Department, Contract AF 30(602)-3415 (October, 1964).

(6) Communication Satellite Project Advent, Final Technical Report "Long Life High Reliability Part Evaluation", General Electric Company, Missile and Space Division, Contract AF 04(647)-476.

(7) "Electronic Parts-Accelerated Life Tests", Final Technical Documentary Report, Contract AF 30(602)-2970, Motorola, Inc., Semiconductor Product Division (January, 1964).

(8) Dodson, G. A., and Howard, B. T., "High Stress Aging to Failure of Semiconductor Devices", Proceedings, 7th National Symposium on Reliability and Quality Control (January, 1961).

(9) Endicott, H. S., and Zoellner, J. A., "A Preliminary Investigation of the Steady and Progressive Stress Testing of Mica Capacitors", Proceedings, 8th National Symposium on Reliability and Quality Control (January, 1961).

Downloaded from https://www.everyspec.com

(10) Thomas, R. E., "When is a Life Test Truly Accelerated?" Electronic Design (January 6, 1964).

(11) Hatch, B. D., Endicott, H. S., and Sohmer, R. G., "Application of the Eyring Model to Capacitor Aging Data", Technical Information Series, No. R625D196, Missile and Space Division, General Electric Company.

Downloaded from https://www.everyspec.com

SECTION 5. RELIABILITY SCREENING PROCEDURES

Downloaded from https://www.everyspec.com

TABLE OF CONTENTS

	Page
5. Reliability Screening Procedures	5-1
5.1 Introduction.	5-1
5.2 Methods of Screening	5-2
5.2.1 Screening by Truncation of Distribution Tails	5-2
5.2.1.1 Step-by-Step Procedure	5-3
5.2.2 Interference Between Stress and Strength Distribution	5-4
5.2.2.1 Mathematical Development of the Screening Procedure	5-5
5.2.2.2 The Screening Problem	5-8
5.2.2.3 Step-by-Step Screening Procedure	5-10
5.2.3 Burn-In Screening	5-10
5.2.3.1 Assumptions in Burn-In Screening	5-12
5.2.3.2 The Environmental Test	5-13
5.2.3.3 The Screening Criteria.	5-14
5.2.3.4 Step-by-Step Procedure for Burn-In Screening.	5-15
5.2.4 Linear Discriminant as a Basis for Screening	5-15
5.2.4.1 Development of a Linear Discriminant	5-16
5.2.4.2 Assumptions Involved in Linear-Discriminant Analysis	5-16
5.2.4.3 The Experimental Procedure.	5-17
5.2.4.4 The Linear-Discriminant Analysis.	5-18
5.2.4.5 Application of the Linear-Discriminant Screening Criterion.	5-21
5.2.4.6 Relation of Linear-Discriminant Analysis and Physical Degradation Laws	5-22

Downloaded from https://www.everyspec.com

TABLE OF CONTENTS
(Continued)

	Page
5.3 Comparison Among Screening Procedures	5-23
5.3.1 Comparison of the Screening Procedures in Constructing a Screening Criterion	5-24
5.3.2 Comparison of the Screening Producers in Application	5-25
5.4 Survey of Precursors and Failure-Sensitive Device Parameters	5-29
5.4.1 Screening Parameters for Resistors	5-29
5.4.1.1 Carbon-Composition Resistors	5-30
5.4.1.2 Carbon-Film Resistors	5-30
5.4.1.3 Metal-Film Resistors	5-30
5.4.2 Screening Parameters for Capacitors	5-31
5.4.3 Screening Parameters for Diodes	5-33
5.4.4 Screening Parameters for Transistors	5-35
5.4.5 Summary	5-36
5.5 Future Role of Screening in Reliability Physics	5-37
5.5.1 Trends in Measured Parameters	5-37
5.5 2 Relation of Screening to Accelerated Testing	5-38
5.6 Physical Methods of Screening	5-39
5.6.1 Resistors	5-40
5.6.1.1 Current or 1/f Noise	5-40
5.6.1.2 Broadband Noise	5-41
5.6.1.3 Parameter Drift	5-41
5.6.1.4 Infrared Emission	5-41

Downloaded from https://www.everyspec.com

TABLE OF CONTENTS
(Continued)

	Page
5.6.2 Capacitors	5-43
5.6.2.1 Anodic and Ionic Reactions	5-43
5.6.2.2 Hot D-C Leakage	5-43
5.6.2.3 Parameter Drift	5-44
5.6.3 Diodes	5-44
5.6.3.1 Reverse Current at Specified Voltage	5-44
5.6.3.2 Reverse Breakdown Voltage	5-44
5.6.3.3 Open-Circuit Voltage Decay Time	5-45
5.6.3.4 Reverse Pulse Recovery Time	5-45
5.6.3.5 Slope of Log Reverse Current Versus Reciprocal Temperature Curve	5-46
5.6.3.6 Avalanche Noise	5-46
5.6.3.7 Electron-Beam Scan	5-47
5.6.3.8 Infrared	5-47
5.6.4 Transistors	5-48
5.6.4.1 Collector-Base Leakage Current, I_{CBO}	5-48
5.6.4.2 Collector-Base Breakdown Voltage, BV_{CBO}	5-48
5.6.4.3 Direct-Current Gain, h_{FE}	5-48
5.6.4.4 Collector-Emitter Saturation Voltage, $V_{CE(sat)}$	5-48
5.6.4.5 Transistor Noise Measurement	5-49
5.6.4.6 Infrared Scanning	5-49
5.6.4.7 Electron-Beam Scanning	5-49
REFERENCES	5-49

Downloaded from https://www.everyspec.com

LIST OF FIGURES

Page

Figure 5.1. Truncation of Tails of Distribution of Quality Parameter Values for a Population of Devices 5-2

Figure 5.2. Device Reliability at Time T Resulting From Screening Out Devices Outside Marginal Tolerance Limits at Time t = 0 5-3

Figure 5.3. Schematic Illustrating Reliability as a Function of Interference Between Stress and Strength Distributions 5-6

Figure 5.4. Component Reliability as a Function of Distance, $\bar{u}/\sigma_u$, Between the Mean Values of the Stress and Strength Distributions E(s) and C(s). 5-8

Figure 5.5. Incremental Gains in Reliability for a Given Incremental Increase in s* as a Function of s*. 5-11

Figure 5.6. Assumed Distribution of Component Strengths Before Screening. 5-12

Figure 5.7. Schematic Representation of Assumed Strength Distributions in Terms of Failure Rate 5-12

Figure 5.8. Schematic Representation of the Theoretical Distribution of z-Values for the Superior and Inferior Components 5-22

Figure 5.9. Time Response of Collector-Base Reverse Leakage Current. . . 5-36

Figure 5.10. Rate of Change in Early Life of Surface Channel Current as a Predictor of Failure 5-38

Figure 5.11. Circuit for Measuring Resistor Noise 5-40

LIST OF TABLES

Table 5.1. Values of Critical Value, s*, Such That R* = 0.95, for Various Values of R 5-11

Table 5.2. Stress Levels and Test Times Commonly Used in Burn-In Test Design 5-14

Table 5.3. Expected Conditional Probabilities of Classification Given $z^* = (\bar{z}_2 + \bar{z}_1)/2$ 5-20

Table 5.4. Expected Conditional Probabilities of Classification . Given $z^* = \alpha\bar{z}_1 + (1-\alpha)\bar{z}_2$ 5-21

Table 5.5. Comparison of Additional Information Included in Successively More Complex Screening Procedures 5-26

Downloaded from https://www.everyspec.com

LIST OF TABLES
(Continued)

		Page
Table 5.6.	Summary Comparison of Attributes in the Construction and Application of the Various Screening Procedures	5-27
Table 5.7.	Summary of Comparative Properties of Various Screening Procedures	5-28
Table 5.8.	Screening Parameters for Resistors	5-29
Table 5.9.	Screening Parameters for Diodes	5-34
Table 5.10.	Screening Parameters for Transistors	5-35

Downloaded from https://www.everyspec.com

5. Reliability Screening Procedures

5.1 Introduction

The purpose of reliability screening is to select from a set of devices those devices having superior reliability or, alternatively, to reject those devices having inferior reliability. Of particular interest is the screening out of component parts that are potential early failures in some specified application. The screening procedures used involve the classification of each device in the set on the basis of initial or early life parameter measurements. Depending on device reliability requirements and the "power" of the screening procedure, devices may be screened into two or more classifications. For example, in screening sets of component parts, one may wish to (1) eliminate the potential early failures, (2) select those parts exhibiting high reliability for application in aerospace systems, and (3) select those parts of intermediate reliability for commercial applications.

It should be noted that reliability screening differs from quality control in several respects. First, it is not the purpose of reliability screening to detect devices which are defective at the time of measurement. In fact, it is assumed all devices are initially good. Second, unlike acceptance sampling, parts qualification, etc., reliability screening is a 100 per cent inspection procedure. Third, classification by a screening procedure is accomplished with respect to life-time requirements and operating conditions involved in the intended application of the device. Thus, it is seen that the concept involved in reliability screening, as opposed to quality control, is not to reject a device strictly because one or more initial parameter measurements lie in some unacceptable region, but to consider such measurements as precursors of early failure (unreliability).

Ideally, a screening procedure would never misclassify a device. That is, a high-reliability device would never be classified as a potential early failure, and conversely, an unreliable device would never be classified as having a potentially long life. In practice, such ideal screening procedures do not exist. The objective of the various procedures described in the following sections is to approach the ideal state as closely as possible. In the case of one procedure, the analysis involves calculation of the probabilities of misclassification. Thus, one can judge the utility of the procedure on the basis of these probability measures. The remaining simpler procedures do not explicitly involve evaluation of utility; however, theoretically the analyses could be extended to include such evaluation.

The screening procedures presented in the following sections are given in order of relative complexity, starting with the least complex. Thus, in order, the screening procedures are:

(1) Screening by truncation of distribution tails

(2) Interference between stress and strength distributions

(3) Burn-in screening

(4) Linear-discriminant screening.

Downloaded from https://www.everyspec.com

The description of each procedure includes a discussion of applications, statement of the assumptions involved, a step-by-step accounting of the procedure, and an illustrative example. In subsequent sections the various procedures are compared, and their future roles in reliability physics are discussed.

5.2 Methods of Screening

The four screening procedures listed above are paired into two groups: (1) screening to obtain quality performance at t = 0 and (2) screening to obtain long life. The essential difference between the two is that the first group involves screening on the basis of initial measurements, whereas the second group involves, in addition to initial measurements, screening on the basis of early life measurements. These may take the form of Δ increments, rates of change in given parameters, etc. Procedures for the second type of screening generally are more powerful, but also are more costly.

5.2.1 Screening by Truncation of Distribution Tails

The primary purpose of screening by truncation is to attain device homogeneity in terms of important device parameters with respect to a design specification. Thus, the screening problem in this case consists in specifying those device parameters to be controlled and in establishing tolerance limits defining the acceptable range of measured values for each parameter. Screening is then accomplished by eliminating those devices from a set having values outside the given tolerance limits. In concept, if the initial parameter values for a population of a given device are described by a probability distribution function, it is seen that this procedure consists in setting tolerance limits on the distribution function such that the distribution tails are truncated (Figure 5.1).

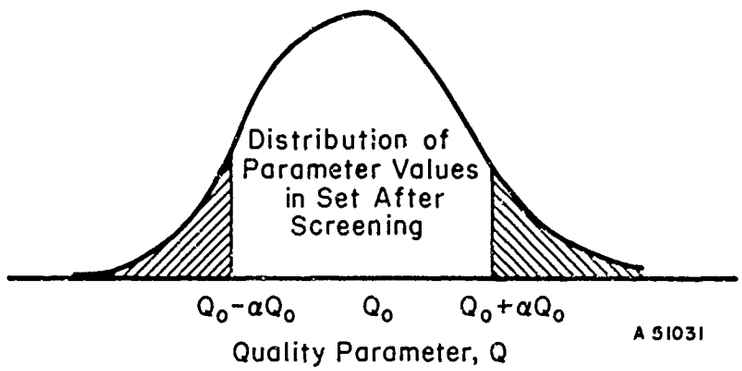


FIGURE 5.1. TRUNCATION OF THE TAILS OF DISTRIBUTION OF QUALITY-PARAMETER VALUES FOR A POPULATION OF DEVICES

As shown in Figure 5.1, the tolerance limits are characteristically expressed in terms of some acceptable per cent deviation from nominal value of the quality parameter. For example, the design specification for a given equipment may specify that all resistors of a given type will have initial resistance measurements within 0.5 per cent of nominal. Thus, the tolerance limits for a 10-K resistor would be T_L = 10 K-(0.005) (10 K) = 9.95 K and, correspondingly, T_U = 10.05 K. For devices having several quality parameters,

Downloaded from https://www.everyspec.com

this procedure can be repeated for each parameter, yielding a set after screening that consists of those devices in the original set having all initial parameter measurements within their respective tolerance limits.

From the standpoint of reliability, the concept involved in this type of screening is characteristic of the marginal tolerance limit problem. That is, it is tacitly assumed that devices whose initial parameter measurements are within marginal tolerance limits (T_L, T_U) are not likely to fail in a specified application. Failure in this case is said to occur if a quality parameter exceeds failure tolerance limits (F_L, F_U). Figure 5.2 shows the distribution of parameter values at t = 0. If those devices having parameter values outside the marginal tolerance limits are eliminated, the resulting truncated distribution will take on the form shown at time T, where the proportion of values inside the failure tolerance limits represents device reliability attained due to the screening procedure.

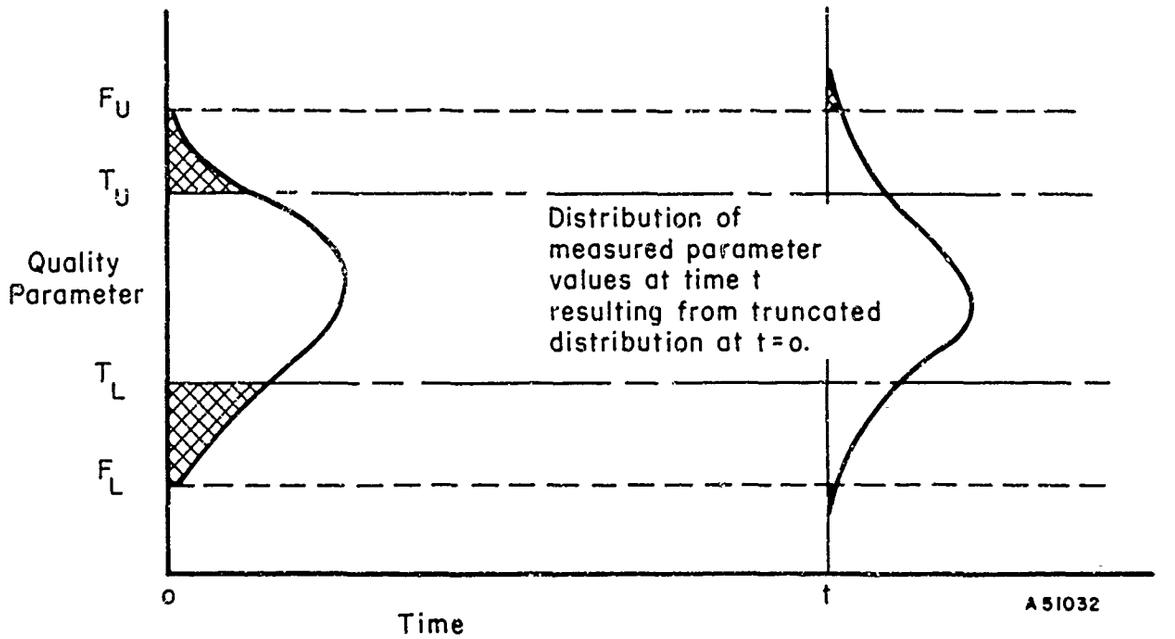


FIGURE 5.2. DEVICE RELIABILITY AT TIME T RESULTING FROM SCREENING OUT DEVICES OUTSIDE MARGINAL TOLERANCE LIMITS AT TIME t = 0

It should be recognized that the device itself may not fail per se, but device output values exceeding F_L or F_U may induce failure in some other component in the system. The importance of this point is that device tolerance limits must be established with respect to the input-output functions of the device in its intended application. Hence, it is seen that appropriate tolerance limits should be established on the basis of an analysis of the dynamics of the system of which the devices being screened are a part.

5.2.1.1 Step-by-Step Procedure

From the foregoing discussion it is seen that screening by truncation of distribution tails consists of four sequential steps:

Downloaded from https://www.everyspec.com

(1) Define the set of quality parameters on the basis of which the set of devices is to be screened

(2) Determine tolerance limits for each quality parameter, based on required device input-output functions for system operation

(3) Obtain initial parameter measurements for each device in the set

(4) Screen out those devices having one or more initial parameter values outside specified tolerance limits.

Thus, screening by truncation of distribution tails has advantages in that it is simple to apply and depends only on parameter measurements at t = 0. On the other hand, the procedure has disadvantages in that device reliability cannot be predicted as a direct result of the screening procedure, and appropriate parameter tolerance limits require an analysis of the system dynamics, which in general may be very complex.

Example. Consider the problem of selecting reference diodes for application in power-supply units. Circuit analysis shows that the output voltage of the power supply (the system) is very sensitive to the stability of the reference diode.

Step (1) The diode parameters that must be controlled to attain stability are reference voltage and zener impedance.

Step (2) Tolerance limits for these parameters are determined at a specified current level according to the functional relation of diode stability and power-supply output. For reference voltage, a two-sided tolerance limit is determined of the form $v_o \pm \Delta v$, and for zener impedance a one-sided limit is determined of the form $R > R_o$ (zener impedance essentially represents the slope of the i-v curve at the reference-current level. The greater the slope, the less sensitive reference voltage is to fluctuations in current.

Step (3) Given these tolerance limits, initial measurements are generated for diodes to be used in the power-supply unit.

Step (4) The diodes are then screened, based on the initial measurements and parameter tolerance limits.

5.2.2 Interference Between Stress and Strength Distribution

The purpose of interference reliability screening is to eliminate devices having strength measurements that may be expected to be exceeded by stress levels in their intended application. The essential difference between this procedure and that given in the preceding section is that here, in addition to obtaining a distribution of device capabilities, it is necessary to generate a distribution of environmental stresses.

Downloaded from https://www.everyspec.com

Reliability is then given by the probability of stress exceeding device strength. Therefore, if those devices having initially low strength measurements are screened out, reliability will presumably be increased.

There are two significant problems that arise in actual application of stress-strength analysis:

(1) In general, the distribution of environmental stresses is difficult to obtain.

(2) The strength distribution will, in general, change as a function of time.

The environmental stresses generally will be combinations of many stress components such as temperature, vibration, and pressure. Moreover, the effect of the stress components on a device will not be additive, but will depend on a function of the interactions of the stress components. Also, the strength distribution will change with time. The rate and type of change will be a function of the environmental stresses actually operating on the device.[1]

The screening problem, in view of the above remarks, is, therefore, to determine an initial value of device strength, s*, such that all devices having measured values less than s* will be eliminated. The choice of the screening criterion value, s*, will depend on the expected environmental stresses in application and on the expected change of the strength distribution as a function of stress.

5.2.2.1 Mathematical Development of the Screening Procedure

Consider the distribution of environmental stress and device strength shown in Figure 5.3.[1] The probability that a device has strength, s_o, or greater is seen from Figure 5.3b to be:

$$P\left\{s > s_o\right\} = \int_{s_o}^{\infty} C(s)\, ds \quad , \tag{5.1}$$

where C(s) = probability density function of component strength.

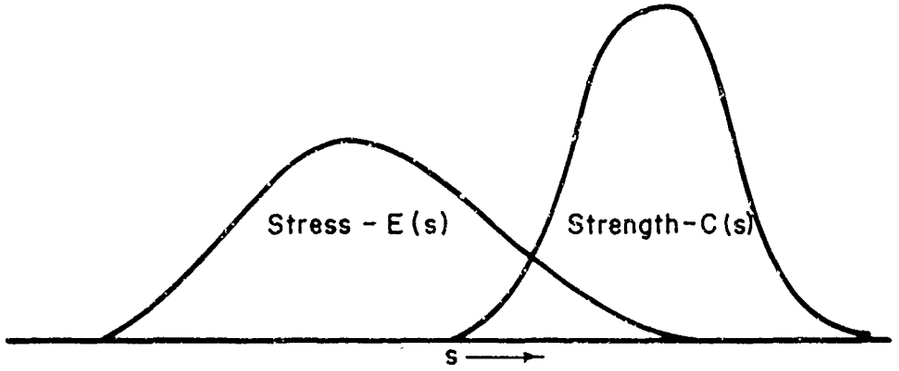
The probability of a stress value occurring in an infinitesimal interval, ds, about the point s_o is given by

$$P\left\{\left(s_o - \frac{ds}{2}\right) \leq s \leq \left(s_o + \frac{ds}{2}\right)\right\} = E(s)\, ds \quad , \tag{5.2}$$

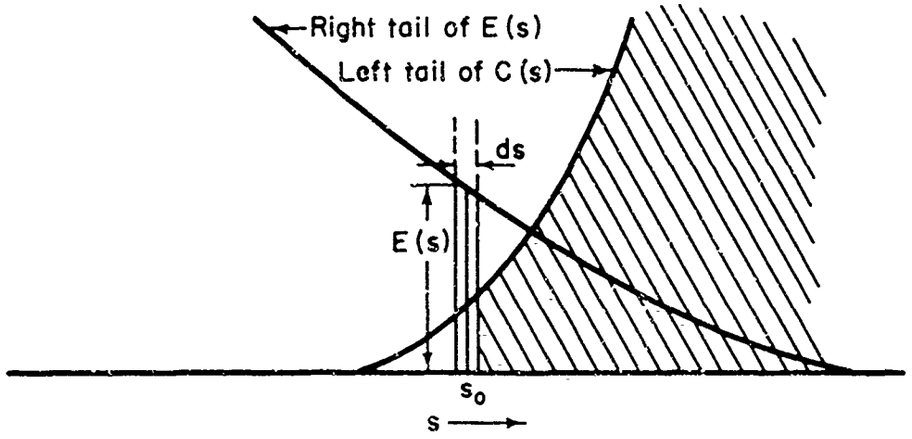
where E(s) = probability density function of environmental stress. Therefore, the conditional reliability given a stress in the interval $\left[s_o \pm \frac{ds}{2}\right]$ is the product of Equations (5.1) and (5.2):

$$dR = \left[\int_{s_o}^{\infty} C(s)\, ds\right] E(s)\, ds \quad . \tag{5.3}$$

Downloaded from https://www.everyspec.com



(a) Interference of stress and strength distributions, E(s) and C(s)



(b) Graphical description of the calculation of reliability as a function of the interference of E(s) and C(s)

A51030

FIGURE 5.3. SCHEMATIC ILLUSTRATING RELIABILITY AS A FUNCTION OF INTERFERENCE BETWEEN STRESS AND STRENGTH DISTRIBUTIONS

Downloaded from https://www.everyspec.com

Component reliability is obtained from Equation (5.3) by integrating (5.3) over all possible values of s,

$$R = \int_o^\infty \left[\int_s^\infty C(s)\, ds \right] E(s)\, ds \quad . \tag{5.4}$$

Equivalently, reliability can be determined by considering the probability of a stress less than or equal to s_o and a component strength in the interval $\left[s_o \pm \frac{ds}{2} \right]$. This leads to the equivalent equation

$$R = \int_o^\infty \left[\int_o^s E(s)\, ds \right] C(s)\, ds \quad . \tag{5.5}$$

Analytic solution of the above equations for reliability is not generally possible. A graphical procedure for evaluating R for arbitrary density functions C(s) and E(s) is given by Kececioglu and Cormier[2]. If C(s) and E(s) are normal density functions, R can be evaluated by considering the variable $u = (s_2 - s_1)$, where s_2 denotes component strength and s_1 denotes environmental stress. Clearly, no failure will occur if $u = s_2 - s_1 > 0$. Since s_2 and s_1 are normally distributed, $u = s_2 - s_1$ is also normally distributed with mean

$$\mu(u) = \mu(s_2) - \mu(s_1) \tag{5.6}$$

and variance

$$\sigma_u^2 = \sigma_{s_2}^2 + \sigma_{s_1}^2 \quad . \tag{5.7}$$

Reliability is, therefore, given by

$$R = P\left\{ u > o \right\} = \frac{1}{\sqrt{2\pi}\,\sigma_u} \int_o^\infty \exp\left\{ -\frac{1}{2} \left[\frac{u - \bar{u}}{\sigma_u} \right]^2 \right\} du \tag{5.8}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\frac{\bar{u}}{\sigma_u}}^\infty \exp\left[-z^2/2 \right] dz \quad , \tag{5.9}$$

where $z = (u - \bar{u})/\sigma_u$. The integral of Equation (5.9) can be evaluated readily by use of the table of the standard normal distribution. In Figure 5.4, reliability is plotted as a function $\bar{u}/\sigma_u$, and the separation between the means of C(s) and E(s) is measured in units of the standard deviation, σ_u. For example, if $\bar{u}/\sigma_u = 1.25$, component reliability is seen to be R = 0.90. Alternatively, if an R = 0.99 is required, a separation of $\bar{u}/\sigma_u = 2.28$ must exist between the means of C(s) and E(s).

Downloaded from https://www.everyspec.com

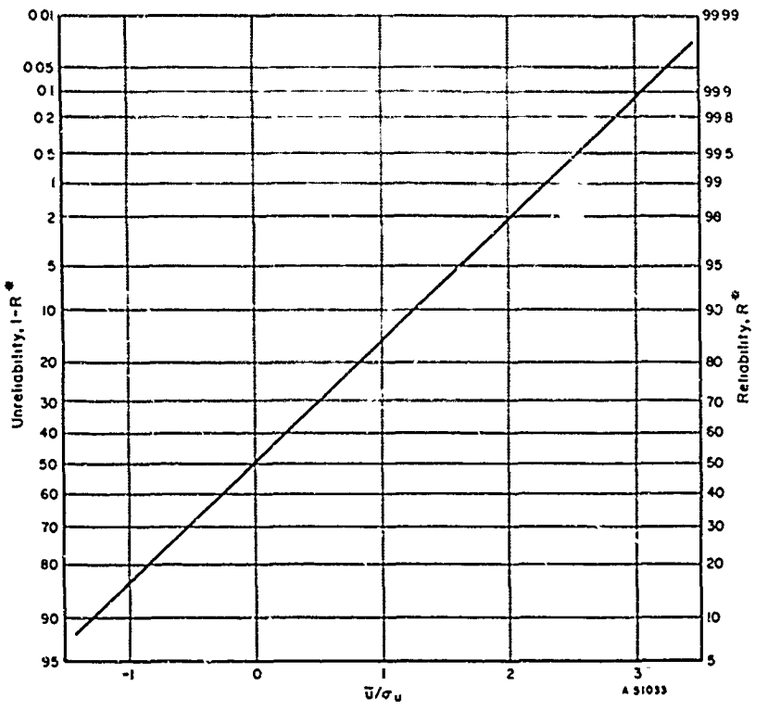


FIGURE 5.4. COMPONENT RELIABILITY AS A FUNCTION OF THE DISTANCE, $\bar{u}/\sigma_u$, BETWEEN THE MEAN VALUES OF THE STRESS AND STRENGTH DISTRIBUTIONS E(s) AND C(s)

5.2.2.2 The Screening Problem

It is clear from Figure 5.4, and the foregoing discussion, that reliability increases as the area of the overlap region of C(s) and E(s) decreases. For given distribution functions, C(s) and E(s), reliability can be increased by screening out those components having strength measurements in some part of the overlap region. The screening problem is, then, to choose a critical value of strength, s*, such that all components having strength measurements less than s* are eliminated. In this sense, the type of screening is similar to that described in Section 5.2.1 in that the procedure consists of truncating the left tail of the C(s).

For given C(s) and E(s), s* should be chosen such that reliability due to screening will attain a specified requirement, R*. If we substitute the truncated distribution of strength values, $C_T(s)$, for C(s) in Equation (5.4) we obtain

$$R^* = \int_{-\infty}^{\infty} \left(\int_{s}^{\infty} C_T(x)\, dx \right) E(s)\, ds \quad , \tag{5.10}$$

where

$$C_T(s) = \begin{cases} C(s) \Big/ \displaystyle\int_{s^*}^{\infty} C(s)\, ds \quad , & s > s^* \\ 0 \quad , & s < s^* \end{cases} . \tag{5.11}$$

Downloaded from https://www.everyspec.com

Since $\int_{s^*}^{\infty} C(s)ds$ for given s^* is a constant with respect to the variable of integration, Equation (5.10) can be written as

$$R^* = \int_{-\infty}^{s^*} E(s)\,ds + K(s^*) \int_{s^*}^{\infty} \left(\int_{s}^{\infty} C(x)\,dx \right) E(s)\,ds \quad , \tag{5.12}$$

where

$$K(s^*) = \left[\int_{s^*}^{\infty} C(s)\,ds \right]^{-1} \quad . \tag{5.13}$$

Equation (5.12) cannot be solved analytically for the value of s^* that will yield a specified reliability, R^*; hence, numerical approximation methods must be used. However, if s^* is sufficiently far out in the left tail of $C(s)$, a close approximation to (5.12) is

$$R^* = K(s^*) \left[\int_{-\infty}^{\infty} \left(\int_{s}^{\infty} C(x)\,dx \right) E(s)\,ds \right] \quad . \tag{5.14}$$

The quantity inside the brackets in Equation (5.14) is seen to be identical to R [Equation (5.4)]. Thus, we obtain

$$R^* = K(s^*)\,R \quad . \tag{5.15}$$

That is, the reliability obtained by screening based on a given s^* is equal to the reliability before screening multiplied by a constant. This approximation will yield a reasonably small overestimate of reliability for values of s^* in the lower 5 per cent tail of the distribution of $C(s)$.

Based on the above, a suitable procedure for determining s^* such that a specified reliability after screening, R^*, is obtained is:

(1) Calculate the required s^*, using the approximate function [Equation (5.15)].

(2) Determine whether the s^* calculated in Step (1) is sufficiently far out in the left tail of $C(s)$; say $s^* \leq s_\alpha$, where s_α is the lower α per cent point of $C(s)$.

(3) If $s^* \leq s_\alpha$, it can be used as the required critical value for screening. If $s^* > s_\alpha$, then a new value of s^* must be determined numerically that satisfies the exact function [Equation (5.12)].

For example, suppose that $C(s)$ and $E(s)$ are normal distributions and that a reliability after screening of $R^* = 0.95$ is required. From Equation (5.15) it is seen that $K(s^*) = R^*/R$, where R depends on $\bar{u}/\sigma_u$ as shown in Equation (5.9) and Figure 5.4. Values of s^* satisfying the requirement $R^* = 0.95$ are given in Table 5.1 for various values of R.

Downloaded from https://www.everyspec.com

Table 5.1 shows as expected that, as the distance between the stress and strength distributions increases, the size of the tail of the strength distribution that is truncated decreases for a given R*. With respect to the acceptability of the approximate values of s* in Table 5.1, it is seen that for R = 0.90 and 0.95 that s* is in the lower 5 per cent tail of C(s) and, therefore, will yield a reasonably good approximation. On the other hand, the s* values for R < 0.90 will yield a significant overestimate of R*. Thus, for the latter cases it is necessary to resort to numerical determination of s* satisfying Equation (5.12).

To make clear the conceptual problem of choosing s*, consider Figure 5.5. As s* is made continuously greater [i.e., s* moves further out in the left tail of C(s)], the area of the overlap region of $C_T(s)$ and E(s) will decrease, and therefore reliability will increase. However, the rate of change of R with increasing s* will decrease. Moreover, as s* increases, the number of components eliminated by screening will increase at an increasing rate. If we assign some cost to each component that is screened out, it is reasonable to assume that at some point the marginal increase in reliability is not worth the cost incurred in attaining it. The important point is that s* should be chosen with respect to both reliability requirements and the cost of attaining the required reliability.

5.2.2.3 Step-by-Step Screening Procedure

(1) Determine the distribution of strength measurements, C(s), for the set of components to be screened

(2) Define the environmental stress variables acting on a component in its intended application

(3) Determine an equation for the combining of stress variables into a single measure, s

(4) Determine the distribution of the environmental stress measure, E(s)

(5) Determine the critical value s* on the basis of a specified reliability requirement, R*, from Equations (5.12) or (5.14), depending on whether the exact or approximate solution is used

(6) Eliminate those components having strength measurements less than s*.

5.2.3 Burn-In Screening

The purpose of burn-in screening is to eliminate components having inferior reliability on the basis of a short-term environmental test. The environmental test may be run under high stress conditions or under normal stress conditions with respect to some intended application. The components that are eliminated as a result of the test may actually fail during the test or may exhibit performance characteristics that are indicative of early failure in actual application. This type of screening can be used to sort out a subset of "substandard" components from a set (population) of standard (reliable) components or, alternatively, to reduce the average hazard rate of the remaining population after screening in those cases where devices have decreasing hazard rates with time.

Downloaded from https://www.everyspec.com

TABLE 5.1. VALUES OF THE CRITICAL VALUE, s*, SUCH THAT R* = 0.95 FOR VARIOUS VALUES OF R

s* is given in units of the standard deviation, σ_s, of strength distribution, C(s) .

R	$\bar{u}/\sigma_u$	K(s*) = R*/R	1/K(s*)	s*
0.86	1.05	1.10	0.906	-1.32
0.88	1.15	1.08	0.927	-1.45
0.90	1.26	1.05	0.948	-1.63
0.92	1.37	1.03	0.969	-1.88

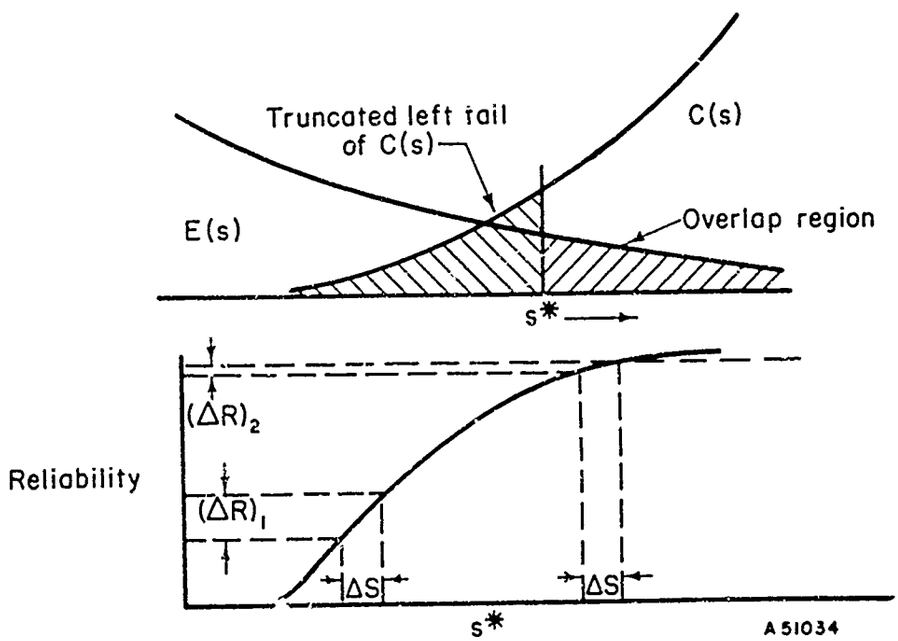


FIGURE 5.5. INCREMENTAL GAINS IN RELIABILITY FOR A GIVEN INCREMENTAL INCREASE IN s* AS A FUNCTION OF s*

Downloaded from https://www.everyspec.com

The development of a burn-in screening procedure in a particular case involves consideration of (1) the assumptions associated with the procedure, (2) the test design, and (3) the screening criterion. The essential aspects of these factors are discussed in the following paragraphs.

5.2.3.1 Assumptions in Burn-In Screening

The underlying assumption in burn-in screening is that a set of components before screening consists of two subsets: (1) a subset of superior components represented by a distribution of strength measurements with a high mean strength and small variance and (2) a subset of inferior components represented by a distribution of relatively low strength measurements.

In terms of failure rate, this assumption, illustrated in Figure 5.6, implies a failure-rate curve of the familiar form shown in Figure 5.7. It is seen that the inferior components fail early, yielding an initially high failure rate. After this initial period (infant mortality), the failure-rate curve levels off until the superior component strengths degrade to the point of wear-out.

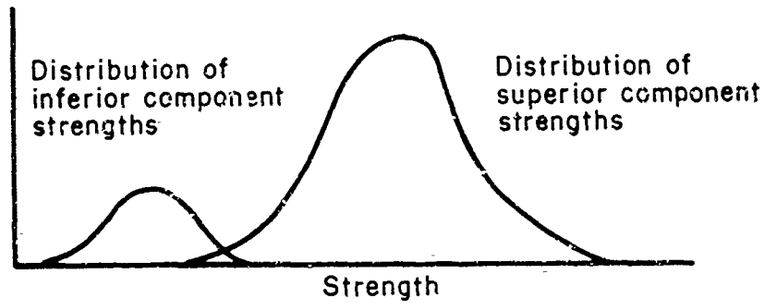


FIGURE 5.6. ASSUMED DISTRIBUTION OF COMPONENT STRENGTHS BEFORE SCREENING

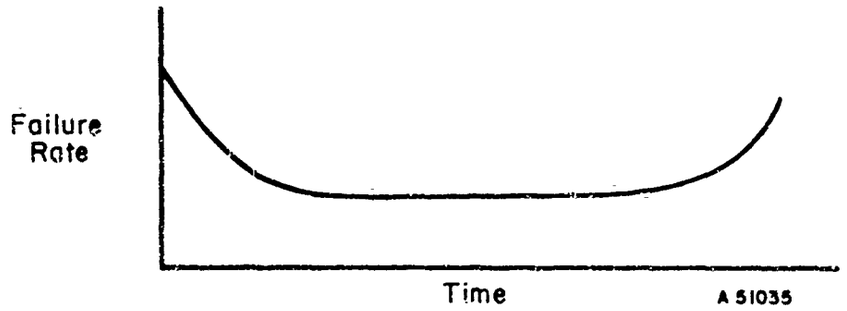


FIGURE 5.7. SCHEMATIC REPRESENTATION OF ASSUMED STRENGTH DISTRIBUTIONS IN TERMS OF FAILURE RATE

Downloaded from https://www.everyspec.com

A second assumption concerns the functional relation of parameters on which screening is based to the failure mechanisms acting in a component. If every failure mechanism and its physical properties for a given component were known, it would be possible to determine what parameter measurements act as precursors of component failure. However, all the failure mechanisms are not generally known or completely describable. Therefore, we must assume (to some degree) which parameters should be measured as precursors of failure.

A third assumption, closely related to the one above, concerns what characteristics of parameter variation are indicative of early failure. These characteristics may include parameter instability, parameter drift in some time increment, or rate of parameter drift. The degree of assumption in this case will depend on the knowledge of the degradation processes acting in a given component type.

It should be noted that the first assumption is statistical in nature, whereas the second and third assumptions are physical in nature. The significant point in this respect is that as our knowledge of the physics of electronic devices increases, the degree of uncertainty in the physical assumptions decreases. This implies that, if we gain complete knowledge of the physical degradation processes acting in a component, it will be possible to predict early failures with near certainty on the basis of short-term burn-in tests.

5.2.3.2 The Environmental Test

The test design for burn-in screening requires specification of

(1) The parameters to be measured

(2) The stress levels for each of the stress factors

(3) The time duration of the test.

The parameters to be measured in a particular instance depend on the degradation processes acting in the component and how these processes relate to the total set of measurable parameters. It will usually be the case that the parameters will be specified on the basis of subjective experience as discussed in the preceding section on assumptions. Once the parameters have been specified, the next step is to determine what measure of the parameter should be used as a basis for screening. As discussed previously, these may include some measures as instability, incremental change over some time period, and rate of change.

The second factor to be specified in the test design is the stress levels under which the test should be run. The governing principle here is to operate the components under a stress environment for a specified time such that the inferior components will fail or exhibit undesirable parameter variation while the superior parts will not prematurely age. Thus, the problem is to select stress levels that are sufficiently severe to yield significant parameter degradation in the inferior components without introducing failure mechanisms not existing under normal stress conditions. This problem is discussed in detail in Section 5.4.

Downloaded from https://www.everyspec.com

The third factor in the test design is the time duration of the test. Ideally, the test time should be that time required to yield statistically significant parameter degradation at the highest possible stress levels giving "true" acceleration. In practice, however, the appropriate trade-off between test duration and stress level must be determined largely from experience of previous test analyses. Again, as knowledge of the reliability physics increases, greater precision can be introduced into the screening-test design. Table 5.2 lists stress levels and test times characteristically used for various component types in burn-in tests. It should be noted that the stress conditions given in this table are intended to be illustrative of stresses that may actually be applied. Alternative stress schemes may be applicable, such as increasing stress testing, short-term overstress, etc.

TABLE 5.2. STRESS LEVELS AND TEST TIMES COMMONLY USED IN BURN-IN TEST DESIGN

Component	Test Time, hours	Stress Levels
Transistors	125-250	Near rated power dissipation Temperature approximately 90% of maximum operating temperature
Diodes	500	Near rated load Temperature approximately at maximum operating temperature
Resistors	750	Electrical stress approximately 30% of rated load Temperatures in excess of rated temperature
Capacitors		Electrical stress at rated voltage Temperature at rated value

5.2.3.3 The Screening Criteria

The final step in developing a burn-in screening program is to determine screening criteria. These criteria generally are in the form of tolerance limits on the measured parameters. The concept in determining these limits is the same as that used in Section 5.2.1, "Screening by Truncation of Distribution Tails". That is, the limits can be considered as marginal tolerance limits in that, if a component parameter is outside of marginal tolerance limits after the burn-in test, the probability is high the parameter will be outside of failure-tolerance limits at some early time of actual operation. These relations can be estimated from long-term component life tests. In Section 5.4, those parameters that have been successfully used in component screening are identified in conjunction with the type of tolerance limits employed and the test conditions.

Downloaded from https://www.everyspec.com

5.2.3.4 Step-by-Step Procedure for Burn-In Screening

(1) Identify the failure mechanisms for the component type to be screened

(2) Determine the measurable parameters that are indicative of the failure mechanisms

(3) Specify the environmental stress levels and time duration for the burn-in test

(4) Specify the desired measure of the indicative parameters, environmental change, rate of change, etc.

(5) Establish the screening criteria: tolerance limits on the measured parameters

(6) Run the test and screen out those components having parameter measures outside of tolerance limits.

5.2.4 Linear Discriminant as a Basis for Screening

Linear discriminant as a basis for screening components may be considered as a logical extension of the techniques of the screening procedures given in the preceding sections. Its objective is to predict early failures on the basis of a weighted average of initial and/or early life parameter measurements. This weighted average is called a linear discriminant. The problems in constructing a linear discriminant are to determine which parameters are significant in predicting failure, to determine the "weight" that should be assigned to these parameters, and to detdrmine what limits on the value of the linear discriminant should be used in order to reject the inferior (or select the superior) components.

In order to construct the linear discriminant, it is necessary to conduct a load-life test on a statistically significant number of the components that are to be screened. The results of this life test are used to define failure (if this has not been previously dictated by the application), to select the parameters that will be used in the linear discriminant, and to evaluate the probability of success of the discriminant in screening subsequent lo's of the component parts. Thus, in contrast to the procedures previously discussed, it is seen here that the parameters on which screening is based and the predictive "power" of the screening function are determined as explicit steps in the analysis.

The screening is carried out by measuring a set of p different parameters selected in the discriminant analysis for each component. These measured values, $x_1, \ldots x_p$, are then substituted into an equation of the form $z = \lambda_1 x_1 + \ldots + \lambda_p x_p$, where the λ's denote optimal numerical "weights" associated with the measured values. This equation is called a linear discriminant. If the weighted average, or z value, is less than a critical value of z, then the component is judged to be a potential early failure. Otherwise, the component is judged to be a potential late failure, or in other words, to have a long life.

Downloaded from https://www.everyspec.com

An ideal screening criterion would never misclassify a component. That is, a superior component would never be classified as inferior, and conversely an inferior component would never be classified as superior. In practice, such ideal screening criteria do not exist. However, the use of linear-discriminant analysis permits the construction of screening criteria for which the probability of misclassification can be computed when certain assumptions are made. Consequently, if the probabilities of correct classification are sufficiently high, the associated screening criterion has practical significance. On the other hand, if the probabilities of correct classification are small, then the screening criterion is of little practical significance, even though the criterion may show statistical significance.

The principal advantage of using the method of linear-discriminant analysis to construct the screening criterion is that no other linear combination of parameter measurements will yield a screening criterion with smaller probabilities of misclassification. Thus, if the linear-discriminant method fails to produce a screening criterion of practical significance, the search for a linear screening criterion may be abandoned. This conclusion must be qualified to the extent that certain assumptions must be made and that various transformations of the data may also be required. If these fail, still other modifications can be attempted in an effort to improve the screening classification.

5.2.4.1 Development of a Linear Discriminant

The development of a linear-discriminant screening criterion requires several steps:

(1) A load-life test must be performed on a statistically significant number of components. The components selected for the test must be representative of the population of components for which the screening criterion is being derived. The length of the test is that time for which reliability prediction is desired.

(2) The linear-discriminant screening criterion is then to be derived from initial, early-life, and end-of-test parameter measurements of the components on test.

(3) The resultant screening criterion should be empirically verified prior to screening components of unknown reliability.

These steps are described in detail in the following paragraphs along with assumptions associated with this screening procedure.

5.2.4.2 Assumptions Involved in Linear-Discriminant Analysis

Linear-discriminant analysis is essentially a form of linear regression analysis. As such, it involves the same assumptions that are pertinent to statistical regression analysis. In particular, linear-discriminant analysis requires that

Downloaded from https://www.everyspec.com

(1) The parameter values used in the linear-discriminant function be normally distributed.

(2) The within-sets variances be homogeneous. That is, the variance of the parameter measurements of the superior components must be the same as that for the inferior components.

Where these assumptions are found not to hold, various transformations may be tried in order to satisfy the assumptions. For example, it is frequently the case that several of the measured parameters are decidedly nonnormally distributed. It has been noted in many of these instances that a logarithmic transformation of the parameter measurements yields satisfactory agreement with the normality assumption.

5.2.4.3 The Experimental Procedure

The first step in deriving a linear-discriminant screening criterion consists of a load-life test on a sample of components drawn from the population of the component type of interest. This experiment consists of the following steps:

(1) Define component failure. Failure may be defined either in attribute form (operative or inoperative) or in terms of tolerance limits on one or more parameters.

(2) Select the time period, t, for which reliability prediction is to be made. Thus, t represents the desired life of the component in its intended application.

(3) Select a sample of components representative of the population for which the screening criterion is being derived. Appropriate statistical randomization procedures should be used in selecting the sample.

(4) Determine the parameters to be used as potential predictors of failure. These may be determined from engineering judgment and/or prior test data.

(5) Obtain initial parameter measurements. If parameter change over a short time period, Δt, is considered as a potential predictor, then parameter measurements must be obtained after the operating interval Δt.

(6) Run the life test for a time period, t, under actual or simulated operating conditions to be encountered in the intended application of the component.

(7) At the end of the test, obtain the parameter measurements in terms of which failure is defined for each component.

(8) Separate the components into two sets: S_1, consisting of those components that did not fail during the test, and S_2, consisting of those components that failed.

Downloaded from https://www.everyspec.com

As an example of a criterion of failure, in a twofold classification, each component may be judged to have failed or not to have failed at the end of the test. The criterion of failure may be based on an attribute definition. For example, those components may be called failures which are not operative at the end of the test. All others are judged not to have failed. Alternatively, it may be that all of the components are operative at the end of test, those component parts being called failures which exhibit a drift in excess of some specified value of some measured parameter. The importance of the failure definition stems from the fact that the screening criterion determined by the linear-discriminant analysis attempts to reproduce from initial or early life measurements the same classification assigned at the end of the test. The prediction classification may be based on initial measurements taken at t = 0 or on early-life measurements taken, for example, at t = 0 and t = 100 hours.

5.2.4.4 The Linear-Discriminant Analysis

For each set of measured parameters, there exists a linear-discriminant screening criterion. For example, if three parameters are measured, it would be possible to construct seven different linear discriminants. Three of these would involve each parameter singly; three would involve a combination of two parameters; and one would involve all three parameters. By examination of the probabilities of misclassification associated with each of these screening criteria, it is possible (1) to determine the minimum number of parameters required for practical screening, (2) to determine the gain in reliability obtained by screening on the basis of two parameters rather than one, three rather than two, etc., and (3) to determine those parameters that may be interchanged for measurement or cost reasons without changing the effectiveness of the screening procedure. In cases where many different parameter measurements are available, it is not feasible to examine all possible subsets and their associated linear discriminants.* In these cases, various methods may be used to select those parameters that appear to be the most suitable for inclusion in a linear discriminant.

The steps below describe the computations required for obtaining a linear-discriminant screening criterion, given the parameters to be used as predictors.[3,4]

The construction of the reliability screening criterion based on a specified set of parameters consists of seeking a quantitative measure of the difference between the superior and inferior components. This measure should be based on initial and early-life measurements and serve to predict the qualities of the component during late life. Let S_1 denote the set of initial and early-life measurements associated with the superior components and S_2 those measurements associated with the inferior components. We want to define a "discriminating" function of the measurements in S_1 which assumes values that are distinctly different from the values assumed by the function for the measurements in S_2. An ideal discriminating function would be one that yields 1 for all measurements associated with superior components and yields 0 for measurements associated with inferior components. In practice, ideal discriminating functions do not exist. Instead, an attractive approximation to such a function is given by a linear combination of the measured parameters:

$$z = \lambda_1 x_1 + \ldots + \lambda_p x_p \quad , \tag{5.16}$$

*The problem in this case is that each discriminant derivation requires the inversion of a matrix, which is an "expensive" process. Obviously, if a sufficiently large computer is available, greater numbers of parameter combinations can be tried.

Downloaded from https://www.everyspec.com

where x_i, i = 1, . . ., p, denotes a measured value of the ith parameter, and the λ's are numerical "weights". When the optimal λ's are determined as described below, the linear combination of measured values is called a "linear discriminant". The optimal λ's are those values of the λ_i that maximize the quantity $(\bar{z}_2 - \bar{z}_1)/s$. That is, the optimal λ's maximize the distance between the means of the distributions of superior and inferior components, $\bar{z}_1$ and $\bar{z}_2$, respectively, expressed in units of the standard deviation, s. The statistical properties of linear discriminants were first developed by R. A. Fisher.[5] Mathematical details of such analyses are given in several statistical texts. In the following, the essential computational steps in deriving a linear discriminant are described.

(1) Partition the initial and/or early-life measurements of parameters $x_1, \ldots, x_p$ into two sets: (a) those associated with components in S_1 and (b) those associated with components in S_2.

(2) Compute the sums of squares and products, within the set of superior components, S^1_{pq}. Repeat for the inferior components, S^2_{pq}. These are given by

$$S^1_{pq} = \Sigma_i \Sigma_j \left(x^1_{ip} - \bar{x}^1_p\right)\left(x^1_{jp} - \bar{x}^1_q\right) , \tag{5.17}$$

where x^1_{ip} = the value of parameter p of the ith component in S_1 and $\bar{x}^1_p$ = the average value of parameter p in S_1. S^2_{pq} is obtained in the same way.

(3) Obtain the sum

$$S_{pq} = S^1_{pq} + S^2_{pq} \quad . \tag{5.18}$$

(4) Form the matrix equation

$$S\lambda = d \quad , \tag{5.19}$$

where

$$S = \begin{bmatrix} S_{11} & \cdots & S_{1p} \\ \vdots & & \\ S_{p1} & \cdots & S_{pp} \end{bmatrix} , \quad \lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{bmatrix} , \quad d = \begin{bmatrix} \bar{x}^1_1 - \bar{x}^1_1 \\ \vdots \\ \bar{x}^1_p - \bar{x}^2_p \end{bmatrix} .$$

(5) Solve Equation (5.19) for λ, which gives the λ-weights for the linear-discriminant [Equation (5.16)].

(6) Compute the quantity

$$\delta = \frac{\bar{z}_2 - \bar{z}_1}{s} = \left[(n_1 + n_2 - p - 1) \sum_{i=1}^{p} \lambda_i d_i \right]^{1/2} , \tag{5.20}$$

Downloaded from https://www.everyspec.com

where n_1 is the number of components in S_1 and n_2 the number in S_2. The quantity δ denotes the distance between the mean values for superior and inferior components in units of the standard derivation, s. δ constitutes a measure of the power of the screening criterion in that the greater δ, the smaller the probabilities of misclassification.

(7) Compute the probabilities of classification as shown in Table 5.3.

TABLE 5.3. EXPECTED CONDITIONAL PROBABILITIES OF CLASSIFICATION

Given $z^* = (\bar{z}_2 + \bar{z}_1)/2$

Assigned Classification	Conditional Probability of Classification	
	Superior	Inferior
Superior	$\Phi(\delta/2)$[a]	$1 - \Phi(\delta/2)$
Inferior	$1 - \Phi(\delta/2)$	$\Phi(\delta/2)$

(a) $\Phi(y)$ = the area to the left of y under the standard normal distribution curve.

For example, the probability of misclassifying a "truly" inferior component as superior is seen to be $1 - \Phi(\delta/2)$.

(8) Repeat Steps (1) - (7) for each combination of parameters of interest and compare the respective "powers" (δ) and resultant probabilities of classification. Acceptance of a given discriminant function will depend on its relative power balanced against the "costs" of applying it.

Because of the assumptions that must be made in the course of developing the criterion and because these assumptions are often incapable of direct verification, it is desirable to make an indirect check on the validity of the linear-discriminant screening procedure. This is accomplished by applying the linear-discriminant screening criterion to screen those components used to generate the test data. That is, the initial and early-life measurements are substituted into the discriminant, and each component part is classified by comparing its z value with the critical value of z. It should be noted that this is principally an internal consistency check and, as such, does not validate the general applicability of the linear discriminant. Ideally the linear discriminant will produce the same classification assigned to the component parts at the end of the test. The extent to which the linear discriminant achieves the end-of-test classification is an indication of the practical effectiveness of the discriminant screening criterion. If the end-of-test classification is not closely approximated by the discriminant screening based on early-life measurements, then it is judged that the screening criterion is of little practical significance. This result may indicate that none of the measured parameters are relevant to predicting the correct classification. Alternatively, it may indicate that some assumption is incorrect and that transformations of the data should be considered as a method to improve the classification.

Downloaded from https://www.everyspec.com

If the linear discriminant yields good agreement with the classification at the end of the test, then it may be applied to other sets of component parts of known quality which were not used in its construction. If the discriminant passes all verification checks successfully, then it may be used to screen component parts of unknown quality having similar statistical properties. The more thorough the verification, the more assurance one has that the screening procedure will perform properly.

5.2.4.5 Application of the Linear-Discriminant Screening Criterion

Once a linear discriminant has been developed, sets of component parts of unknown reliability can be screened. This is accomplished by

(1) Determining a critical value, z^*

(2) Obtaining a z value for each component to be screened

(3) Classifying as superior those components for which $z < z^*$ and classifying as inferior those components for which $z > z^*$ (this assumes that $\bar{z}_2 > \bar{z}_1$).

The expressions for the classification probabilities given in Table 5.3 assume that z^* is midway between $\bar{z}_1$ and $\bar{z}_2$. In general, however, z^* may be chosen to satisfy some specified criterion function. Such criteria may include (1) minimization of cost of misclassification or (2) specification of some acceptably small probability of misclassifying an inferior component as superior. Thus if we define the critical value by

$$z^* = a\,\bar{z}_1 + (1 - a)\,\bar{z}_2,\ 0 \le a \le 1 \quad , \tag{5.21}$$

then the expected conditional probabilities of classification are as given in Table 5.4. Thus, in general the critical value, z^*, will not be located midway between $\bar{z}_1$ and $\bar{z}_2$ as illustrated in Figure 5.8.

TABLE 5.4. EXPECTED CONDITIONAL PROBABILITIES OF CLASSIFICATION

Given $z^* = a\,\bar{z}_1 + (1 - a)\,\bar{z}_2$

Assigned Classification	Conditional Probability of Classification	
	Superior	Inferior
Superior	$\Phi\,[(1 - a)\,\delta\,]^{(a)}$	$1 - \Phi\,[(1 - a)\,\delta]$
Inferior	$1 - \Phi\,(a\,\delta)$	$\Phi\,(a\,\delta)$

(a) $\Phi(y)$ denotes the area to the left of y under the standard normal distribution curve

Downloaded from https://www.everyspec.com

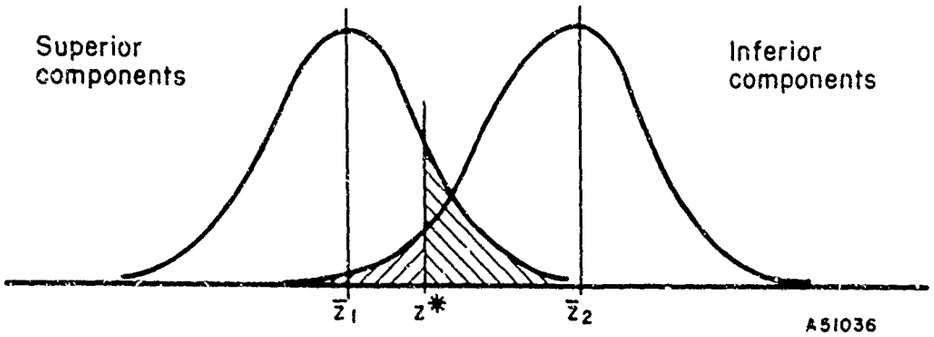


FIGURE 5.8. SCHEMATIC REPRESENTATION OF THE THEORETICAL DISTRIBUTION OF z-VALUES FOR THE SUPERIOR AND INFERIOR COMPONENTS

5.2.4.6 Relation of Linear-Discriminant Analysis and Physical Degradation Laws

The parameters characteristically used in a linear-discriminant screening criterion are operational parameters. That is, they are generally those parameters in terms of which component performance is measured. For example, a criterion for screening transistors may be based on initial or early-life parameter measurements such as collector cutoff current, emitter cutoff current, collector breakdown voltage, etc. How well this class of variables serves as predictors of component reliability is determined from statistical analysis of test data rather than theoretical considerations. The net effect of reliability prediction based on this class of variables is that their potential "power" is bounded by the degree of statistical variation in the operational parameters serving as measures of the fundamental aging process acting in a component.

If, however, progress in reliability physics can uncover those fundamental parameters characterizing aging processes such as activation energies, frequency factors, and diffusion coefficients, then significantly more powerful screening criteria can potentially be developed. Moreover, the screening criteria based on this more fundamental class of parameters may take on or be closely related to the form of the physical laws governing aging processes.

For example, suppose activation energy and the frequency factor were taken as the parameters for a linear-discriminant screening criterion. The analysis may show that the linear discriminant

$$z = \lambda_1 x_1 + \lambda_2 x_2 \tag{5.22}$$

should take on the form

$$z = \ln (\text{frequency factor}) - \frac{1}{kt} (\text{activation energy})$$

$$= \ln A - \frac{1}{kt} \Delta E \quad , \tag{5.23}$$

Downloaded from https://www.everyspec.com

where $\lambda_1 = 1$, the coefficient of the pre-exponential term, and $\lambda_2 = -1/kt$.

From Equation (5.23), it is seen that the rate of aging, R, is related to the discriminant z-value by

$$R = e^z = Ae^{-\Delta E/kt} \tag{5.24}$$

Thus, in terms of linear-discriminant screening, this example states that the optimal discriminant is dependent on the logarithmic aging rate of a component. In other words, the discriminant that maximizes the distance between the distributions of z-values for superior and inferior components depends only on a single parameter – the aging rate.

An important result of deriving linear discriminants from more fundamental parameters directly related to aging processes acting in components is that several significant problems in executing a linear-discriminant analysis may be alleviated. These include:

(1) Physical theory may preclude the necessity of testing many parameter combinations to find that which is optimal. Only a small number of tests may be required, each involving a small number of parameters. Thus, the computational problem will be greatly reduced.

(2) Physical theory may indicate the type of parameter transformation that should be made, thus yielding the best agreement between expected and observed screening results.

(3) Physical theory may provide a basis for verification of the λ-weights in the linear discriminant.

(4) Physical theory may permit the derivation of linear discriminants under accelerated test conditions that can be used to screen components having intended application in various other stress environments.

In general, the potential advantage to component screening of development in reliability physics is that screening criteria based on a more fundamental class of variables can be derived more cheaply, be validated more easily, and may be significantly more powerful. It should be recognized, however, that these inferences are desired objectives and, as such, remain to be demonstrated.

5.3 Comparison Among Screening Procedures

In Section 5.2, four procedures for screening components were described whose purpose is to eliminate potentially unreliable components prior to their use in a specified application. These procedures were arranged in order of increasing complexity to develop and/or apply. The selection of which procedure is most appropriate in a given instance depends on the properties of the component type, the screening power required, and the allowable "costs" in developing the screening criterion. In the following paragraphs, this selection problem is treated in terms of a comparison of the attributes of constructing and applying the four screening procedures.

Downloaded from https://www.everyspec.com

5.3.1 Comparison of the Screening Procedures in Constructing a Screening Criterion

To construct a screening procedure, it is required to (1) select those component parameters on the basis of which screening will be accomplished, (2) develop the screening function, and (3) determine the performance criteria for component classification. For screening by truncation of distribution tails, stress-strength screening, and burn-in screening, component-parameter selection procedures are not explicitly given as part of the analysis. In these methods, the parameters on which screening is based are determined from component tests not formally a part of the screening procedure, prior tests on the same or similar components, or engineering judgment. In contrast, for linear-discriminant screening the parameters yielding the best discriminant function are determined explicitly by the formal analysis in the derivation procedure. Moreover, in linear-discriminant analysis, screening is based on a single dependent variable, which is a linear combination of the relevant component parameters, whereas screening by the other three methods is based on one or more component parameters treated independently.

Development of a screening function (the second general step in deriving a screening procedure) for screening by truncation of distribution tails and burn-in screening follows directly as a product of Steps (1) and (3). That is, no direct experimentation or analysis is conducted relating initial parameter measurements to potential early failures. For these methods, construction of the screening procedure is completed by setting tolerance limits on those performance parameters in terms of which failure is defined. For stress-strength testing, it is necessary to determine a distribution function characterizing the stress environment to which the components will be subjected. The screening procedure is then completed by setting a lower tolerance limit on the distribution of part strengths that will act to eliminate those components that are potential failures in the overlap region of the stress-strength distributions. In the case of linear-discriminant screening, the construction of a screening function requires that a load-life test be conducted on a sample of components for the time period on the basis of which subsequent component lots will be screened. A linear-discriminant screening function is then derived from the life-test data. The construction of this procedure is then completed by determining a critical value yielding acceptable probabilities of misclassification.

An additional factor that should be recognized in the construction of a screening procedure is the assumptions associated with the process. These assumptions may be either explicitly involved in the derivation of a screening criterion or implicit in the application of the screening criterion. In all four of the screening procedures presented in this section, explicit assumptions must be made as to the distribution form of the screening parameters. In screening by truncation of distribution tails, stress-strength screening, and burn-in screening, tolerance limits are determined in part with respect to this assumption. Moreover, in stress-strength screening an explicit assumption must be made as to the distribution of stresses. In linear-discriminant screening, the probabilities of misclassification are calculated on the basis of this assumption in addition to the assumption that the variances of the z-values for the superior and inferior components are equal.

In addition to these explicit assumptions, screening by truncation of distribution tails, stress-strength screening, and burn-in screening carry implied assumptions with respect to changes in component performance parameters over time. For screening by truncation of distribution tails and stress-strength, the assumption of a degradation rate

Downloaded from https://www.everyspec.com

may be considered as a strong assumption since screening is accomplished based on initial parameter measurements. In the case of burn-in screening, the assumption is somewhat weaker since actual parameter changes can be observed for a short period of time before screening is actually accomplished. It is this assumption that leads to the marginal tolerance limit concept for the three screening procedures. Linear-discriminant screening, on the other hand, does not involve this assumption since the screening function is not derived from an extrapolation of expected component performance over time.

5.3.2 Comparison of the Screening Procedures in Application

The application of a screening procedure requires that (1) the appropriate parameter measurements be obtained and (2) these measurements be compared against the screening criterion to classify the components. The important questions to be answered with respect to choosing among the four screening procedures are:

(1) What effort, in terms of time and cost, is required for a given procedure?

(2) What is the screening power attained by the respective procedures?

Screening by truncation of distribution tails and stress-strength screening are the easiest of the four procedures to apply, since screening is accomplished at time $t = 0$ based on initial parameter measurements. Linear-discriminant screening may possibly be based on a linear combination of initial measurements although in practice this is not generally the case. In most instances, more powerful linear-discriminant screening criteria result when some incremental parameter change measurements are included in the discriminant. Burn-in screening, as implied by its name, requires that screening be accomplished on the basis of parameter measurements after some short period of operation. Moreover, a burn-in test is characteristically run at stress levels at or near rated operating conditions of the component. Thus linear-discriminant screening and burn-in screening are generally more expensive to apply since a test period and equipment for conducting the test are required in addition to the instrumentation needed in all cases to obtain the required parameter measurements.

The screening "power" attainable in applying the various procedures parallels the "costs" of application described in the preceding paragraph. The increase in lot reliability attained in screening by truncation of distribution tails and stress-strength screening is characteristically weaker than that attained in burn-in and linear-discriminant screening. An important consideration in this respect is that the "power" in screening by truncation of distribution tails and burn-in screening is not directly measurable. In order to achieve such measures in these two cases, it is necessary to conduct external life tests and to correlate performance parameter changes to the setting of the marginal tolerance limits for the screening procedure. In stress-strength screening, power, in terms of increase in reliability, can be measured in theory; however, in practice, the actual calculations are extremely difficult and must usually be determined by approximating methods. In linear-discriminant screening, on the other hand, power, in terms of probabilities of misclassification is explicitly measurable as a part of the discriminant analysis. Moreover, it is a straightforward procedure to determine the trade-off in power against the number of parameters in the discriminant, exclusion of certain parameters for "cost" reasons, and variation of the critical value.

Downloaded from https://www.everyspec.com

In summary, it is seen that in choosing among the four screening procedures, (1) screening by truncation of distribution tails, (2) stress-strength screening, (3) burn-in screening, and (4) linear discriminant screening, each succeeding procedure takes into account more information at some cost either in the construction or application of the procedure, or both. These comparative properties of the screening procedures are further summarized in Tables 5.5, 5.6, and 5.7.

TABLE 5.5. COMPARISON OF ADDITIONAL INFORMATION INCLUDED IN SUCCESSIVELY MORE COMPLEX SCREENING PROCEDURES

Screening Procedure	Factors Included in the Screening Function
Truncation of distribution tails	Marginal tolerance limits on initial parameter measurements
Stress-strength	Marginal tolerance limits on initial component strength
	Specification of distribution of stresses
Burn-in	Marginal tolerance limits on parameter measurements or drift after short operating period
	Testing under specified stress environments
Linear discriminant	Parameters selected by analysis to maximize screening power
	Screening power directly computed in analysis for specified life time and stress environment

Downloaded from https://www.everyspec.com

TABLE 5.6. SUMMARY COMPARISON OF ATTRIBUTES IN THE CONSTRUCTION AND APPLICATION OF THE VARIOUS SCREENING PROCEDURES

Attribute	Truncation of Distribution Tails	Stress-Strength Screening	Burn-In Screening	Linear-Discriminant Screening
		Construction		
Screening parameters determined by analysis				X
Screening function based on individual parameters	X	X	X	
Screening function based on combination of parameters				X
Stress explicitly recognized in screening procedure		X		X
		Application		
Procedure based on initial measurements	X	X		
Procedure based on parameter change measurements (short-term test required)			X	X
Screening power directly measurable		X		X
Relatively "inexpensive" to apply	X	X		
Relatively "high" in power			X	X

TABLE 5.7. SUMMARY OF COMPARATIVE PROPERTIES OF VARIOUS SCREENING PROCEDURES

Screening Procedure	Properties				
	Concept	Assumptions	Requirements in Development of Screening Criteria	Difficulty in Application	Power
Truncation of Distribution Tails	Marginal tolerance limits on screening parameters	Parameter distribution forms Form of degradation process	Design analysis to determine parameters and tolerance limits Life test not explicitly required	Easily applied, based on initial parameter measurements	Not explicitly determined by analysis; generally weak
Stress-Strength	Elimination of components having inferior strength	Form of stress-strength distributions Model for combining stress elements	Derivation of combined stress distribution Life test not explicitly required	Easily applied, based on initial strength values	Slightly more powerful than truncation; increase in static reliability can be calculated (in theory)
Burn-In	Marginal tolerance limits on screening parameters	Parameter distribution forms Extrapolation of degradation process	Design analysis to determine screening parameters and tolerance limits Life test not explicitly required	Requires short-term test usually at high stress levels	Not explicitly determined by analysis; generally significantly more powerful than truncation
Linear Discriminant	Optimum linear classification model	Statistical regression analysis Does not assume extrapolation of degradation process	Life test of sample of components Discriminant analysis of life test data	May require short-term test; some computations required; easily programmed for computer	Generally most powerful of techniques presented; power directly measurable in terms of probabilities of classification

Downloaded from https://www.everyspec.com

Downloaded from https://www.everyspec.com

5.4 Survey of Precursors and Failure-Sensitive Device Parameters

The search for those component parameters that act successfully in the various screening procedures appears to be a never-ending procedure. A parameter that is a good predictor of early failures for metal film resistors may not work at all for wire-wound resistors. Moreover, the parameter selection problem is highly interacting with the type of the applied stresses (temperature, voltage, etc.) in a screening test, the method of applying the stress (step voltage input, temperature cycling, etc.), the application for which the component is designed, and the measured characteristics of the parameter (incremental change, stability, dynamic response, etc.).

It is the objective of this section to identify those component parameters that have been successfully used in various screening applications. The following types of electronic components are considered:

(1) Resistors

(2) Capacitors

(3) Diodes

(4) Transistors.

5.4.1 Screening Parameters for Resistors

Table 5.8 lists measurement parameters that have been used in developing screening criteria for each of three types of resistors.

TABLE 5.8. SCREENING PARAMETERS FOR RESISTORS

Screening Parameter	Resistor Type		
	Carbon Composition	Carbon Film	Metal Film
Resistance	X	X	X
Temperature coefficient of resistance	X	X	
Voltage coefficient of resistance	X		
Short-time voltage overload	X	X	X
Short-time humidity exposure		X	
Thermal noise	X		
1/f noise	X		X

Downloaded from https://www.everyspec.com

5.4.1.1 Carbon-Composition Resistors

The resistance values of carbon-composition resistors is decreased by the embrittlement and carbonization of the epoxy-resin binder due to thermal aging. Also, resistance values may be increased by vapor penetration into the porous epoxide-carbon pellets under operation at mild stress levels. Thus, depending on the intended application, the measurement of resistance change in high-thermal or high-humidity environment is used as a screening parameter for carbon-composition resistors. Also, bondage imperfections between the epoxide and carbon particles may be identified by measurements of temperature coefficient of resistance, voltage coefficient of resistance, 1/f noise, or thermal noise[3].

Potential early failures may slso be identified by the change in resistance resulting from short-time voltage overload. Permissible overloads depend on whether the resistance values of the resistors are above or below their critical resistance values. The critical resistance is defined for a class of resistors with a given value of rated wattage and rated voltage, e.g., 1/2-watt, 500-volt resistors, as the ratio of the square of the rated voltage to the rated wattage,

$$R_c = \frac{V^2 \text{ (rated)}}{W \text{ (rated)}} \quad . \tag{5.25}$$

In the case of 1/2-watt, 500-volt resistors, the critical resistance is 500 K ohms. Thus, if the resistance of a specific 1/2-watt, 500-volt resistor is less than 500 K ohms, it will be wattage limited, whereas if it is greater than 500 K ohms, it will be voltage limited. This suggests that the failure mechanisms operative in a given resistor type may depend on whether the resistance value is greater or less than the critical resistance. Generally, resistors with $R < R_c$ may withstand higher voltage overloads but shorter application times than those with $R > R_c$.

5.4.1.2 Carbon-Film Resistors

The resistance values of carbon-film resistors are affected by thermal aging of the acrylic or resin coatings covering the films. These resistors are generally characterized by a decrease in resistance during early life, followed by an increase in resistance. Thus, resistors with deteriorated film coatings may be identified by a decrease in resistance from burn-in procedures.

As is the case with carbon-composition resistors, short-time voltage overload is an appropriate screening parameter for carbon film resistors. Also, short-time humidity exposure is used for screening carbon-film resistors that are to be used in high-humidity environments.

5.4.1.3 Metal-Film Resistors

Because of their high-temperature characteristics, metal-film resistors whose resistances are less than the critical resistance can be screened by the application of three times the voltage required for rated power dissipation for a period of 10 seconds. For resistors with resistance values greater than the critical value, the over-voltages should be

Downloaded from https://www.everyspec.com

limited to two times the voltage required for rated power dissipation but may be applied for longer periods of time. The wattage dissipation associated with excessive over-voltage tends to damage the resin coating of the resistor, increasing its conductivity and thus providing a parallel conduction path that would tend to lead to thermal runaway.

Excessive 1/f noise, measured in tin oxide film resistors, has been used successfully to identify resistors demonstrating changes in resistance and in temperature coefficient of resistance during subsequent load-life tests[6].

5.4.2 Screening Parameters for Capacitors

At the present time, most reliability screening techniques for thin-film capacitors are more nearly process quality-control measures than reliability screening measures. There are several reasons for this. First, a thin-film capacitor is usually part of a network of thin-film passive devices deposited on an insulating substrate. It is more reasonable and sometimes necessary to screen the circuit as a whole rather than an individual component of the circuit, since the individual components are not replaceable. Second, the technology of thin-film circuitry is new. Emphasis thus far has been placed on achieving desired levels of parametric performance in thin films rather than on devising reliability screening techniques. Third, the technology of thin-film circuitry is changing. New techniques and materials are being tried frequently. Since the properties and characteristics of thin-film circuitry are highly sensitive to processing methods, screening techniques of long-term validity will probably not be found while methods of manufacture are undergoing frequent change. On the other hand, it is considered appropriate to discuss mechanisms that have been observed in thin-film capacitors which could possibly lead to device failure and to indicate also the modes of failure which these mechanisms are likely to precipitate. It should be pointed out that these observations have been made by investigators in the course of device fabrication research, and the mechanisms described have not been used as the bases for screening techniques.

There are several ways in which a capacitor can fail: capacitance may drift out of tolerance, dissipation factor may increase out of tolerance, leakage current may become excessive, or the dielectric may break down, either causing the capacitor plates to short circuit or causing other undesirable changes in the capacitor characteristics. These failures may be caused by electrochemical reactions, impurities in the dielectric, porousness, pinholes or cracks in the dielectric, diffusion of atoms from the metal capacitor plate into the dielectric, or irregularities in the dielectric thickness.

Ion migration is believed to have caused a pattern of dielectric breakdown observed by Best, et al.[7], in a potash-lead-silicon glass capacitor as a result of voltage step-stress testing. Capacitors subjected to short-time (30 minute) periods at each voltage in the stress program failed by dielectric breakdown occurring at the capacitor edges. Capacitors subjected to long-time periods (several hours) at each voltage failed by dielectric breakdown at the center of the capacitor. Capacitors in the latter category broke down at higher voltages than those in the former. The reason hypothesized for this behavior is that, under prolonged application of voltage, sodium ions migrate toward the negative electrode, causing a space-charge layer to form. The space-charge layer caused the electric field in the region of the plate to be uniform. Since normally the electric field in a capacitor is much higher at the dielectric edges than at the center, breakdown is more likely to occur at the edges. Accumulation of a space-charge layer in the

Downloaded from https://www.everyspec.com

capacitors observed by Best tended to equalize the electric field and reduce the effect of fringing at the foil edges. In this instance, judging from the available information, the aging stress tended to improve the capacitors. However, it seems likely that over the long run such a phenomenon would have a deleterious effect on capacitor performance.

A phenomenon believed to be migration has been observed in silicon monoxide used to overcoat a Nichrome film resistor.[8] Alkali ions presumably contributed by the glass substrate caused discoloration and cracking of the SiO film. The effect occurred only when elevated temperature, an electric field, and water vapor were present. Although this effect has not been related to capacitor failure, the widespread use of SiO as a thin-film capacitor warrants consideration of the effect as a possible failure mechanism.

Klein has studied nonshorting dielectric breakdown in silicon monoxide thin-film capacitors.[9] Nonshorting breakdown did not destroy the capacitor but left a small hole in the dielectric. The consequence of these holes was to reduce capacitance owing to decrease of effective area. The effect was more pronounced in the presence of water vapor. It was believed that the effect was due to pores in the dielectric. When water vapor adsorbed on the pore wall, a conducting path was provided. The electric field in the dielectric between the bottom of the pore and the second electrode was thus higher than that between the two electrodes, and premature breakdown occurred.

Diffusion of tantalum atoms into the Ta_2O_5 dielectric has been observed in tantalum thin-film capacitors.[10] Such diffusion is greater in the presence of elevated temperatures and applied electric field. The observed diffusion could be used to explain capacitance versus frequency characteristics of this kind of capacitor. It is possible that diffusion of metal atoms into the dielectric could cause increased capacitance, increased current leakage, and decreased breakdown voltage.

In a continuation of the work described in Reference (10), the diffusion of gold, silver, chromium, nickel, and cobalt into boro-alumina silicate and thermally grown SiO_2 dielectrics is being studied.[17] As yet, metal diffusion has not been established as a mechanism of failure in tantalum oxide capacitors.

The above-mentioned effects are not used for screening and have not been related to thin-film capacitor failure. They are included here because it is felt that, when they are better understood and more readily recognized, they will be suitable bases for reliability screening tests.

Two test methods are known which could be used to inspect capacitors for flaws in the dielectric or other defects that could cause failure. These are infrared scanning[12-14] and observation of anion reactions[15]. Infrared scanning would show hot spots on the electrode, which would be indicators of high current density and high electric field. Anion reactions would indicate visually the presence of pinholes or voids in the dielectric. It is not likely that these tests could be performed on an encapsulated device, however, since they are both forms of visual inspection.

A promising screening technique now under development for use with solid tantalum electrolytic capacitors may be useful in screening thin-film tantalum capacitors.[16] In

Downloaded from https://www.everyspec.com

this test, capacitors are put in an oven at 125°C, and a voltage of 120 per cent of the rating voltage is applied. After 1/2 hour and at the end of 5 days, leakage current is measured. High leakage currents or a large change between the initial and final values indicate a device prone to failure. Screening efficiencies as high as 10* have been achieved.

To summarize, techniques have not been devised for reliability screening of thin-film capacitors. Several inspection techniques may be applicable to unfinished devices but not to completed, encapsulated units. It is very possible, since thin-film capacitors will usually be mounted on a substrate inseparable from other circuit elements, that screening techniques as such will not be devised for thin-film capacitors but will be devised for integrated circuits instead. Even in that case, however, it will be desirable to have knowledge of mechanisms by which thin-film capacitors can fail, and to that end several such mechanisms have been included here.

5.4.3 Screening Parameters for Diodes

Table 5.9 lists parameters that have been used in developing screening criteria for diodes.

Screening by the first six parameters in Table 5.9 is done by the truncation of the distribution tails of operating parameters on the assumption that marginally acceptable parameters will tend to drift out of tolerance.

Thermal storage at greater-than-maximum operating temperature may be used to accelerate adsorption, desorption, or chemical reaction of contaminants that may exist on or be transported to a sensitive surface.

*Screening efficiency is defined as the fraction of 'bad' devices removed from the population divided by the fraction of the population judged as 'potentially bad' by the screening criterion[17]. This is illustrated by the following three examples.

Example	Lot Size	No. of Devices Removed From Lot	Total No. of Failures	No. of Failures in Removed Group	Ratio of No. of Failures Removed to Total No. of Failures	Ratio of No. of Devices Removed to Total No. of Devices	Efficiency
(1) Referenced screening procedure	1,000	90	10	9	0.9	0.09	10
(2) Random selection	1,000	100	10	1	0.1	0.1	1
(3) Perfect selection	1,000	10	10	10	1	0.01	100
	10,000	10	10	10	1	0.001	1,000

It will be noted from Example (3) above that the screening efficiency is not independent of the fraction of bad devices in the lot.

Downloaded from https://www.everyspec.com

TABLE 5.9. SCREENING PARAMETERS FOR DIODES

(1)	Reverse current at specified voltage
(2)	Reverse breakdown voltage – voltage at specified current level
(3)	Resistance at specified forward voltage
(4)	Open-circuit voltage decay time
(5)	Reverse recovery time
(6)	Power operation at near-rated maximum temperature
(7)	Thermal storage
(8)	Thermal cycling
(9)	Slope of the forward current-voltage characteristic, as $\frac{\Delta \log I_F}{\Delta V}$
(10)	Slope of the reverse current-reciprocal temperature curve, as $\frac{\Delta \log I_R}{\Delta T^{-1}}$
(11)	Avalanche noise

Thermal cycling is used to accelerate the effects of physical defects such as non-hermetic seals and the tendency to fracture of adjacent materials with different thermal coefficients of expansion. The effects of the physical defects are detected by the measurement of one or more operational parameters.

Excessive surface channel current may be identified by the parameter numbered (9) in the table. The diode equation for forward current may be written as $I_F = I_o \exp (qV/nkT)$, where $\frac{1}{n}$ is proportional to the slope of $\log I_F$ versus V curve. Sah[18] has shown the value of n is always less than 2 for minority-carrier diffusion current and for bulk or surface generation-recombination current, but $2 < n < 4$ for surface channel current. Determination of the value of n, therefore, may be used to screen against devices with predominant surface channel current.

In cases where surface channel current may be abnormally high, and yet may not be predominant, the Parameter (10) (Table 5.9) may be used to identify the surface channel component of the reverse current[19,20]. In this technique, cooling the diode to appropriately low temperatures "freezes out" contributions to the current from generating

Downloaded from https://www.everyspec.com

mechanisms with higher activation energies, i.e., minority-carrier diffusion and generation-recombination current, revealing the magnitude of current components with low activation energies, e.g., surface channel current.

The measurement of the noise current associated with the onset of avalanche conduction through localized regions of a reverse biased diode may be used to screen those units with localized breakdown within operating voltage. Although the magnitude of the reverse current from conduction through a single microplasma may not be sufficient in some cases to disqualify a diode for service, continuous operation of the diode with localized avalanching would be expected to result in localized heating and contribute to early degradation or failure.

5.4.4 Screening Parameters for Transistors

Table 5.10 lists measurement parameters that have been used in developing screening criteria for transistors.

TABLE 5.10. SCREENING PARAMETERS FOR TRANSISTORS

I_{CBO},	Collector-base reverse leakage current
h_{FE},	Direct-current gain
BV_{CBO},	Collector-base breakdown voltage
I_{EBO},	Emitter-base reverse leakage current
P_{CE},	Collector-emitter power dissipation
E_e^2,	Excess noise (1/f noise)
N,	Current-noise voltage

Depending on the type of screening procedure to be applied, the problem is what characteristic of these parameters should be measured, and under what test specifications the measurements should be made.

Collector-base reverse leakage current, I_{CBO}, has been the most widely used screening parameter. Burn-in screening, linear-discriminant screening, and an initial measurement procedure have been successfully applied using I_{CBO}. In burn-in testing of a silicon mesa transistor, change in leakage current, ΔI_{CBO}, measured at a constant collector voltage over a 350-hour test period successfully eliminated 85 per cent of the early failures.[21] Environmental stresses for this test consist of operating the transistor at maximum rated power dissipation and at near-rated operating temperature. In linear-discriminant screening, a discriminant function of the form

$$z = \lambda_1 \ln I_{CBO} + \lambda_2 \ln \Delta I_{CBO} \quad , \tag{5.26}$$

Downloaded from https://www.everyspec.com

where I_{CBO} was measured at V_C = +50 volts, yielded high probabilities of correct classification that were statistically significant. A third procedure involved measuring the time response of I_{CBO} to a voltage step input.(22) The time-response characteristics of I_{CBO} are given in Figure 5.9. The figure shows that I_{CBO} rises quickly to a sharp peak and decays exponentially to a fixed leakage current value for potentially reliable transistors. For potential early failures, however, the I_{CBO} time response is seen to go through a relatively long oscillation before stabilizing. Screening on the basis of this time-response characteristic to predict degradation failure was shown to be highly efficient on a PNP, germanium, high-power alloy transistor.

Considerable evidence also exists indicating that 1/f noise measured at either the emitter base or collector base is a good predictor of catastrophic failure. In this case, measurement of 1/f noise under reverse bias at near-rated power dissipation but below onset of avalanche was indicated as most effective.(22)

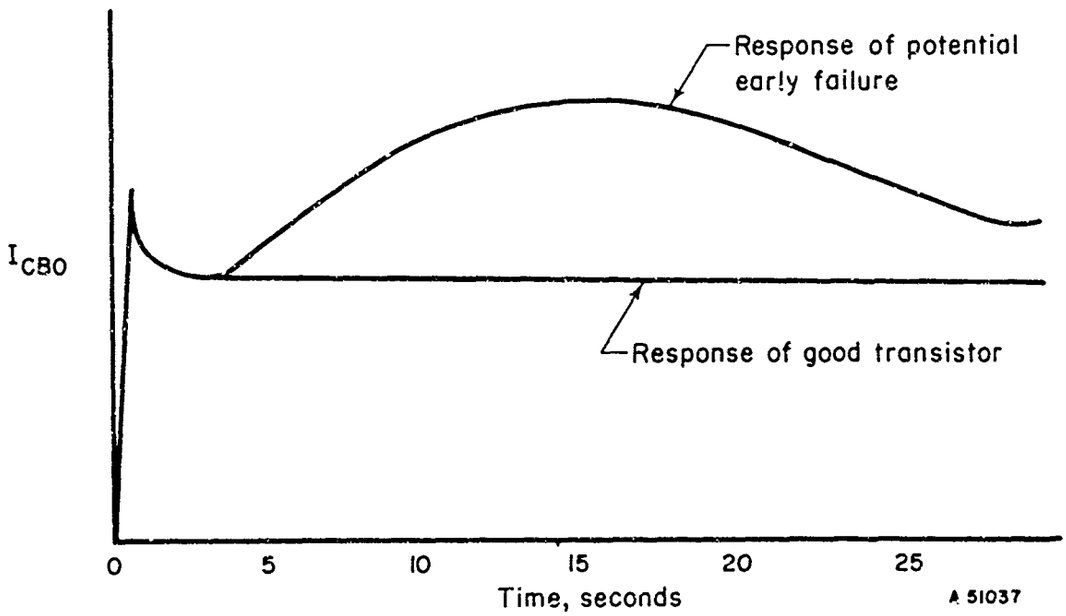


FIGURE 5.9. TIME RESPONSE OF COLLECTOR-BASE REVERSE LEAKAGE CURRENT

5.4.5 Summary

Examination of the type of parameters illustrated above used for screening shows that, at the present state of the art, screening is accomplished at the operational level. That is, this class of component parameters predict at best potential early failure occurring in observable failure modes. For example, excessive collector leakage current, I_{CBO}, during early life may be a precursor of transistor early failure due to the drift of BV_{CBO} or I_{CBO} outside of operational tolerance limits, or 1/f noise may act as a precursor of early failure due to open or shorted contacts.

Downloaded from https://www.everyspec.com

It is hoped that [illegible] research in reliability physics theoretical physical models will ultimately be derived that relate the observable electrical characteristics of a component to basic failure mechanisms. Thus, effectively, a failure mechanism serves as a theoretical model at the atomic or molecular level which explains observed failure modes[23]. Sections 2 and 3 of this Notebook are directed toward collecting and organizing research efforts in the area of aging and failure mechanisms. As the state of the art in this area is advanced, it is reasonable to expect that screening criteria will be developed based on a much more fundamental class of parameter measurements. This topic is further probed in Section 5.5.

5.5 Future Role of Screening in Reliability Physics

Application of the screening procedures described in the preceding section are essentially statistical in nature at the present state of the art in reliability. That is, in burn-in screening, tolerance limits are set on the screening parameters on the basis of expected parameter degradation as determined from the statistical analysis of life-test data. Similarly, in linear-discriminant screening, expected probabilities of misclassification are predicted in screening lots of unknown reliability based on statistical correlation between the new components and the component data from which the discriminant was derived.

Developments in reliability physics, however, suggest that future screening procedures will be based on models that (1) include fundamental physics parameters characterizing a component, as opposed to operating parameters and (2) derive their functional form from theoretical analysis of the aging processes acting in a component in contrast to statistical hypothesis of a function. The impact of these developments on component screening may be expected to yield procedures that are significantly more powerful and screening models that are applicable to a wide range of operating conditions.

5.5.1 Trends in Measured Parameters

Empirical investigations of failure mechanisms and their relation to the stresses acting on electronic components, and the development of mathematical models characterizing degradation processes, continually add to our understanding of how components fail. This fact, in conjunction with developments in instrumentation and allied test equipment, suggests that future application of the screening procedures described in Section 5.2 will be based on fundamental physical parameters as opposed to operational performance parameters characteristically used now.

To illustrate this trend, consider the case of a diode type whose failure mode is an increase in surface channel current after an appreciably long time, t, of stable operation under a given environmental stress. It must be concluded that the physical processes responsible for the increased surface channel current have been in operation from $t = 0$. Such a process could be oxidation of a surface, decreasing its resistance. During early life, the surface component of reverse current would have been negligible with respect to junction current and therefore would not have been identifiable from the measurement of operating parameters. However, by cooling the diode to an appropriately low temperature, the junction current, characterized by a higher activation energy than the surface

Downloaded from https://www.everyspec.com

channel current, could be "frozen out", and the surface channel component of the reverse current separately identified. Ideally, it would be possible by burn-in screening to measure the rate of increase of channel current during early life. Since the magnitude of the surface channel current in this ideal case is proportional to the oxide thickness, the rate of increase of the channel current would be proportional to the oxidation rate. A knowledge of the rate law of the oxidation process, based on an understanding of the physical and chemical properties of the diode, would permit a prediction of the continued rate of increase of surface channel current, as depicted in Figure 5.10. I_T represents the total reverse current under operating conditions. I_S represents the surface channel current that has been determined by operation of the diode during the burn-in period, t_o to t_1, to obey the rate law f(t). Substitution of the critical value, I_F, of the reverse current for I_S in the expression $I_S = f(t)$, yields t_F, the time to failure.

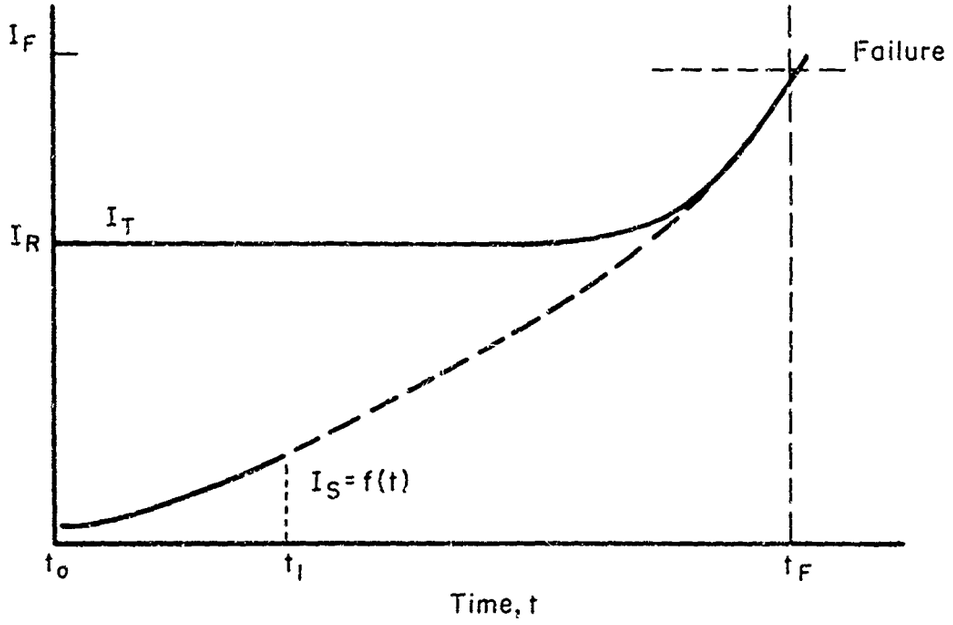


FIGURE 5.10. RATE OF CHANGE IN EARLY LIFE OF SURFACE CHANNEL CURRENT AS A PREDICTOR OF FAILURE

Assuming it can be measured by "freezing out" junction current.

It should be noted that this example represents an ideal case. In practice, such cases will be partially masked by "noise" and other complicating factors. However, it is the task of the screening models to sort out the "true" effects. Of prime importance here is that the example does illustrate specifically the kind of measurements that may ultimately be expected to yield powerful screening criteria.

5.5.2 Relation of Screening to Accelerated Testing

A basic problem in the development of screening procedures is to determine the set of stress conditions for which a given screening criterion is applicable. For example, a

Downloaded from https://www.everyspec.com

linear-discriminant screening criterion is applicable only for screening components whose intended application is for a period of time in a stress environment that is the same as that for the component life-test data from which the discriminant was derived. Thus, if one wishes to use the component type in another application, it is necessary to derive a new screening criterion.

Clearly, if models can be developed that relate component operating time and stress environment, it would be possible to obtain a screening criterion, derived under one set of conditions, that would be applicable over a wide range of conditions. This, however, is exactly the accelerated testing problem described in Section 4. In that section, theoretical models were presented that exhibit great potential in describing aging rates under various stresses in terms of basic materials properties and environmental interactions. Thus, when research in reliability physics can demonstrate the acceptability of these models, they can be directly applied in the development of screening models. Specifically, based on these results:

(1) Accelerated testing theory can serve as a basis for adapting screening functions derived for one operating environment to other environments.

(2) Accelerated testing theory can serve to identify the parameters to be included in a screening function that will yield maximum screening power.

Thus, it is seen that potential future developments in reliability physics tend to tie closely together the theory that underlies both screening and accelerated testing. The fundamental characteristic of these developments is that, in both instances, the trend is from purely statistical treatment of the problems to a more theoretically oriented consideration based on the physics of rate processes.

5.6 Physical Methods of Screening

The purpose of this subsection is to describe and characterize the particular types of tests that have been used to provide data inputs for the reliability screening analyses described earlier in this section. Information will be presented as to the actual device characteristics that are measured in the screening tests, how the measured characteristics are analyzed for failure-prediction purposes, the range of applicability of the tests, and the results obtained by workers who have used the tests. The information will be presented for resistors, capacitors, diodes, and transistors. Unusual equipment requirements of the tests will also be noted where appropriate.

It should be pointed out that many of these tests are applicable to only one device and that the efficiency of many of the tests has not been demonstrated. In some cases, there may be a physical basis for expecting a test to be useful for screening, but there are no statistical data available to support the expectation. This kind of information is provided where necessary.

Downloaded from https://www.everyspec.com

5.6.1 Resistors

5.6.1.1 Current or 1/f Noise

Curtis[6] has established a definite relationship between current noise in tin oxide film resistors and defects – blemishes, scratches, and bridges – in the resistor film. He also observed in performance tests that all resistors exhibiting abnormal TC or excessive change of resistance during 1000-hour load-life test were also relatively noisy. However, not all noisy resistors exhibited abnormal TC or excessive change of resistance.

Current noise is expressed quantitatively in terms of the noise index. It is measured using a circuit like that shown in Figure 5.11. Noise index is defined by Conrad as follows[24]:

$$\text{Noise Index} = 20 \log \frac{V_{rms}}{V} \text{ db in a decade,} \tag{5.27}$$

where

V_{rms} = number of microvolts of open-circuit rms current noise in a frequency decade

V = number of volts dc applied to the resistor.

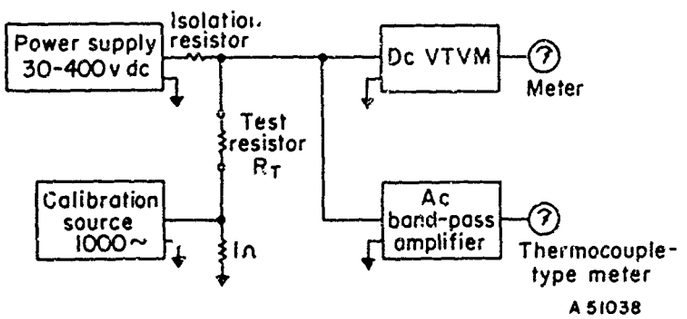


FIGURE 5.11. CIRCUIT FOR MEASURING RESISTOR NOISE

Determination of the noise index requires measurement of system noise, total noise, and d-c voltage. Noise index may then be calculated using the following formula:

$$\text{Noise Index} = D - f(T-S) - T \quad , \tag{5.28}$$

where

D = 20 log applied d-c voltage

T = total noise in db after voltage is applied

$f(T-S) = -10 \log \left(1 - 10^{-(T-S)/10}\right)$ is the correction factor for system noise.

Downloaded from https://www.everyspec.com

The paper by Conrad referred to above should be used as a guide in making this measurement. It develops the theoretical background of the measurement, gives preferred voltage levels for different size resistors, tabulates f(T-S), and presents details of equipment requirements.

5.6.1.2 Broadband Noise

Noise of an erratic nature that does not show any characteristic frequency dependence is known as "broadband noise". It has been associated with poor electrical contacts and with chemical processes occurring in a device. Because its occurrence is random, a difficulty arises in describing it quantitatively so that a numerical rejection criterion based on it may be stipulated.

Card and Mavretic[25] have described a system for detecting and recording this kind of noise. Although they were not able to correlate their results with device parameter drift, it is conceivable that correlation can be established with certain types of devices.

5.6.1.3 Parameter Drift

A method of resistor screening calling for determination of parameter drift may be desired. Resistor parameters that can be used for this type of screening are resistance, temperature coefficient of resistance, and voltage coefficient of resistance. The techniques used for performing these measurements are well known[26] and are very nearly the same for the different kinds of resistors.

In testing for parameter drift, it is usually necessary for the resistors to be under load and under environmental stress – high temperature or high humidity or both. The particular conditions desirable for a particular resistor must be determined for each resistor type.

5.6.1.4 Infrared Emission

Because the reliability of electronic parts is strongly affected by temperature, with high temperatures being detrimental to reliability, the presence of abnormally hot spots in a component can be an indication of early failure. All bodies emit electromagnetic radiation of intensity W according to the Stefan-Boltzmann law,

$$W = \sigma A \epsilon T^4 \quad , \tag{5.29}$$

where

σ = Stefan-Boltzmann constant, $5.67 \times 10^{-12} \frac{\text{watt}}{\text{cm}^2\ \text{deg}^4}$

A = area of the emitting surface

ϵ = emissivity of the emitting surface

T = Kelvin temperature.

Downloaded from https://www.everyspec.com

The wavelength distribution of radiant energy is given by Planck's law:

$$W_\lambda = 2\pi c^2 h \lambda^{-5}\left[\exp\left\{\frac{hc}{\lambda kT}\right\} - 1\right]^{-1} , \qquad (5.30)$$

where

W_λ = radiant flux per unit area per unit increment of wavelength

c = velocity of light, 2.99×10^8 m/sec

h = Planck's constant, 6.63×10^{-34} watt sec^2

k = Boltzmann constant, 1.38×10^{-23} joule/deg

λ = wavelength of radiation.

The value of λ for which W_λ is maximum, λ_m, is given by Wien's Displacement Law,

$$\lambda_m = \frac{w}{T} , \qquad (5.31)$$

where w = Wien's displacement constant, 2898 micron degrees.

Substitution into these equations of the temperatures at which electronic components operate will demonstrate that the radiation emitted by electronic components has maximum intensity in the range of 5 to 10 microns.

Since the intensity of this radiation varies as T^4, measurement of this radiation emitted by an electronic part is a very sensitive means of determining its temperature at closely spaced points over its entire surface. Such temperature data, when systematically analyzed, has shown promise of being an effective indicator of early failure or drift of electronic parts. Vanzetti has used an infrared scanning technique to detect differences in power dissipation of several hundredths of a watt in resistors.[13] Walker and Roschen have used infrared scanning to obtain temperature profiles of sputtered tantalum thin-film resistors on glass substrates and have correlated maximum resistor temperature with per cent change of resistance of four resistors.[12] Feduccia has used infrared scanning in tests performed on deposited Nichrome films. After obtaining infrared profiles on these resistors, he put them on life test. The majority of resistors that failed, 64 per cent, were characterized by a low intensity of IR emission. This number accounted for 17 per cent of all the resistors tested.[14]

It is evident that infrared emission data can be used to screen resistors. Standards of temperature profile need to be established for each type and configuration of resistor to be screened.

The apparatus required for infrared scanning consists of a radiation detector; an optical system for focusing the emission from a small point on the sample to the detector; a rotating slotted disk to periodically interrupt the radiation beam, thereby converting the desired signal from dc to ac; a scanning system; and a low-noise, narrow-band-pass

Downloaded from https://www.everyspec.com

amplifying and indicating system. The indication of signal may be film exposed in an infrared camera or a voltage. Vanzetti has found that the film indication is somewhat less sensitive to small changes in power dissipation of the part than is a voltage indication. Using the voltage method, Walker found that he could resolve temperature differences as low as 0.5°C[12]. He could also spacially resolve different components separated by as little as 0.0014 inch.

The references cited above give more complete details as to equipment and experimental techniques required for using infrared scanning for screening resistors.

5.6.2 Capacitors

5.6.2.1 Anodic and Ionic Reactions

Corl[15] has described a technique which purports to be useful for inspecting thin-film capacitors deposited on semiconductor substrates. The technique involves observation of irregularities in electrolytic reactions occurring at metal surfaces in contact with n- or p-type regions of semiconductor. One of its uses is the detection of flaws in the oxide dielectric of a metal-oxide-semiconductor capacitor.

The technique, basically, is as follows: A circuit containing the capacitor is immersed in an electrolyte which causes an anodic or cathodic reaction to take place at metal surfaces, depending on whether the metal is connected to an n-region or to a p-region of semiconductor. Corl used an acidic chromate electrolyte solution, and the metal surfaces were aluminum. The surface of the aluminum connected to the p-region turned brown, while the surface of the aluminum connected to the n-region remained bright. Thus, if the bottom plate of a capacitor was actually p-type silicon and the top electrode was connected to a region of n-type silicon, the top electrode should remain bright when the circuit was immersed. However, if a hole in the oxide permitted the top electrode to contact the p-type silicon underneath, the surface of the top electrode over the hole would indicate a p-type reaction. Such occurrences were demonstrated by Corl. The principal requirement for the electrolyte is that it react at connections to either but not both n- or p-type regions of semiconductor.

5.6.2.2 Hot D-C Leakage

A screening technique currently under consideration for use with solid tantalum electrolytic capacitors has been described by Howard[16]. The method is said to have exhibited a screening efficiency of 10.

In this method, the capacitors to be tested are placed in an oven at 125°C, and a voltage with a magnitude of 120 per cent of the rated voltage is applied. After 1/2 hour at this voltage and temperature, leakage current is measured. The voltage is removed, but the capacitors are left charged and remain in the oven for 5 days, after which leakage current is again measured. A device exhibiting a large change in leakage current between initial and final measurements is regarded as a potential early failure.

This procedure has not been used on other types of capacitors but may be applicable to thin-film tantalum capacitors.

Downloaded from https://www.everyspec.com

5.6.2.3 Parameter Drift

Parameters that may be measured in parameter-drift screening tests of capacitors are capacitance, dissipation factor, and, for electrolytic capacitors, leakage current. The testor must establish parameter criteria that are suitable for the device he is testing. It is likely that the efficiency of the tests will be enhanced if they call for the capacitors to be subjected to elevated temperature and high humidity.

Capacitance and dissipation factor are both measured with a-c bridges. There are many types of such bridges. The theories and operating characteristics of some of them are described by Michels[27]. Null-type bridges which must be manually balanced and automatic bridges with digital readout are commercially available. Accuracies of the order of 0.1 per cent are common for such instruments.

The techniques for measuring capacitance and dissipation factor will depend upon the types of instruments used. It is considered appropriate to note here that, in measuring capacitance, one must insure that distributed capacities in cables and connections are held to a minimum and are taken into account in the calculations. Distributed capacitance can be hundreds of picofarads. In a parameter-drift measurement, uncertainties of this magnitude may obscure significant changes.

Reverse leakage current of electrolytic capacitors may also be used for parameter-drift screening. Measurement of this parameter calls for a constant-voltage power supply capable of supplying the rated voltage of the capacitor and an ammeter capable of resolving at least 10^{-7} ampere. Since the leakage current of an electrolytic capacitor decreases for some time after voltage is applied, the time after application of voltage when the measurement is taken must be carefully controlled. Temperature at which this measurement is taken is very critical and must also be controlled.

5.6.3 Diodes

5.6.3.1 Reverse Current at Specified Voltage

Measurement of diode reverse current requires a constant-voltage power supply and an ammeter capable of measuring 10^{-10} ampere. Voltage should be increased slowly to the specified amount.

5.6.3.2 Reverse Breakdown Voltage

Reverse breakdown voltage can be determined on a transistor curve tracer oscilloscope or by making point-by-point measurements of the reverse current-voltage characteristics of the diode. Care should be taken that voltage and current levels used do not exceed rated values for the device.

If the breakdown characteristic is not sharp, i.e., shows a "soft knee", there may be difficulty in specifying the exact breakdown voltage. One solution to this problem is to

Downloaded from https://www.everyspec.com

choose a current above which the diode is not to operate and designate the reverse voltage at which that current is reached as the breakdown voltage.

The ammeter used in this measurement should have at least three decade ranges.

5.6.3.3 Open-Circuit Voltage Decay Time

Measurement of open-circuit voltage decay time affords a measure of the semiconductor minority-carrier lifetime in the diode. The essence of the measurement is to have a periodically interrupted forward current in the diode. Voltage across the diode is displayed on an oscilloscope. The voltage trace has three stages. The first stage is a horizontal trace corresponding to the part of the cycle when a steady current flows in the diode. The second stage is a vertical trace occurring at the instant the current supply is interrupted. This corresponds to the instantaneous drop in voltage across the ohmic component of the diode resistance. The third stage is an exponentially decreasing curve terminating at $V = 0$. This part of the curve is a result of the residue of minority carriers injected into the diode base by a forward current across the p-n junction.

The point on the time scale at which the curve crosses the $V = 0$ line is the parameter of interest in this measurement. Changes in this parameter are meaningful only if the forward current level is kept the same each time the measurement is made.

5.6.3.4 Reverse Pulse Recovery Time

When a properly contacted p-n junction is forward biased, charge carriers cross the junction and cause a higher than equilibrium carrier concentration to exist in the semiconductor material. If the forward bias is removed, the carrier concentration decays at a rate inversely proportional to the minority-carrier lifetime of the carriers – holes or electrons – in this particular material. If a reverse pulse of sufficient amplitude is superimposed on the forward bias, there will be a high reverse current which decays exponentially to the reverse saturation current. This high reverse current is the manifestation of the flow of the large, nonequilibrium, concentration of charge carriers in the diode under the influence of the reverse pulse. As these carriers recombine, this reverse current decays proportionately. The time required from application of reverse pulse until current begins to decay, t_f, is an indicator of charge carrier lifetime, τ, in the material. The value of the charge-carrier lifetime may be obtained from the delay time, t_f, and the ratio of forward to reverse steady-state currents from the following expression[27]:

$$\operatorname{erf}(t_f/\tau)^{1/2} = \frac{1}{1 + (I_f/I_R)} \quad . \tag{5.32}$$

This technique can be used to measure lifetimes as low as 10^{-8} second. A pulse generator and oscilloscope with a fast rise time (10^{-9} second) are required for making this measurement. For the purpose of screening, however, it may not be necessary to calculate the carrier lifetime, τ. A measurement of the time lapse, t_f, before the reverse current begins to decay may be sufficient.

Downloaded from https://www.everyspec.com

5.6.3.5 Slope of Log Reverse Current Versus Reciprocal Temperature Curve

The slope of a plot of log reverse current versus reciprocal temperature is linear for the temperature range over which a given electron activation process predominates. Such a plot can be used to calculate the activation energy of the charge carriers that contribute to reverse current over this temperature range. If the plot is made over a sufficiently wide temperature range, more than one linear portion of the curve may be seen, indicating the presence of additional charge-carrier energy states. Both surface and bulk energy states may be detected by this technique. In diodes where the space-charge region at the surface is less than a diffusion length, surface currents are usually ohmic and the associated activation energies are in the range of 0.01 to 0.1 electron-volt. Where the space-charge region is much wider than a diffusion length, surface currents are of the generation-recombination type, and activation energies are approximately the same as for bulk states, i.e., 0.1 to 0.5 electron-volt. Changes in these states over a period of time can be caused by any one of several deleterious mechanisms in the material.

The measurement is made by applying 1 volt reverse bias to the diode and monitoring the current over the rated temperature range. A thermocouple may be used to measure diode temperature, but care must be taken to assure adequate thermal contact to reduce errors in temperature readings due to thermal lag across high thermal resistances. Temperature and current readings may be taken at 1 to 2-degree temperature intervals. In silicon junction devices made from high-resistance material, the current range may be as much as 10 orders of magnitude over a temperature interval from -65°C to 200°C.

It should be pointed out that, if the rated temperature range is exceeded, the performance of this measurement may impose thermal stresses on the diode that may in themselves have deleterious effects on diode properties.

5.6.3.6 Avalanche Noise

A diode biased in the region of the reverse characteristic where direct current is beginning to increase sharply will exhibit a pulsating dc, called avalanche noise. This noise signal will rise, reach a peak, and decrease to the ambient noise level as the d-c bias is increased. The entire noise envelope may be traversed as the d-c voltage is changed by as little as 1 volt. As the dc is increased further, additional noise envelopes may be detected.

The sources of these noise signals are commonly known as microplasmas. They are actually current pulses of constant amplitude which turn off and on with very high frequency. In the initial part of a given envelope, the ratio of on-to-off time of the pulses is much less than one. As d-c bias is increased, the ratio of on-to-off time increases toward infinity as the microplasma becomes continuously conducting. The a-c signal detected on an electrodynamic a-c meter is a maximum when the ratio of on-to-off time is unity. As on-time begins to exceed off-time, the detected a-c signal decreases, becoming zero as the microplasma current becomes continuous. As d-c bias is increased further, new microplasmas may be formed, and the process described above is repeated.

The microplasma or avalanche noise current may be measured by connecting an a-c voltmeter across a precision resistor in series with the diode. The peak noise currents are of the order of 5×10^{-5} ampere, so the size of series resistor will be determined by the range of the a-c detection equipment.

Downloaded from https://www.everyspec.com

The voltage at which the noise envelope occurs may be used as a screening criterion. The appearance of a noise envelope at a voltage lower than the rated voltage would indicate the existence of avalanche conduction, which is normally accompanied by excessive joule heating.

5.6.3.7 Electron-Beam Scan

Two techniques for performing extremely detailed inspections of microcircuits by means of a scanning electron beam have been described by Everhart, et al[29]. In the first, a low-energy beam impinges on the surface of material being inspected. As a result, secondary electrons are emitted by the surface and collected by a collector. The collector current is used to modulate the intensity of a CRT display, which is scanning in synchronization. The number of electrons reaching the collector, and hence brightness of the CRT display, depends on the potential of the emitting surface. An area 0.1 micron in diameter may be inspected this way. Thus, variations in potential are observed visually in minute detail. A second method is useful for examining p-n junctions that are buried in the crystal or obscured by thick metallic overlays. In this method, a photovoltage generated in the p-n junction by the electron beam is measured and observed on an oscillioscope. P-n junctions can be delineated in this way.

The electron-beam scan has been used to detect conditions on a microcircuit that may indicate poor reliability. Such conditions are lateral diffusion of impurities, diffusion flaws, uneven thickness or breaks in conductive layers, undercutting of oxide layers by etchants following photomasking, and induced inversion layers.[30]

The scanning electron beam makes possible a visual examination of the electrical performance of a circuit in fine detail without seriously perturbing the circuit. On the other hand, the equipment required is quite elaborate, and the sample being inspected must be placed in a high vacuum.

A scanning electron beam is also used in electron-probe microanalysis. In this technique, electrons accelerated through a potential of 25 to 35 kv impinge on the target, exciting X-rays characteristic of the target material. The chemical species and surface concentrations of the target material are determined from the wavelengths and intensities of the emitted X-ray spectrum. The area analyzed by the beam is typically 2.5 to 4.5 microns in diameter. Spatial resolution depends on surface roughness and focusing of the beam. Typically, interfaces between different elements can be delineated within several microns.

5.6.3.8 Infrared

Observation of infrared emission intensity, discussed previously, should be applicable as a screening technique for diodes. The exact manner in which infrared could be used would have to be determined by experience with the part to be screened.

Downloaded from https://www.everyspec.com

5.6.4 Transistors

5.6.4.1 Collector-Base Leakage Current, I_{CBO}

Collector-base leakage current is measured by reverse biasing the collector-base junction and leaving the emitter terminal open. In a PNP transistor, the base lead should be positive and the collector lead negative. In an NPN transistor, the base should be negative and the collector positive.

Voltage across the collector-base junction should be measured with a voltmeter whose impedance is high compared with that of a reverse biased p-n junction. Currents are typically of the order of 10^{-9} ampere, and current-measuring equipment should therefore be capable of resolving 10^{-11} ampere.

The voltage at which the value of I_{CBO} is taken should, of course, be specified.

5.6.4.2 Collector-Base Breakdown Voltage, BV_{CBO}

This measurement is also made with the transistor treated as a diode, i.e., the emitter terminal is open. The measurement is made as it is for a diode; see Section 5.6.3.2.

5.6.4.3 Direct-Current Gain, h_{FE}

Direct current gain, h_{FE}, is defined as the ratio of collector current, I_C, to base current, I_B, in a transistor connected in the common emitter configuration with no a-c signal applied. (It should not be confused with small signal gain, h_{fe}, or B_s.) A specification of h_{FE} is meaningful only if collector-current and collector-emitter voltage, V_{CE}, are also specified. A high collector leakage current will render the determined value of h_{FE} meaningless.

Determination of h_{FE} involves measurement of I_C, I_B, and V_{CE}. These measurements may be made by connecting meters at appropriate points in the circuit. The General Electric Transistor Manual[32] describes a circuit that provides for direct measurement of h_{FE} by potentiometric determination of the ratio of I_C to I_B.

5.6.4.4 Collector-Emitter Saturation Voltage, $V_{CE(sat)}$

Collector-emitter saturation voltage is the voltage from collector to emitter when the collector is in saturation. It is usually less than a volt and should be measured on a high-impedance voltmeter capable of resolving millivolts. In order for the specified value of $V_{CE(sat)}$ to be meaningful, the appropriate values of I_C and I_B should be specified.

Downloaded from https://www.everyspec.com

5.6.4.5 Transistor Noise Measurement

Transistor noise is a complex subject that will not be discussed here. There are three significant noise parameters that can be measured:

(1) Noise figure, $NF = 10 \log \left\{ \frac{\left(\frac{\text{Signal power}}{\text{Noise power}}\right)_{in}}{\left(\frac{\text{Signal power}}{\text{Noise power}}\right)_{out}} \right\}$

(2) Equivalent noise voltage, $\left(\overline{e_n^2}\right)^{1/2}$

(3) Equivalent noise current, $\left(\overline{i_n^2}\right)^{1/2}$

Techniques for measuring these quantities are described in References (29), (30), and (31).

5.6.4.6 Infrared Scanning

The measurement of infrared emission intensity was used by Vanzetti[13] to screen transistors. For a description of the technique, see Section 5.6.1.4.

5.6.4.7 Electron-Beam Scanning

The electron-beam scan is potentially useful for inspecting transistors, particularly for discovery of diffusion flaws and imperfections in the passivating oxide layer. For a description of this technique, see Section 5.6.3.7.

REFERENCES

(1) Bratt, M. J., Reethof, G., and Weber, G. W., "A Model for Time Varying and Interference Stress-Strength Probability Density Distributions With Consideration for Failure Incidence and Property Degradation", Proceedings, SAE-ASME-AIAA Aerospace Reliability and Maintainability Conference (July, 1964).

(2) Kececioglu, Dimitri, and Cormier, David, "Designing a Specified Reliability Directly Into a Component", Proceedings, SAE-ASME-AIAA Aerospace Reliability and Maintainability Conference (July, 1964).

(3) Thomas, R. E., "Component-Part Screening Procedures Based on Multiparameter Measurements", IRE Transactions of the Professional Group on Component Parts, Vol. CP-6, No. 4.

(4) Thomas, R. E., "Reliability Screening of Electronic Component Parts", Technical Report No. 4, Contract NONR-2864(00), Project NR 042-210 (October, 1960).

Downloaded from https://www.everyspec.com

(5) Fisher, R. A., "The Use of Multiple Measurements in Taxonomic Problems", Contributions to Mathematical Statistics, Wiley (1950), Chapter 32.

(6) Curtis, J. G., "Current Noise Measurement as a Failure Analysis Tool for Film Resistors", Physics of Failure in Electronics, 1, edited by Goldberg and Vaccaro, Spartan Books, Inc., Baltimore, Maryland, 204-213 (1963).

(7) Best, G. E., Bretts, G. R., McLean, H. T., and Lampert, H. M., "The Determination and Application of Aging Mechanisms Data in Accelerated Testing of Selected Semiconductors, Capacitors, and Resistors", Physics of Failure in Electronics, 3, to be published.

(8) Smith, P. C., and Genser, M., "An Investigation of Thin-Film Resistor Failure", Physics of Failure in Electronics, 3, to be published.

(9) Klein, N., Gafni, H., and David, H. J., "Mechanisms of D-C Electrical Breakdown in Thin Silicon Oxide Films", Physics of Failure in Electronics, 3, to be published.

(10) "Study of Failure Mechanisms at Surfaces and Interfaces", Final Report, RADC-TDR-63-152, on AF 30(602)-2593 by Motorola, Inc., Solid State Systems Division, Phoenix, Arizona, for Rome Air Development Center, Griffiss Air Force Base, New York.

(11) "Failure Mechanisms at Metal-Dielectric Interfaces", Third Technical Documentary Report, Contract AF 30(602)-3266, by Motorola, Inc., for Rome Air Development Center, Griffiss Air Force Base, New York.

(12) Walker, M., Roschen, J., and Schlegel, E., "An Infrared Scanning Technique for the Determination of Temperature Profiles in Microcircuits", IEEE Trans. on Electron Devices, ED-10 (4), 263 (1963).

(13) Vanzetti, R., "Infrared Techniques Enhance Electronic Reliability", Solid State Design, 29-37 (August, 1963).

(14) Feduccia, A. J., "Reliability Screening Using Infrared Radiation", 1964 IEEE International Convention Record, 12, Pt. 9, 18-23.

(15) Cori, E., "Anion Reaction for Failure Analysis of Microcircuit Components", Physics of Failure in Electronics, 3, to be published.

(16) Howard, Lowell, P. R. Mallory and Company, Indianapolis, Indiana, Private Communication. Information obtained on Contract AF 30(602)-3291, "Nondestructive Reliability Screening of Tantalum Capacitors", Final Report to be published in June, 1965, as an RADC Technical Documentary Report.

(17) Bevington, J. R., and Ingle, L. V., "Nondestructive Reliability Screening of Electronic Parts", Final Report, RADC-TDR-64-311, on AF 30(602)-2972, 4-39 (September, 1964).

Downloaded from https://www.everyspec.com

(18) Sah, C. T., "Effect of Surface Recombination and Channel on P-N Junction and Transistor Characteristics", IRE Trans. on Electron Devices, ED-9 (1), 94-108 (January, 1962).

(19) Gorton, H. C., Zacaroli, A. R., Reid, F. J., and Peet, C. S., "Preparation and Properties of Grown P-N Junctions of InSb", J. Electron. Chem. Soc., 108 (4) (April, 1961).

(20) Thomas, R. E., and Gorton, H. C., "Research Toward a Physics of Aging of Electronic Component Parts", Physics of Failure in Electronics, 2, edited by Goldberg and Vaccaro, Cato Show Printing Company, 25-60 (1964).

(21) Von Alven, W. H., Semiconductor Reliability, 2, Engineering Publishers (1962).

(22) Bevington, J. R., and Ingle, L. V., "Non-Destructive Reliability Screening of Electronic Parts", Technical Documentary Report No. RADC-TDR-64-311.

(23) Vaccaro, J., "Reliability Physics at RADC", Physics of Failure in Electronics, 3, to be published.

(24) Conrad, G. J., Newman, N., and Stansbury, A. P., "Recommended Standard Resistor-Noise Test System", IRE Transactions of the Professional Group on Component Parts, Vol. CP-7, 71 (September, 1960).

(25) Card, W. H., and Mavretic, A., "Burst Noise in Semiconductor Devices", Physics of Failure in Electronics, 2, 268 (1963).

(26) Wellard, C. L., Resistance and Resistors, McGraw-Hill Book Company, Inc., New York (1960), Chapters 3 and 6.

(27) Michels, W. C., "Electrical Measurements and Their Applications", D. Van Nostrand and Co., Inc., Princeton, New Jersey (1957), Chapter 5.

(28) Lax, B., and Neustadter, S. F., "Transient Response of a p-n Junction", J. Appl. Phys., 25 (9), 1148-1154 (September, 1954).

(29) "IRE Standards on Methods of Measuring Noise in Linear Two Parts, 1959. Standard 59 IRE 20. S1", Proc. IRE, 48, 60 (1960).

(30) Hanson, G. H., and Van Der Ziel, A., "Shot Noise in Transistors", Proc. IRE, 45, 1538 (1957).

(31) Transistor Manual, Seventh Edition, General Electric Company, 499-507 (1964).

Downloaded from https://www.everyspec.com

SECTION 6. THE USE OF STATISTICAL METHODS IN RELIABILITY PHYSICS EXPERIMENTS

Downloaded from https://www.everyspec.com

TABLE OF CONTENTS

	Page
6. The Use of Statistical Methods in Reliability Physics Experiments	6-1
6.1 Introduction.	6-1
6.2 Use (and Misuse) of Averages.	6-1
6.2.1 Conclusions	6-4
6.3 The Problem of Statistical Inference	6-4
6.4 Statistical Techniques for Analyzing the Functional Relation Between Two Variables.	6-6
6.4.1 Techniques for Testing for Significant Differences.	6-8
6.4.1.1 Analysis of Variance	6-8
6.4.1.2 Techniques Dealing With Medians and Proportions	6-10
6.4.1.3 The Question of Assumptions.	6-10
6.4.2 Techniques of Curve Fitting	6-12
6.4.2.1 The Fitting and Testing of a Linear Equation	6-12
6.4.2.1.1 Fitting the Equation.	6-13
6.4.2.1.2 Testing "Goodness of Fit"	6-14
6.4.2.2 The Fitting and Testing of a Nonlinear Equation: The Arrhenius Model	6-15
6.4.2.2.1 Fitting the Equation.	6-16
6.4.2.2.2 Testing "Goodness of Fit"	6-16
6.4.2.3 Fitting and Testing Other Nonlinear Functions	6-20
6.4.2.4 The Accuracy of Constants in a Fitted Equation	6-21
6.5 Extensions to Multi-Variate Analysis	6-21
6.5.1 Analysis of Variance	6-22
6.5.1.1 Two Independent Variables	6-22
6.5.1.2 More Than Two Independent Variables	6-26
6.5.1.3 Fractional Replications Design	6-26

Downloaded from https://www.everyspec.com

TABLE OF CONTENTS
(Continued)

Page

6.5.2 Surface Fitting 6-29

6.5.2.1 Fitting an Equation 6-29

6.5.2.2 Testing Goodness of Fit 6-30

6.5.2.3 The Accuracy of Constants in a Fitted Equation 6-31

6.6 Error Reduction 6-31

6.6.1 Error Reduction by Experimental Control 6-31

6.6.2 Error Reduction by Statistical Control. 6-32

REFERENCES 6-33

LIST OF FIGURES

Figure 6.1. Time Histories for ICES at 10 Volts 6-2

Figure 6.2. Mean Time Histories for Three Environmental Stress Levels . . . 6-5

Figure 6.3. Mean Time History for ICBO at 10 Volts (μ Amperes) 6-8

Figure 6.4. Mean Time History for a Parameter and Line of Best Fit 6-13

Figure 6.5. Arrhenius Equation Fitted to Example Data (Single Solution) . . . 6-17

Figure 6.6. Arrhenius Plot for Example Data (Single Solution). 6-17

Figure 6.7. Arrhenius Equation Fitted to Entire Temperature Range and to Two Separate Subranges of Temperature. 6-19

Figure 6.8. Arrhenius Plots for Example Data (All Three Solutions). 6-19

Figure 6.9. Graphs of Data From Table 6.4 6-23

LIST OF TABLES

Table 6.1. Use of Averages When Components are Dropped From Test. . . . 6-3

Table 6.2. Example Data: Degradation Rates for Six Levels of Junction Temperature 6-15

Downloaded from https://www.everyspec.com

LIST OF TABLES
(Continued)

		Page
Table 6.3.	Some Useful Nonlinear Equations.	6-20
Table 6.4.	Average Rate of Change (for a Given Parameter) for Various Combinations of Ambient Temperature and Collector Voltage . . .	6-22
Table 6.5.	Combinations of Parameters Investigated	6-28
Table 6.6.	An Experimental Design to Control for Lot Differences	6-32

Downloaded from https://www.everyspec.com

6. The Use of Statistical Methods in Reliability Physics Experiments

6.1 Introduction

The objective of this section is to list and describe classical statistical techniques useful to the reliability physicist in analyzing data generated from experiments. The emphasis is on the research questions that can be answered by judicious application of such techniques, the information these techniques provide, and the practical application for such techniques. Thus, this section is almost exclusively output oriented – how to apply the various techniques is not covered.

No formal statistical training is assumed on the part of the reader. Therefore, only those statistical terms, concepts, and definitions are given that are required to understand the information provided by the various techniques. Inasmuch as possible, such concepts and definitions are given in nontechnical language.

The emphasis is on analysis of data generated from experiments, rather than on questions concerned with designing, planning, and conducting experiments. The type of data considered for analysis is almost exclusively the change and the rates of change of continuous parameter values, as functions of time and environmental stress conditions, rather than failure-rate data. Further, emphasis is on analysis of the functional relation between one independent and one dependent variable. Extensions and application of techniques and concepts to the case of several independent variables is shown in less detail. Liberal use is made of examples, and hypothetical (but realistic) data are used.

With the above orientation, it is believed that the reliability physicist will be in a better position to judge the validity of the use of statistical methodology as a tool for providing information on relevant research issues; he will also be in a better position to work and communicate effectively with the statistician in analyzing problems of a statistical nature.

All the statistical techniques described here deal with averages. The over-all organization and content of this section is as follows. First, a critical discussion on the use of averages is given. This is followed by a definition and discussion of the general problem of statistical inference in a reliability physics context. Next, various statistical inference techniques for analyzing the functional relation between two variables are given, including various techniques for testing statistical significance, curve fitting, and related topics. In the part dealing with curve fitting, special attention is given to the Arrhenius equation because of its importance in reliability physics and because, by using it as an example, several concepts and techniques of curve fitting can be illustrated. The next major part deals with multivariant analysis, including analysis of variance and the topic of surface fitting. Finally, a brief discussion of error reduction and associated experimental designs is given.

6.2 Use (and Misuse) of Averages

Averages, of one sort or another, are frequently found useful in summarizing and describing a large body of data. However, they are frequently used indiscriminantly,

Downloaded from https://www.everyspec.com

leading to erroneous conclusions. The usefulness of statistical techniques involving averages can be no greater than the validity of the averages concerned.

Consequently, some attention is given below to the topic of "averaging". Since the main problems of data averaging in reliability physics are associated with time-dependent data, the case is treated in which the investigator wishes to use averages to generally summarize and describe how the behavior of individual components changes over time.

Consider Figure 6.1, which presents some hypothetical data. Four time histories for each of four individual transistors are shown. The parameter of concern, ICES at 10 v, is measured at the indicated operating times. The mean time history is also shown, i.e., the mean parameter value* for each operating time that measurements were made.

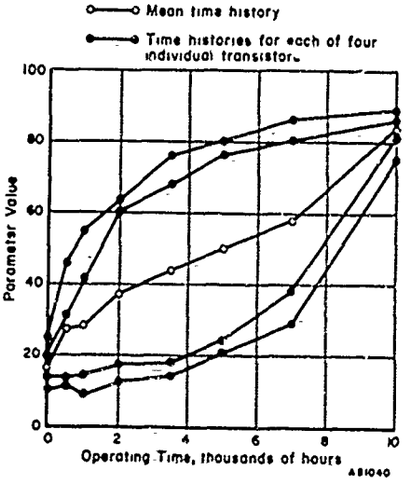


FIGURE 6.1. TIME HISTORIES FOR ICES AT 10 VOLTS

It can be seen that the shape of the mean curve does not represent very well the shape of any individual curve – the top group of two curves have a decreasing slope and the bottom group of two curves exhibit an increasing slope, whereas the mean curve gives the impression of a roughly linear trend. In this sense, the mean curve is misleading (as would also be a median or any other average curve). Further, if the investigator were not to examine each individual curve, he would not discover these two groups of components, which could reflect two separate aging mechanisms. In short, examination, analysis, and interpretation of only a mean curve for all the data may not only be misleading, but may also result in loss of important information contained in the data. In this example, the investigator might well plot two average curves, one for each group of components, and perform separate analyses for each group.

*The arithmetic mean of a set of measurements is defined as the sum of these measurements divided by the number of measurements involved.

Downloaded from https://www.everyspec.com

In general, where components show large individual differences in their aging behavior (not uncommon in reliability studies), and where these differences cannot be largely attributable to errors in measuring devices and various uncontrolled environmental disturbances, the investigator can legitimately question the use of any over-all average parameter curve. In any event, there can be no substitute for examining each individual time history to see whether an average curve can generally represent all of the individual curves and, if it cannot, to see the possible "groupings" of individual curves. In this connection, where there is a great deal of data, such as several individual time histories for each of several environmental stress conditions, computer techniques can be used to advantage. A digital computer can be used to plot individual component time histories and to compute and plot average curves for different homogeneous groups of time histories.

The problem of averaging parameter values becomes more difficult when some of the components in an aging study are dropped from test after differing amounts of operating time. The problems that arise in this situation are illustrated in Table 6.1, which shows, for three operating times (t_1, t_2, and t_3), the measured values of some parameter for each of seven components. At t_2, Components 1, 2, and 3 were dropped from test, and thus no parameter values for these components exist for t_2 and t_3. At t_3, Component 4 was dropped from test.

TABLE 6.1. USE OF AVERAGES WHEN COMPONENTS ARE DROPPED FROM TEST

Component	t_1	t_2	t_3	
1	15	--	--	Parameter values for each component
2	12	--	--	
3	10	--	--	
4	6	8	--	
5	4	5	5	
6	3	4	4	
7	2	4	4	
	7.4	5.2	4.3	← Mean parameter values

In this situation, in order to describe "average" parameter variation over time, one might simply average using the available data. The mean parameter values, computed by using the available data for each component, are shown. Such averages can be highly misleading. From t_1 to t_2, the mean parameter values show a decrease, whereas in reality, each single component remaining on test has shown an increase in the parameter value. From t_2 to t_3, no single component has shown a decrease in the parameter value, yet the mean indicates that, "on the average", the parameter decreases. The reason for this distortion is because the components dropped from test manifested

Downloaded from https://www.everyspec.com

relatively high parameter values for the previous measurement. In any case, when components are dropped from test, the use of any over-all average parameter curve can lead to grossly false conclusions. One solution, apart from using a "percentage survival" measure for the entire group as a function of time, is to analyze each group of components separately – the group from which no components are dropped from test and the group from which components are dropped from test. For the former group, a mean parameter curve can be computed, but for the latter group only a percentage survival measure would be meaningful.

6.2.1 Conclusions

Averages of any sort, and especially averages dealing with aging behavior, can be misleading, and preoccupation with such averages may not only lead to erroneous conclusions, but also result in loss of information as well. The investigator should carefully examine the averages he computes in summarizing and describing his data, and he should carefully consider just what it is that is being averaged together. The problem is not so much whether one should use averages, but more a question of how one should average, and especially what data should and what data should not be averaged together in order to provide results that are meaningful and useful from an engineering-physics point of view.

The technique of grouping components that exhibit "similar" aging behavior can probably be used to advantage in analysis of aging-behavior data from many reliability physics experiments. Averages can then be validly utilized for each of the groups thus formed. Computer techniques can be advantageously used in formulation of groups and associated calculations.

The techniques presented here are used to draw conclusions and make inferences about averages, and such techniques are valid and useful only to the extent that the averages themselves are valid and useful.

6.3 The Problem of Statistical Inference

Statistical inference is the problem of inferring characteristics of a population from information contained in a sample of elements from that population. Sample results are usually considered to be of limited generality, meaning, and usefulness, and the interest is usually in what is true in the larger population. Thus, for example, the problem of opinion polling is a problem in statistical inference. The pollster attempts to achieve a sample of individuals that yield opinions representative of the opinions of the public at large (the population). Sometimes serious errors are made, but most often sample statistics, e.g., per cent of people in the sample for or against a given issue, correspond quite closely to population values of the statistics.

In experimental research in reliability physics, the problem of statistical inference is much more complex. For example, the sample size may be quite small because of cost and other practical considerations; components may differ widely with respect to the measures of interest; instruments designed to measure parameters of interest may be of limited accuracy, introducing variable measurement error into the data; and different variable and uncontrolled environmental stress disturbances may serve to distort the

Downloaded from https://www.everyspec.com

sample results. As a result of all these factors, sample averages (e.g., mean-time histories or average degradation rates) based on a limited number of sample observations may deviate considerably from the corresponding averages that would be achieved in a much larger population of components. The problem, then, is to make inferences about such population averages, or "true" values as they sometimes are called, from information contained in limited and fallible sample data. Such population results are of greater generality and, as compared with sample results, are relatively free of variable error that results from limited measuring devices and uncontrolled environmental disturbances, since such error tends to "average out" over a large population of observations. Population values are analogous to the "signal" and sample values represent "signal" plus "noise"; the problem is to separate the two.

The problem is illustrated in Figure 6.2, which shows the type of experimental data that will be of major concern in this section. Mean time histories are shown for three

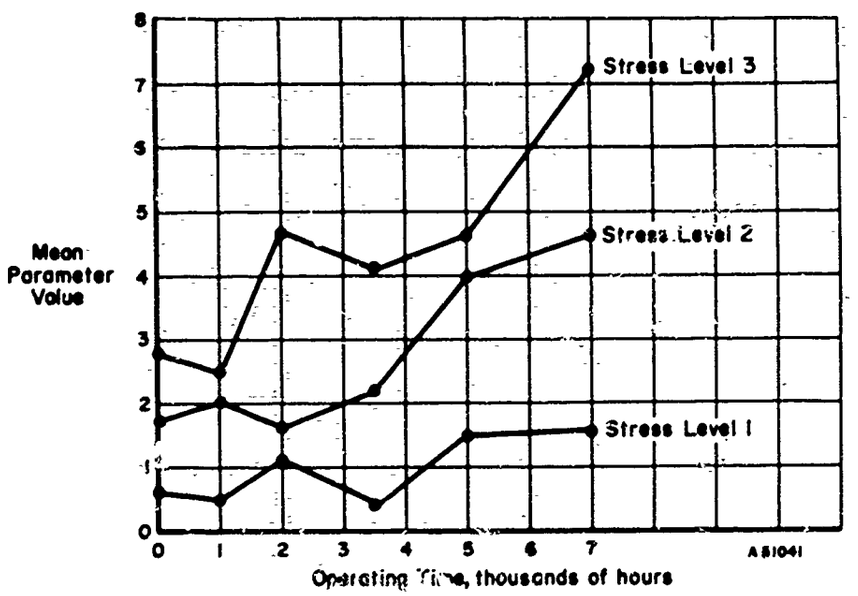


FIGURE 6.2. MEAN TIME HISTORIES FOR THREE ENVIRONMENTAL STRESS LEVELS

levels of some environmental stress, e.g., ambient temperature, for some parameter of concern. The data are for a group of 15 components, 5 components under each stress level. Each group of 5 components represents a sample* from a very much larger group (or population) of components of the same type made by the same manufacturing process. The problem is to make inferences about the aging behavior of the larger population of components, with knowledge only of the sample results. That is, if the entire population of components could be observed under each stress condition, instead of only 5 components, what would the mean time histories look like? What conclusions can be drawn

*It is assumed that each group of components is a random sample. The concept of random sampling is discussed subsequently.

Downloaded from https://www.everyspec.com

about the way they would change as a function of operating time? For example, is it reasonable to believe that the "true" (population) means are actually linear over time and that the only reason the observed means do not fall on a straight line is because of error, i.e., that the observed nonlinearity is a result only of the characteristics of the particular group of components that happen to be in the sample (sampling error) and/or other variable error? If it is not reasonable to believe that the relationship in the population is linear, might the relationship conform to some smooth exponential function, the observed irregularities in the data simply being a function of one or more sources of error? For Stress Level 1, can the investigator confidently conclude that there is any relation at all in the population, or might the observed "trend" in the means be only a function of error, the "true" means not differing at all among themselves? Finally, if it can be concluded that there is a real relationship, can this relationship be assumed to be an increasing one and any reversals only a function of error?

The problem increases in complexity when not only aging effects under a given stress condition are being considered, but when the effects of various stress levels on aging behavior are also of concern, e.g., the effect of stress on rate of change of the parameter. In the case when a different sample of components is used under each stress condition, the observed effects of stress may be considerably in error, simply because the small sample of components under one stress condition "happens" to be different than the sample of components in another stress condition -- i.e., the effects of stress may be confused with individual differences between components. However, if a very large number (population) of components were to be observed under each stress condition, such a contaminating effect would be equalized from one stress condition to another. So, the general question is still the same: If the entire population of components could be observed under each stress condition, what would be the effects of stress on aging behavior? What can be concluded about such conceptual population (true) results with only knowledge of the failable sample data?

There is a wide variety of statistical inference techniques designed to answer questions such as those given above, some of which are described in the remaining parts of this section. They all have the same general objective -- to infer something about population (true) results from sample data which is only partly true and partly error (sampling error, plus measurement error and other error). Since such error is "noisy" (random-like), its magnitude varying from sample to sample and its exact amount never known in any given experiment, these techniques necessarily make inferences of a probabilistic nature by taking into account the "likely" amount or range of error that can be expected in a given situation and for a given experiment.

6.4 Statistical Techniques for Analyzing the Functional Relation Between Two Variables

By the functional relation between two variables, x and y, one variable x considered independent and the other variable y considered dependent, is meant the relation between x and some average value of y -- for example, the mean or median value of y. The investigator is interested in the nature of this relation. For example, the investigator may be interested in how (if at all) the mean value of ICBO at 10 volts for a particular transistor type changes as a function of time, or he may be interested how (if at all) the median time to failure changes as a function of junction temperature.

Downloaded from https://www.everyspec.com

The techniques given below are used to make inferences about the "true" functional relation – that is, the functional relation that obtains in the population by taking into account information contained in the sample data. Two major classes of techniques for making such inferences are of most use to the reliability physicist: (1) techniques of testing for significant differences and (2) techniques of curve fitting.

For the techniques of testing for significant differences, the first concern is whether or not there is any relation at all in the population between two variables – i.e., whether or not the investigator can confidently conclude that the true average value of some dependent variable changes at all as a function of some independent variable. If the decision is that there is a real relation, then this class of techniques can also be used to provide information of where the function increases or decreases. Thus, this class of techniques is maximally useful when the investigator cannot confidently assume that there is any relation between the two variables, or when he cannot confidently judge the trend of the data over the range of the independent variable being studied. Thus, this first class of techniques is maximally useful where little is known about a particular relation and where the study and analysis of the relationship is in the beginning stages. This class of techniques includes analysis of variance methods and associated techniques, such as the t model, and also includes various tests dealing with medians and proportions.

It is important to note that, by use of this first class of techniques, no information is provided on how, or how much, the average value of a dependent variable changes as a function of some independent variable, apart from inferences about whether there is any relation at all in the population and apart from possible ordinal information obtained. If the investigator is interested in making more specific inferences about the nature of the functional relation, then the second class of techniques (techniques of curve fitting) is to be employed. More specifically, curve fitting, as used in this section, concerns itself with the mathematical form of the functional relation between two variables. The concern is with the form of the equation(s) that might be used to describe the way in which an average value of some dependent variable in the population changes as a function of some independent variable and also the estimation of the values of constants in any such equation. The values of these constants may have either theoretical meaning, or may only be used empirically in an equation to provide a prediction of the "true" (population) average for any given value of the independent variable and thus provide an estimate of how (and how much) the true value of the dependent variable changes as a function of the independent variable. In any case, for curve fitting, it is assumed that the two variables in question are truly related, this assumption being supported by applying a test of significance to the data, or by previous studies and engineering knowledge, or both.

In the part dealing with curve-fitting methods, least-squares techniques are presented for fitting to a body of data both linear and nonlinear functions. Techniques for testing "goodness of fit" of a fitted function are also discussed, including analysis-of-variance techniques. For nonlinear functions, the fitting and testing goodness of fit of the Arrhenius equation is used as an example. Finally, the topic of accuracy of constants in a fitted equation is briefly discussed.

Downloaded from https://www.everyspec.com

6.4.1 Techniques for Testing for Significant Differences

6.4.1.1 Analysis of Variance

Analysis of variance and associated techniques, such as the t model, are the major techniques used in testing for significant differences. To illustrate the use of the analysis of variance, consider the data shown in Figure 6.3*. This figure shows the mean value of ICBO at 10 volts for five operating times of 0, 1000 hours, 5000 hours, 10,000 hours, and 12,000 hours. The data are for a sample of, say, eight transistors, aged at some given environmental stress condition.

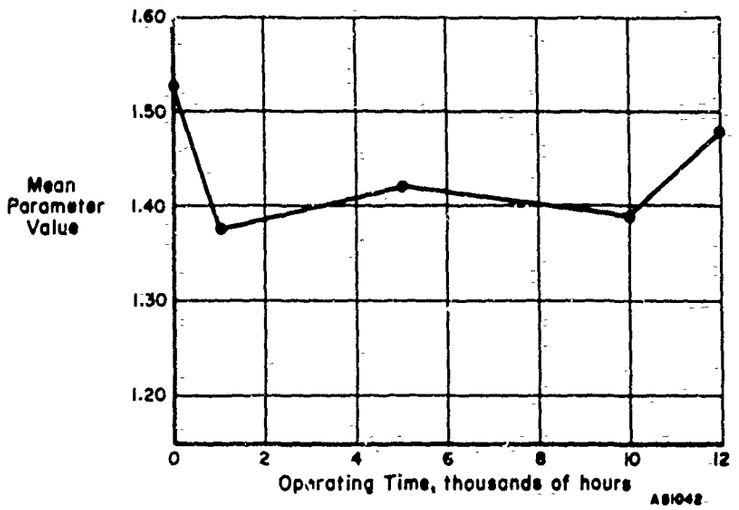


FIGURE 6.3. MEAN TIME HISTORY FOR ICBO AT 10 VOLTS (μ AMPERES)

The first question is: In the population, can one confidently conclude that there is any change at all of the mean parameter value as a function of time? There are observed differences between the means (an observed relationship), but such differences might be entirely a combination of sampling error and other sources of error, and there may be no difference at all between the true means. By techniques of analysis of variance, it is possible to calculate the probability that the observed differences are in fact due to error by taking into account such factors as sample size and differences in aging behavior between components. Then, if the probability is quite low that the observed differences are due only to error, say less than 0.05**, which is a traditionally used significance level, it is concluded with a high degree of confidence that there is a real relationship between operating time and parameter level. That is, since it is so unlikely that the

*See Winer[1], Chapter 4, "Repeated Measures Designs", for specific procedures and computational formulas in analyzing data of this nature.

**The other frequently used significance level is the 0.01 level. What level to use depends on a variety of considerations, but whatever level is used should be specified before any data are collected.

Downloaded from https://www.everyspec.com

observed differences are due only to error, it is concluded with confidence that there is in fact a real relationship – that there must be some differences (some relationship) between the true means. In this case, the differences between the sample means are said to be statistically significant. On the other hand, if it is sufficiently likely that the observed differences could be explained on the basis of error (the probability value is greater than 0.05), then the investigator cannot conclude, with confidence, that there are any differences between the true means, i.e., the observed differences are too easily explained by the operation of various error sources to conclude that they represent anything but error variations. In this case, the differences between the observed means are said to be not statistically significant.

It is important to note that nonsignificant differences in no sense prove that there is no true relationship between two variables. It only means that the observed differences are too easily explained on the basis of error to confidently conclude, with the available data, that there is a true relationship. More concretely, nonsignificant differences mean that the investigator cannot (with the level of confidence implied by the level of significance) order the different means from high to low. Thus, he cannot conclude, from the available data, whether the means increase or decrease with time or have any idea at the given level of significance where the mean value may increase or decrease.

In any case, nonsignificant differences end a statistical enquiry of this sort. But, what if, in Figure 6.3, the differences between the five means were to be statistically significant? This means only that there is some relationship between ICBO and time, that is, that one or more means in the population differ from one or more other means in the population. In order to obtain more specific information on which population means differ from other population means, significance tests can be made between separate pairs of means. For example, the data in Figure 6.3 indicate an increase in mean performance from 10,000 hours operating time to 12,000 hours. However, this increase observed in the sample may be due to error, and in truth the mean parameter value for 12,000 hours may be equal to, or even less than, the mean parameter value for 10,000 hours. The investigator might like to know if he can safely conclude that there is a true increase since, if there is, this could be interpreted as the effect of an aging mechanism operating from 10,000 to 12,000 hours that is different from the aging mechanism accounting for parameter change preceding 10,000 hours.

In order to determine whether or not there is a true increase, a test can be performed to see whether these two means differ significantly. Such a test would be exactly analogous to the previously described over-all test to determine whether there were significant differences among any of the means. If the difference between the two observed means is statistically significant (at a given level of significance), then the investigator can be highly confident that there is a true increase in the mean parameter value. If the difference between the two observed means is not statistically significant, then it cannot be concluded with confidence that there is a true increase; the two population means may be equal, or the population means may be in a direct opposite to that indicated by the sample means.

Several other pairs of means could also be tested for significance. For example, if the mean parameter value for 1000 hours were to be significantly lower than the mean parameter value initially, and assuming the previously discussed contrast yielded a significant increase in the mean from 10,000 to 12,000 hours, then the investigator would

Downloaded from https://www.everyspec.com

have corroboration for the "bathtub" hypothesis – that is, a decrease and then an increase in the mean parameter value over time.

6.4.1.2 Techniques Dealing With Medians and Proportions

There are several other specific techniques, apart from analysis of variance as illustrated above, used to test for significant differences among averages of one sort or another. They all provide essentially the same information about the relation between two variables, i.e., they all provide a means for deciding whether there is any relationship between two variables and, if so, provide information as to where the function might increase or decrease. Which particular technique to use depends, in large part, upon the particular "average" that is computed for the data. Analysis of variance is generally used in connection with arithmetic means, as illustrated in the previous example. However, in several instances where the functional relation between two variables is being studied, it is more appropriate to use median values of the dependent variable, rather than means, because of the occurrence of "extreme" data points. For example, the concern may be with median time to failure, as a function of some stress condition, like ambient temperature. In this case, the median test (and extensions of the median test) can be used to test for significant differences between medians.[2,3] All interpretations of such a test, both for the over-all relationship and for significant differences between pairs of medians, would be the same as those given for the previous example. The only difference is that inferences were being made about median values in the population rather than mean values.

Finally, in cases where a proportion is studied as a function of some independent variable, there are several techniques that can be used. For example, the concern might be with examining failure rate of components as a function of applied voltage. The failure rate is observed at various levels of voltage for samples of components in each voltage level. The first interest is again in determining if there is any true change, and next in testing various pairs of failure rates to see whether they differ significantly. Several test techniques are available for the purpose of testing differences between proportions, including (1) the McNemar test for the significance of changes, (2) the Fisher exact-probability test, (3) the chi-square test for two independent samples, (4) the chi-square test for k independent samples, and (5) the Cochran Q test. Siegel[3] and Bradley[2] show how these tests can be used in various situations.

6.4.1.3 The Question of Assumptions

In applying any of the above techniques, as well as techniques subsequently presented, it is important that certain well-defined assumptions concerning characteristics of the data be met. The issue of assumptions is a complex one and cannot be discussed in any detail here. The assumptions that need to be met differ with the particular technique used and, also, for a given technique, they may differ with the particular application. For example, in analysis of variance and most other techniques as well, in the case where the same components are observed at each value of the independent variable (as is the case for the independent variable time when observing the aging behavior of

Downloaded from https://www.everyspec.com

components under a constant environmental stress), the assumptions are different than when a different sample of components is observed at each value of the independent variable (as is the case when mean time to failure of components is studied as a function of different levels of environmental stress). Further, whatever the assumptions are for application of a given technique, these assumptions do not have to be met exactly. For example, the assumption of a normal distribution is frequently required. This assumption of normality does not have to be met exactly. How far the data can deviate from normality without affecting the validity of the technique is a complex problem and depends on the particular situation. The topic of assumptions itself, and the validity of various techniques under various conditions of deviating from the assumptions, has been given considerable study. In any case, the services of a statistician are essential for the checking of assumptions in each specific context and for the exercise of judgment in deciding whether the assumptions are "sufficiently" well met. If they are not, the assumption of normality, for example, might be met by transforming the data, such as by using the logarithm of the dependent variable.

In the use of a technique that deals with a certain type of average, the assumptions might not be met; but if a different type of average is used, the assumptions required for significance tests dealing with this average might very well be met. For example, the assumptions required for tests dealing with medians and proportions are generally less stringent than are assumptions required for those techniques concerned with means, namely analysis-of-variance and related techniques. However, it is also generally true that tests dealing with medians and proportions are less powerful than those dealing with means. Therefore, if a mean is a logically defensible summary measure, and if the required assumptions with tests dealing with means can be met (using some data transformations, if necessary), then use should be made of these more powerful techniques.

The one assumption that the above techniques and most other techniques of statistical inference have in common is that of random sampling. This fundamental issue itself is a complex one and often ignored in the practical application of statistical-inference techniques. To validly apply any of the above techniques, the sample or components under study must be, or must be "considered" to be, a random sample from the larger population of components of interest. This is because techniques of statistical inference are based on the "laws of chance", and chance must be allowed to operate without restriction, as far as samples are concerned. If the population of interest toward which inferences are made is a real and finite population, i.e., 5000 transistors of Type X, then it is a relatively simple matter to draw a random sample from this population. This would be done by first enumerating the 5000 transistors and then using some technique to draw a sample of a given size at random from the 5000. The problem of random sampling, however, becomes more complicated when the investigator is interested in drawing conclusions about an infinitely large conceptual population, e.g., "all possible components for Type X that are and ever could be made by a given manufacturing process". This is usually the situation in experimental research. In this case, obviously, all members of the population cannot be enumerated, and thus a random sample cannot be drawn from such a population. What can be done (but probably is not done very often), is to pick a sample at random from a very large group of available components that can safely be considered, in terms of experimental findings, to be essentially the same as the infinitely large population of interest. Then the investigator can fairly contend that any inferences he makes from sample data to the infinite population are legitimate.

Downloaded from https://www.everyspec.com

One of the most frequent situations in experimental research is that of simply using components that are "available", where no formal attention has been given to selecting such components at random from a defined population because of practical considerations or neglect. In this situation, in order to use techniques of statistical inference, it is necessary (and often done) to "consider" the sample as being a random one from a larger (infinite) population of components "like" the ones in the sample, i.e., that have generally the same statistical characteristics. The difficulty with such a position is that such a conceptual and loosely defined population may not be nearly equivalent to the defined population of interest, but statistical inferences can be made only about this conceptual population. Thus, if possible, this situation is to be avoided. If it cannot be avoided, any "extrapolation" from the conceptual population to the real population must be made on other than a measurable statistical basis.

The major point here is this: Practically all techniques of statistical inference, and all techniques covered in this section of this Notebook, require the assumption of random sampling from the defined population of interest. To the extent that this assumption can be considered to be met, then to this extent only are these techniques useful. This fundamental fact is often ignored in the application of statistical inference techniques.

One final point should be made about assumptions. It may have been implied from the examples given that the above techniques can be used only for a quantitative independent variable, that is, a variable where the levels can be ordered from low to high, such as time or voltage. These techniques can also be used for a qualitative independent variable, that is, where the levels of the independent variable cannot be ordered. An example of the use of analysis of variance in such a situation will be given later when discussing multivariate analysis.

6.4.2 Techniques of Curve Fitting

The second major class of techniques used to make inferences about the functional relation between two variables is curve-fitting techniques. Techniques of curve fitting are illustrated below in some situations encountered in reliability physics experiments. Both linear and nonlinear functions are considered.

6.4.2.1 The Fitting and Testing of a Linear Equation

The fitting and testing of a linear equation is the simplest case. As an example, consider the data in Figure 6.4, where mean values for some parameter of interest for a sample of transistors are shown (solid points) for various operating times.

Downloaded from https://www.everyspec.com

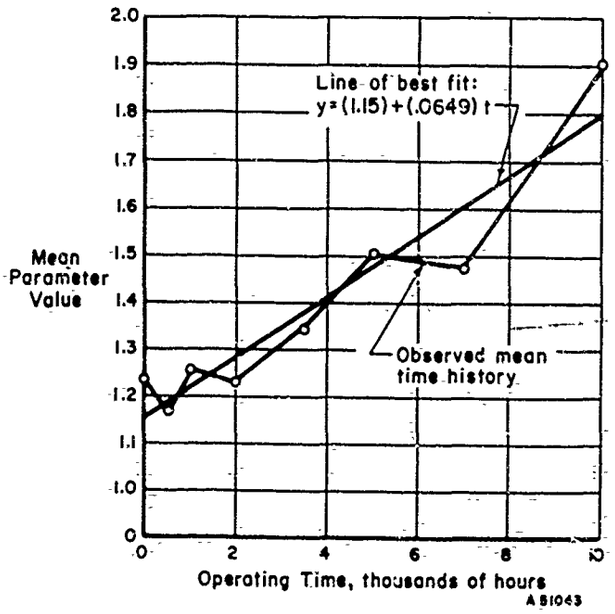


FIGURE 6.4. MEAN TIME HISTORY FOR A PARAMETER AND LINE OF BEST FIT

These data are for a given level of junction temperature, and represent only part of the total experimental data. That is, there are analogous mean parameter time histories not shown for, say, each of five other junction temperature levels. For the sake of illustration, it is assumed that the investigator is ultimately interested in applying the Arrhenius model to average rates of change of the parameter value and is therefore interested in two questions: (1) For each junction temperature level, is the assumption of linearity required for an Arrhenius analysis a reasonable one? That is, is it reasonable to assume, for any given junction temperature level, that the "true" (population) mean parameter values are linear (or approximately so) over time, any observed departures from linearity being attributed to error? and (2) If the assumption of linearity for any given junction temperature level is plausible, how is an "average" rate calculated, the average rate for each of the six temperature levels serving as the input to the Arrhenius model?

In order to answer these two questions, a linear equation must first be "fit" to the sample mean-time histories for each of the six junction temperature levels. Next, a goodness-of-fit test must be made for each temperature level to see whether the assumption of linearity is a plausible one. How this is done will be illustrated only for the data for one junction temperature level shown in Figure 6.4.

6.4.2.1.1 Fitting the Equation

In fitting the equation $y = a + bx$ to the observed data, where y is the mean parameter value and x is operating time, t, the general problem is to determine values for a and b such that the resulting graph of the equation "best fits" the observed data points or passes "as close as possible" to the observed data points shown in Figure 6.4.

Downloaded from https://www.everyspec.com

The almost universally accepted criterion for determining the values of a and b is the least-squares criterion. Use of this criterion provides values for a and b such that the sum of the quantities $(y_m - y_c)^2$ is a minimum, where y_m is an observed mean and y_c is the output of the linear equation. In the present example, there are eight observed means, i.e., eight values of y_m. Corresponding to these eight values are eight values of y_c. Using the least-squares criterion, values of a and b are obtained that minimize the sum of squares of the eight discrepancies between y_m and y_c.

The best fitting linear equation is shown in Figure 6.4 with a = 1.15 and b = 0.0649 as the least-squares values for a and b, which are easily calculated by well-known computational formulas. The value for b (0.0649) is taken as the "average rate" (per 1000 hours) to be used as an input to an Arrhenius analysis for this given temperature level, provided the assumption of linearity is a plausible one. This is the topic of "goodness of fit" and is discussed below.

6.4.2.1.2 Testing "Goodness of Fit"

The problem of goodness of fit for the example data is in deciding whether the true means can be assumed to be linear (or approximately so), i.e., to decide whether the observed departures from linearity can be accounted for by various sources of error. If so, the linear equation is said to provide a good fit. If not, the linear equation does not fit.

A frequently used way of testing goodness of fit for data such as shown in Figure 6.4 is to examine the "scatter" of the observed means from the fitted line. If the observed points fall "close" to the fitted line, and, perhaps more importantly, if the observed points do not indicate any "systematic" tendency to be nonlinear, then it is argued that the assumption of true linearity (or approximate linearity) is a plausible one, i.e., that a linear equation "fits". Otherwise, it is concluded that the assumption of linearity is not plausible. The difficulty with such an approach is that it has little quantitative basis and is quite subject to the investigator's personal interpretations. There are available various correlational measures of the "degree" of scatter, but application of these measures would seem to provide little additional information over and above visual inspection of the data and, in no practical sense, do such measures reduce any ambiguity.

Probably a better way, and certainly a more objective and rigorous way, of testing the linearity assumption is provided by means of analysis of variance techniques*. By using the analysis of variance, it is possible to determine whether the trend of observed means differs "significantly" from linearity. That is, is it possible to calculate the probability that the observed departures from linearity can be accounted for on the basis of various sources of error. If this probability is sufficiently low, say less than the traditionally used 0.01 or 0.05 criterion (significance) levels, then it can be concluded with confidence that the true relation is not linear but behaves according to some

*The data in Figure 6.4 represent a case of correlated observations. Lewis[4] provides analysis of variance goodness-of-fit tests for this case for any function, linear or otherwise.

Downloaded from https://www.everyspec.com

curvalinear function, and therefore the linearity assumption cannot be made or a linear equation does not "fit". In this case, various transformations* might be applied to the data in order to make it more nearly linear, and the same analysis of variance test applied to the transformer data.

On the other hand, if the probability is relatively high (in particular, greater than 0.01 or 0.05) that the observed departures from a straight line can be accounted for on the basis of error, then it can be concluded that the assumption of linearity (or approximate linearity) is a plausible one or that a linear equation does "fit". In this case, however, in no sense is it proved that the true relationship is linear. The only conclusion that can be drawn is that the investigator can reasonably assume the relation to be linear (or approximately linear) or that the observed data do not contradict the assumption of linearity. In any case, this is valuable information and, if the outcome of linearity tests for the other five thermal-stress levels were the same, it would be legitimate to proceed with an Arrhenius analysis using as an input the slopes of the fitted lines for each thermal stress-level. The fitting and testing of the Arrhenius model is described below.

6.4.2.2 The Fitting and Testing of a Nonlinear Equation: The Arrhenius Model

The techniques and principles of fitting and testing goodness of fit of a nonlinear function to a body of data can be illustrated by using the Arrhenius equation. In Table 6.2 below, some hypothetical data are given. For each of six junction temperature levels for a sample of ten transistors in each junction temperature level, average degradation rates are given for some parameter of interest. Each average rate was obtained by calculating the slope of the best-fitting least-squares line to mean-time history data, analogous to the data shown previously in Figure 6.4. It is assumed that, for each junction temperature level, a test for linearity was made and that in each case the assumption of linearity was found to be plausible, either for the original parameter values or for some transformed values of the parameter.

TABLE 6.2. EXAMPLE DATA: DEGRADATION RATES FOR SIX LEVELS OF JUNCTION TEMPERATURE

Junction Temperature		
Centrigrade	Absolute	Degradation Rates
25	298	0.0040
50	323	0.0050
75	348	0.0080
100	373	0.0170
125	398	0.0300
150	423	0.0700

*A log transformation might be used for the present data. If any transformation is used, it must be applied to each individual component time history, and then mean values calculated. If a transformation is applied to the original means, then the analysis of variance test for linearity cannot be used. Transformations applied to a large body of data are easily and efficiently accomplished by use of a digital computer.

Downloaded from https://www.everyspec.com

6.4.2.2.1 Fitting the Equation

The first problem is to fit the Arrhenius equation to the observed average rates. The Arrhenius equation is

$$R(T) = e^{A-B/T} \quad , \tag{6-1}$$

or, in log form

$$\ln R(T) = A - B(1/T) \quad , \tag{6-2}$$

where R(T) is the average rate for a given value of junction temperature T (degrees absolute), and A and B are constants to be empirically determined from the observed data.

The criterion used in determining values for A and B is the least-squares criterion, just as for a linear function. That is, values for A and B are determined such that the sum of the quantities $(y_m - y_c)^2$ is a minimum, where y_m is an observed rate and $y_c = R(T)$ is the output of the Arrhenius equation. Lewis[4] and Williams[5] give techniques for determining least-squares values of constants in nonlinear functions.

The least-squares solution is shown in Figure 6.5, where it is seen that A = 6.75 and B = 4008. The observed average rates from Table 6.2 are indicated by solid dots. The fit looks respectable, but "how good" it is must be determined by a test of goodness of fit, to be described subsequently. Figure 6.6 shows the log plot of the equation against reciprocal temperature. The logarithms of the observed rates (solid points) are also shown.

It is important to note that the estimates for A and B in Equation (6.1) were not obtained by use of the logarithmic form of the equation (Equation 6.2), i.e., by fitting a least-squares straight line to a log plot of the observed rates versus reciprocal temperature and taking the slope of this line as the value for B and the intercept of this line as the value for A. Such a linear reduction process for solving for A and B seems to be the technique generally used, but it does not provide least-squares values of A and B for the original data, i.e., for the data as shown in Table 6.2 or Figure 6.5. Such values are often a good approximation to the true least-squares values but can often be appreciably in error. From a logical point of view, the original data should be "best fit" by the equation. Further, the goodness-of-fit test to be described below deals with the original values and cannot be properly used unless a true least-squares solution is provided on the original rates.

6.4.2.2.2 Testing "Goodness of Fit"

The next step, after fitting the equation, is to test goodness of fit. The problem is to decide whether the observed departures of the average rates from the fitted Arrhenius model can be attributed to sampling error and other sources of error. That is, in Figure 6.5, can the departures of the solid points from the curve of best fit be reasonably attributed to error?

Downloaded from https://www.everyspec.com

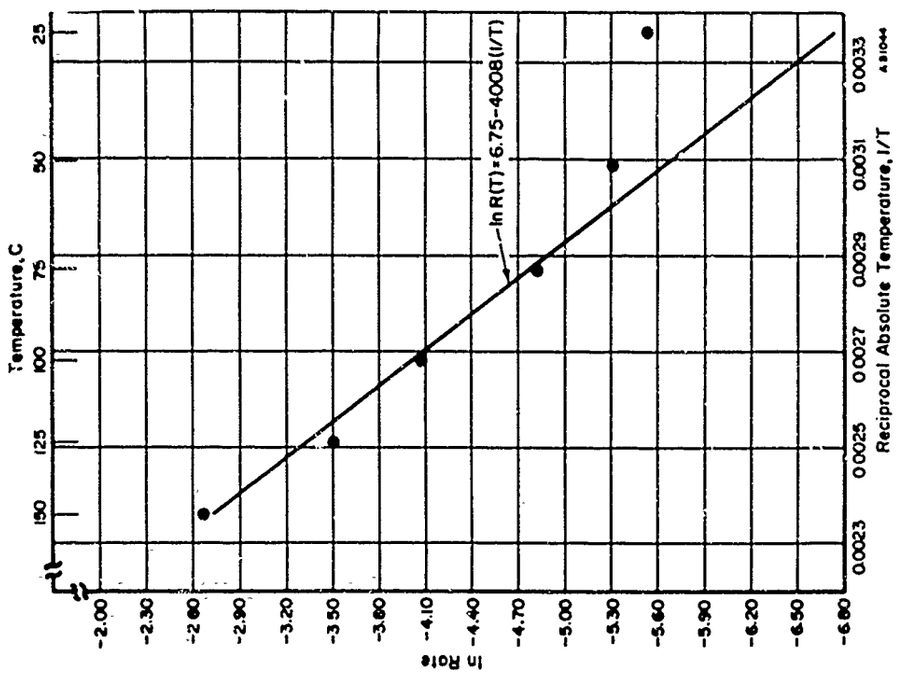


FIGURE 6.6. ARRHENIUS PLOT FOR EXAMPLE DATA

(Single Solution)

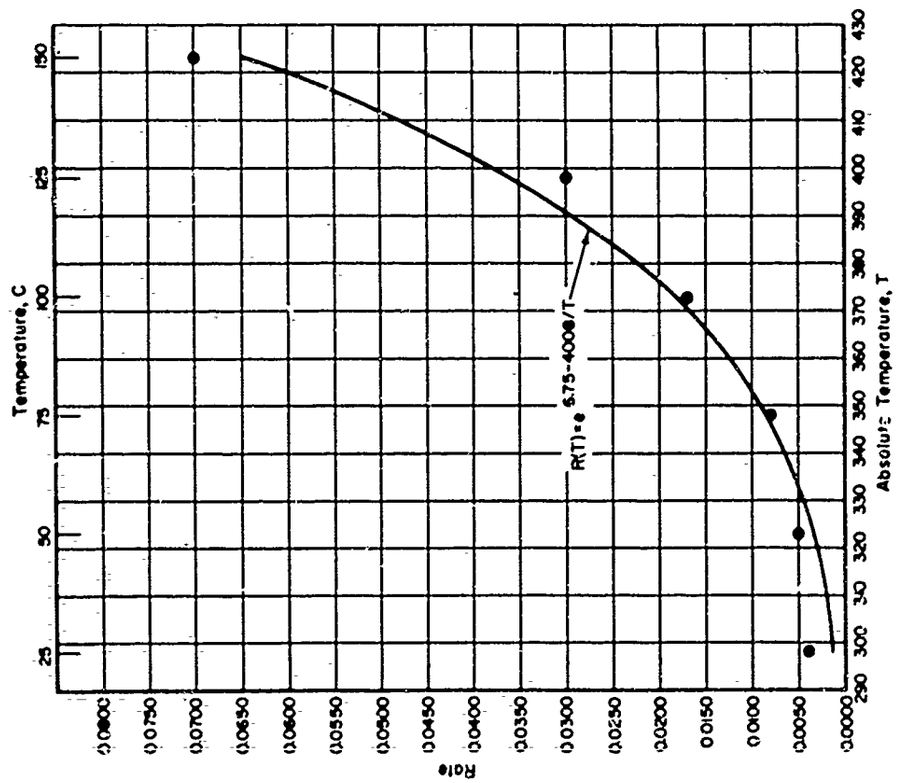


FIGURE 6.5. ARRHENIUS EQUATION FITTED TO EXAMPLE DATA

(Single solution)

Downloaded from https://www.everyspec.com

The problem is exactly analogous to the problem of goodness of fit of a linear function and, again, techniques of analysis of variance can be used. That is, the probability that the observed departures from the curve in Figure 6.5 are due solely to error is calculated. If this probability is quite low, i.e., if it is unlikely that the departures from the Arrhenius model are due solely to error, then it is concluded that the model does not fit – that the "true" rates do not conform to an Arrhenius equation but conform to some other equation. If the probability is relatively high that the observed departures can be attributed solely to error, then it is reasonable to conclude that the "true" average rates can be described (or approximated) by an Arrhenius equation.* Again, this latter outcome would in no sense prove the Arrhenius model to be true: other equations, yielding other calculated curves would also fit within the error limits of the data. It only means that the Arrhenius model is one plausible explanation of how the true rates vary as a function of junction temperature or that the observed data do not contradict the Arrhenius model. Nevertheless, of the several functional equations that would fit the data, the Arrhenius model would be chosen because of its theroetical and practical usefulness, and the above test provides a means for determining whether it is in fact permissible to retain the model.

In any case, if the Arrhenius model fits, in the above sense, then a single-aging mechanism over the entire range of junction temperature investigated can be assumed to be plausible. Also, the estimated value of B can be used to compute acceleration factors, and the estimated values for A and B in the equation can provide an estimate of the "true" average rate for any given junction temperature within the range of temperature investigated.

However, if the Arrhenius model does not fit over the entire range of junction temperature, it is possible to fit two Arrhenius equations, one equation to one subrange and another equation to another subrange. This was done for these data, yielding two solutions, as shown in Figure 6.7. The original fit to the entire range is also shown. Figure 6.8 shows the log plots for all three solutions. One solution is for the range from 25 to 75 C, the other solution for the range from 75 C to 150 C.

If both these equations fit the points in the respective subranges of junction temperature, as judged by using the above test of goodness of fit, then it can be concluded that

*In order to apply a test of this sort, it is necessary to deal, literally, with arithmetic means and to have a measure of variability around each of these means. The solid points in Figure 6.5 can in fact be considered to be arithmetic means. That is, it can be shown that a rate (slope) determined by fitting a least-squares line to a mean-time history, as these six rates are determined, is exactly equivalent to the arithmetic mean rate determined by fitting a least-squares line to each individual component time history and then getting the mean of the slopes of the individual component time histories (assuming that no components are dropped from test over the time interval of concern). Also, the variability around any one of these means, which enters into the error term in the test, can be obtained by using a computer to calculate each individual component rate (which is efficient, rapid, and not costly).

In order to use this test, the equality of variance assumption must be met. If it cannot be met, then separate tests on each mean rate can be performed using the point on the curve as the hypothesized value in a t ratio.

Downloaded from https://www.everyspec.com

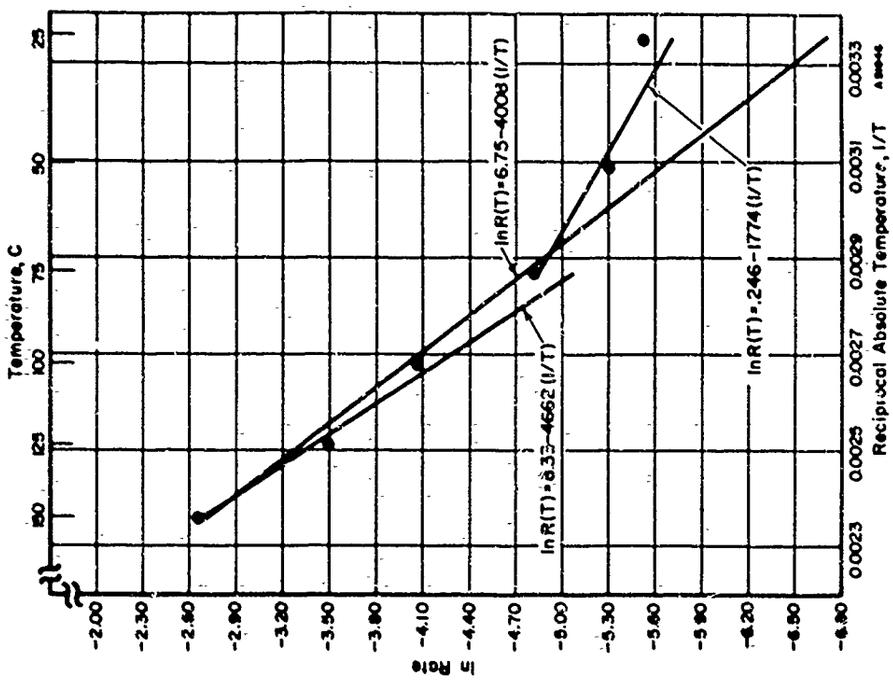


FIGURE 6.8. ARRHENIUS PLOTS FOR EXAMPLE DATA (ALL THREE SOLUTIONS)

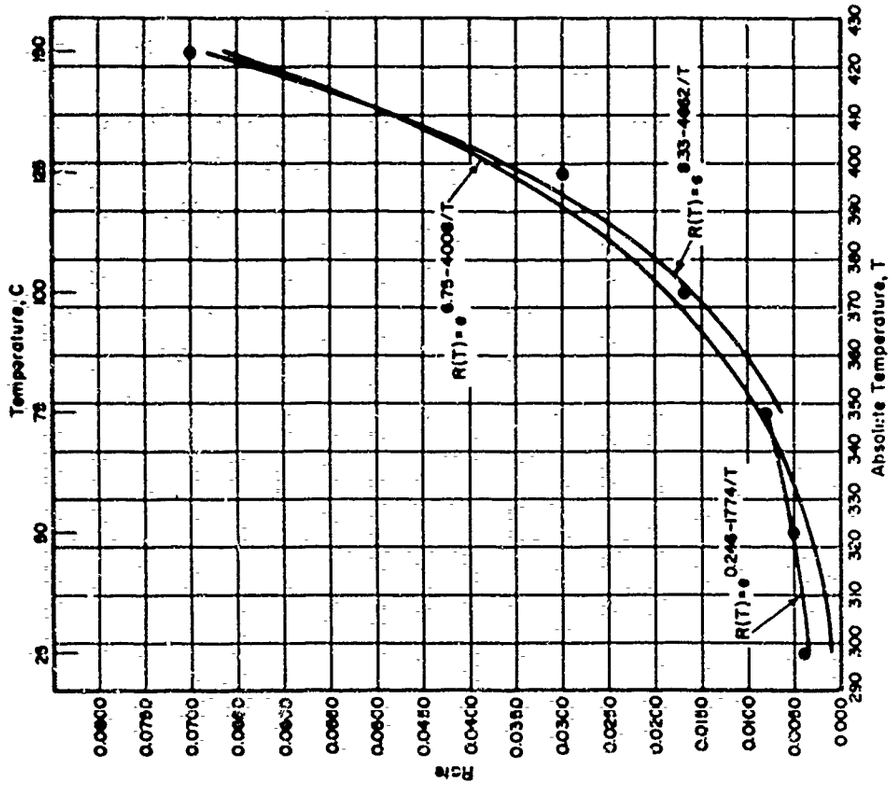


FIGURE 6.7. ARRHENIUS EQUATION FITTED TO ENTIRE TEMPERATURE RANGE AND TO TWO SEPARATE SUB-RANGES OF TEMPERATURE

Downloaded from https://www.everyspec.com

there may be two different aging mechanisms operating over the entire temperature range – one mechanism from 25 C to 75 C and a different mechanism from 75 C to 150 C. Acceleration factors can then be computed for each separate solution. However, it is to be recognized that in applying goodness-of-fit tests to selected subranges of data, such tests are not completely valid, since selection of the subranges involves capitalizing of the particular error that happens to be peculiar to certain data points. If such "after-the-fact" analysis is done, and one, or both, equations is found to fit, a separate experiment, using a different sample of components, would be desirable in order to corroborate the results.

6.4.2.3 Fitting and Testing Other Nonlinear Functions

In fitting and testing other nonlinear functions, the principles are essentially the same. The objective in fitting a specified equation to the data is to obtain least-squares values for the constants in the equation. Various techniques are available for this purpose. The goodness of fit of the equation can then be assessed by an analysis-of-variance test, which can be used for any nonlinear function.* If the equation fits, it can be used to predict the "true" (population) average value of y for any given value of x over the range of data to which the equation was fitted and thus provide an estimate of how the dependent variable changes as a function of the independent variable. Any extrapolations beyond the range of data to which the equation is fit must be made on logical grounds rather than empirical or statistical grounds.

There is a wide variety of equations with a wide variety of characteristics that might be fit to a given body of data. Some equations often found useful are shown in Table 6.3 Each equation has its special characteristics. For example, Equation (8) (a polynomial), as contrasted with the other equations, has the capability of increasing and then decreasing (or conversely), whereas the other equations are all increasing (or decreasing) functions. Equation (6), the Gompertz curve, often used to describe growth processes, can yield an inflection point, i.e., produce an "s" shaped curve if this is indicated by the data.

TABLE 6.3. SOME USEFUL NONLINEAR EQUATIONS

	Equation	Linear Form of Equation
Parabolic and Hyperbolic	(1) $y = ax^b$	$\log y = \log a + b \log x$
	(2) $y = ax^b + c$	$\log (y - c) = \log a + b \log x$
	(3) $y = \frac{x}{a + bx}$	$\frac{1}{y} = a\frac{1}{x} + b$
Exponential	(4) $y = ae^{bx}$	$\log y = \log a + (b \cdot \log e)\, x$
	(5) $y = ae^{bx} + c$	$\log (y - c) = \log a + (b \cdot \log e)\, x$
	(6) $y = vg^{h^x}$	$\log (\log v - \log y) = \log (-\log g) + (\log h)\, x$
Logarithmic	(7) $y = a + b (\log x)$	--
Polynomial	(8) $y = a + bx + cx^2$	--

Note: All logarithms are to the base ten.

*Reference is made to Lewis[4] and Williams[5] for computational formulas and descriptions of such analysis-of-variance tests for both linear and nonlinear functions.

Downloaded from https://www.everyspec.com

The linear form of these equations, where appropriate, is also shown in Table 6.3. By plotting observed data on scales implied by the linear form of the equation, a preliminary idea can be obtained as to whether the equation will fit before any actual calculations of least-squares constants are made. For example, considering Equation (1), if the logarithm of the observed data points is "approximately" linear when plotted against the logarithm of the independent variable x, then the equation $y = ax^b$ may fit the data. However, tests of goodness of fit should be made on the original untransformed data by using analysis-of-variance techniques.

6.4.2.4 The Accuracy of Constants in a Fitted Equation

If a given equation is found to fit the data as a result of applying an analysis-of-variance test, the problem still remains as to how "close" the calculated constants in the equation might be to the true values of these constants in the population. Goodness-of-fit tests provide no direct information on this equation – they only help one decide whether the form of equation is a plausible one or not. For example, the calculated value of B is more or less in error even if the goodness-of-fit test indicates that the Arrhenius model is plausible. To the extent that it is in error, any calculations of acceleration factors are in error or any inferences made from the value of B about the nature of an aging mechanism may be in error. Further, even a small error in B and an associated acceleration factor may be seriously magnified when extrapolations of parameter values are made beyond the operating time on which test data are collected.

The accuracy of calculated values of constants depends on sample size and on a number of other factors as well. For a given situation, just how accurate the estimated values of the constants are is an extremely complex issue and cannot be discussed here. Further, the conventional techniques apply only for the case of uncorrelated observations (e.g., different components under each of several stress levels) and not to the case where observations are correlated (e.g., the same components being observed over time). Reference is made to Ezekiel[6] and Williams[5] for techniques of assessing the reliability of constants in linear and various nonlinear equations.

6.5 Extensions to Multi-Variate Analysis

The concern in this part is with the change in the average value of a given dependent variable as a function of two or more independent variables that are being varied in a single experiment. For example, the investigator may be interested in mean time to failure for a particular transistor type, as a function of both junction temperature and collector voltage; or he may be interested in the mean value of ICBO at 10 volts, as a function of both operating time and junction temperature.

Again, as was the case when analyzing the functional relation between one dependent and one independent variable, the concern here is with making inferences about the "true" functional relation, i.e., the functional relation that obtains in the population. Two major classes of techniques are quite useful for making such inferences: (1) analysis of variance and associated techniques and (2) surface fitting. The use of these two classes of techniques is described below.

Downloaded from https://www.everyspec.com

6.5.1 Analysis of Variance

As contrasted with the case of one independent variable, the unique aspect of analysis of variance for two or more independent variables is that it provides a means for testing the interactions of the independent variables, i.e., the possible nonadditive effects of two or more variables on the average value of a dependent variable. The use of and the information provided by analysis of variance is described and illustrated below for various situations occurring in reliability physics experimentation where this technique might be applicable. Of the several possible analysis-of-variance designs and associated analyses, only those felt to be of most applicability to the reliability physicist are covered*.

6.5.1.1 Two Independent Variables

The use of analysis of variance for this case can be illustrated by reference to the hypothetical data given in Table 6.4 and Figure 6.9. In the cells of Table 6.4 are shown average rates of change (per 1000 hours) for some parameter of interest, for various combinations of ambient temperature and collector voltage. Each average rate shown is based on a sample of five transistors. For example, the average rate for T = 50 C and V_c = 0 was obtained from the mean time history of five transistors aging under this stress condition, the parameter being measured at selected operating times out to 10,000 hours. The rate was obtained by taking the slope of the least-squares line fitted to the observed means over time, analogous to the data shown in Figure 6.4. The other rates were similarly obtained.

TABLE 6.4. AVERAGE RATE OF CHANGE (FOR A GIVEN PARAMETER) FOR VARIOUS COMBINATIONS OF AMBIENT TEMPERATURE AND COLLECTOR VOLTAGE

Collector Voltage, V_c	Ambient Temperature, T: T = 50 C	T = 75 C	T = 100 C	T = 125 C	
V_c = 0	.020	.024	.019	.031	.024
V_c = 10	.025	.029	.033	.050	.034
V_c = 20	.027	.027	.043	.056	.038
V_c = 30	.029	.036	.051	.053	.042
	.025	.029	.036	.048	

Main Effect of Temperature

Main Effect of Voltage

*Reference is made to Winer[1] and Cochran and Cox[7] for the treatment of the analysis-of-variance designs covered here.

Downloaded from https://www.everyspec.com

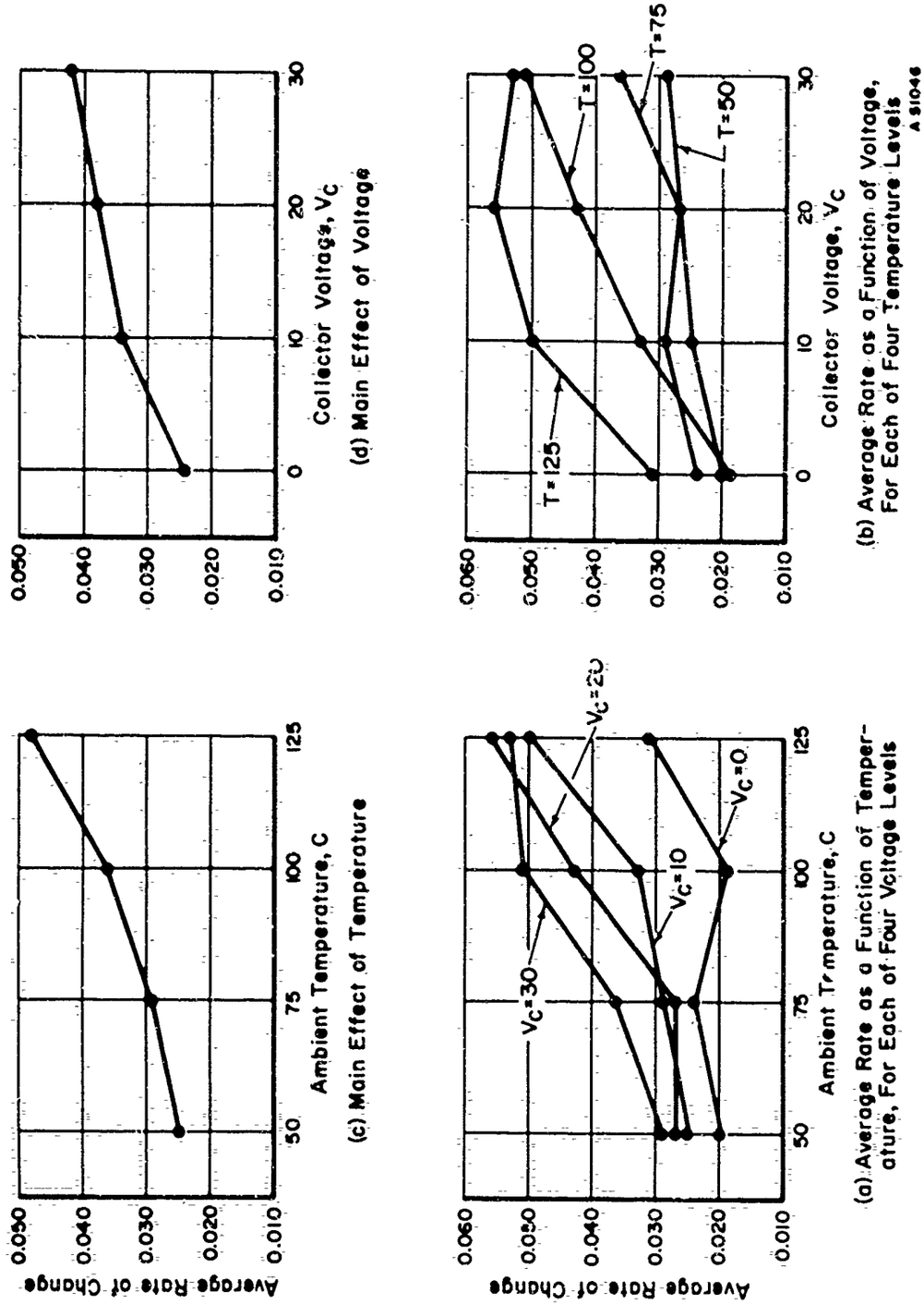


FIGURE 6.9. GRAPHS OF DATA FROM TABLE 6.4

Downloaded from https://www.everyspec.com

The data in Table 6.4 are shown graphically in Figure 6.9. Figure 6.9(a) shows a plot of the average rate as a function of ambient temperature for each of the four voltage levels. The same data is shown in Figure 6.9(b), but plotted by voltage for each of the four temperature levels.

Thus, average rate (the dependent variable) is being studied as a function of the two independent variables of temperature and voltage. Examination of the observed averages indicates certain trends. The problem is to take into account the error (sampling error, measurement error, and other sources of error) that is inherent in the observed trends, and to make inferences about the "true" (population) effect of temperature and voltage on the average rate of change of the parameter.

In analyzing such data in an analysis-of-variance framework*, the first concern would be with the interaction effects of temperature and voltage. Referring to Figure 6.9(a), an observed interaction is said to exist between temperature and voltage if the four curves (one for each voltage level) are not parallel, i.e., if the relative effect that temperature has on average rate of change depends on the particular level of voltage, or the temperature effect itself differs from one voltage level to another. If the four curves were to be parallel (have the same shape), then there would be no observed interaction. Further, if the effect that temperature has depends on the particular level of voltage [as seen to be the case from Figure 6.9(a)], then it also follows that the effect of voltage depends on the particular level of temperature [as seen from the nonparallel curves in Figure 6.9(b)]. Conversely, parallel lines in a plot such as in Figure 6.9(a) would also mean no interaction (parallel lines) in a plot such as in Figure 6.9(b).

Thus, it is evident that there is an observed interaction between temperature and voltage, since the four curves in Figure 6.9(a) and the four curves in Figure 6.9(b) are not parallel. However, this observed interaction may be due only to error, i.e., if each average rate were to be based on a much larger sample (population) of components, instead of only five components, there may be no interaction at all, or all the "true" curves may be parallel. By means of analysis-of-variance techniques, it is possible to calculate the probability that the observed interaction is in fact due only to error. If this probability is quite low, say less than the 0.01 or 0.05 significance levels, then it is concluded with confidence that there is a true interaction, and the observed interaction is said to be statistically significant. On the other hand, if the probability that the observed interaction is due only to error is not sufficiently low (does not reach the 0.01 or 0.05 significance levels), it cannot be concluded with confidence that the two variables, in truth, interact. In this case, the observed interaction is not statistically significant. That is, it is plausible to assume (but in no sense is it proven) that there is no real interaction between the two variables, the observed interaction being within "error limits" of the data, or the observed interaction is too easily accounted for on the basis of error to conclude that any real interaction exists.

Mathematically, a significant interaction would mean that the true average rates, as a function of temperature and voltage, would need to be described by the following functional form:

*In order to analyze these data in an analysis-of-variance framework, it is necessary to consider each of the 16 average rates as arithmetic mean rates of individual rates. This can be done (see the footnote on page 6-18). For the error term in an analysis-of-variance test, the variability of the individual rates within each cell of the matrix is used.

Downloaded from https://www.everyspec.com

$$R(T) = f(T) + f(V) + [f(T)]\,[f(V)] \quad , \tag{6.3}$$

where R(T) is the average rate, f(T) is some function of temperature, and f(V) is some function of voltage. Thus, a significant interaction means that a multiplicative [f(T)] [f(V)] term is required to describe the true relationship, or that the true effects of temperature and voltage are not simply additive. On the other hand, if the interaction is not significant, then it is permissible to assume that the multiplicative term in Equation (6.3) is zero, or it is plausible to assume that the true effects are simply additive.

In any case, if an analysis-of-variance test indicates nonsignificant interaction, zero (or negligible) interaction in the population is generally assumed. That is, it is assumed that the true effects of temperature for each of the four given levels of voltage are quite similar, and that the true effects of voltage for each of the four temperature levels are quite similar. This being the case, attention is then focused on the main effect of both temperature and voltage. The observed main effect of temperature is obtained by averaging across rows in Table 6.4, and is shown at the bottom of Table 6.4 The observed main effect of temperature is thus the average of the four individual curves (one for each of the four voltage levels) shown in Figure 6.9(a). This main effect is plotted in Figure 6.9(c). The main effect for voltage is analogously defined, and is plotted in Figure 6.9(d).

Under the assumption of zero (or negligible) interaction, the main effect of temperature is then taken as the best estimate of how the true average rate changes as a function of temperature, regardless of the particular level of voltage. This estimate would then replace the four individual curves shown in Figure 6.9(a). Similarly, the main effect of voltage is taken as the best estimate of how the true average rate changes as a function of voltage, regardless of the particular level of temperature. This estimate would then replace the four individual curves shown in Figure 6.9(b). Further, significance tests can be applied to each of the main effects shown in Figures 6.9(c) and 6.9(d), exactly analogous in interpretation to the significance tests described previously. For example, if the four means shown in Figure 6.9(d) do not differ significantly (at a prescribed level of significance), then it cannot be concluded with confidence (on statistical grounds) that voltage has any effect at all on average rate of change of the parameter. On the other hand, if the four means shown in Figure 6.9(d) differ significantly, it is concluded with high confidence that there is in fact a real relationship between temperature and voltage. In this case, individual pairs of average rates, shown in Figure 6.9(d), can then be tested for statistical significance, with the same interpretations as previously given (see Section 6.4.1.1).

The main effect of temperature can also be tested for significance, as well as individual pairs of average rates [Figure 6.9(c)], to see whether it can be concluded that there is any relation at all between temperature and average rate of change, and, if so, to see which pairs of average rates differ significantly.

If the analysis-of-variance test indicates a significant interaction, the subsequent analysis proceeds along quite different (and more complex) lines. In this case, since the true effects of temperature vary, depending on the level of voltage, and the true effects of voltage vary, depending on the level of temperature, the two main effects lose much of their meaning. It is then necessary to focus attention and subsequent analysis on the four individual curves shown in Figure 6.9(a), and the four individual curves shown in Figure 6.9(b). These eight curves are referred to as simple effect curves (as contrasted to

Downloaded from https://www.everyspec.com

the main effect curves). Each simple effect can be tested for statistical significance. For example for T = 50 [Figure 6.9(b)], voltage may not have a significant effect on average rate of change, but for T = 125, a test may reveal a significant affect of voltage. Again, comparisons between individual pairs of average rates, in Figures 6.9(a) and 6.9(b), can be made. Exactly what significance tests are made, with respect to simple effects and individual pairs of means, depends on the purposes of the investigator.

Another example of analysis-of-variance techniques, as applied to the case of two independent variables, would be the analysis of data such as was shown in Figure 6.2. For example, after completion of "initial" effects (e.g., after 1000 hours of operating time), the investigator may want to make a test to see whether the observed interaction between time and the stress variable is statistically significant, i.e., to see whether the effect of operating time on the mean parameter value depends on the particular stress level under which the components are aged. Also, main effects of operating time could be examined (the average of the three curves shown in Figure 6.2), and simple effects could be tested for significance (i.e., each of the three curves shown in Figure 6.2).

6.5.1.2 More Than Two Independent Variables

Figure 6.2 shows data for the two independent variables of time and some stresses, e.g., ambient temperature. If such data were to be had for each of, say, three levels of some other stress, e.g., voltage, then this would be an example of a three-variable analysis, i.e., the mean value of some parameter as a function of operating time, ambient temperature, and voltage. Analysis-of-variance techniques could be used too for testing interactions among these three variables, and for testing the main and simple effects of these variables. However, the analysis would be much more complex than for the case of two independent variables. For example, there would be the possibility of an interaction between operating time and ambient temperature, between operating time and voltage, and between temperature and voltage. That is, there would be these three two-way interactions to consider. Further, in addition to, say, a possible interaction between operating time and temperature, the nature of this interaction itself may depend on the particular level of voltage, which is termed a three-way interaction between time, temperature, and voltage.

In general, when a large number of independent variables are being investigated (say, more than three), frequently both conceptual and methodological difficulties are involved in carrying out and interpreting the analysis, especially in analyzing for interactions involved. For this reason, such designs are often not employed.

A useful analysis-of-variance design, when several variables of interest are being evaluated in a single study, is the fractional replications design, described below.

6.5.1.3 Fractional Replications Design

In the case of several independent variables, each of which can assume several levels, the number of combinations to be evaluated can be quite large. In turn, this necessitates that a large number of components be used, which may often be impractical. In such a situation the fractional replications design[7,1], a type of analysis-of-variance

Downloaded from https://www.everyspec.com

design, is highly useful. In using this design, data are not collected for all possible combinations of interest, but only for some fraction of the total number of possible combinations, e.g., 1/2 of the total number. The particular combinations data are collected for are chosen so as to provide the maximum amount of relevant information.

The use of a fractional replications design can be illustrated by reference to a study by Gorton[8], concerned with the development of a dosimeter system for measuring fast-neutron irradiation. The purpose of the study was to determine optimum material and device properties and processing procedures to be used in the fabrication of a radiation detector device. In particular, various combinations of the following four independent variables were studied:

(1) Three different material suppliers

(2) Three different levels of material lifetime

(3) Three different process variations

(4) Three different levels of base width.

The effects of these variables on each of several device parameters were investigated (e.g., average preirradiation carrier lifetime in the device).

Thus, a total of 81 (i.e., 3^4) possible combinations for fabricating the device were considered. These 81 combinations are shown tabularly in Table 6.5. To construct and evaluate a number of specimen devices for each of the 81 combinations was prohibitively costly. Therefore, device specimens were made for only 27 of the possible 81 combinations. These 27 combinations on which performance data were collected are indicated by code numbers in Table 6.5, the dashed lines indicating that no devices were made and evaluated for these combinations. For each of the 27 combinations that were evaluated, a sample of 10 devices were constructed, giving 270 devices in all. The particular 27 combinations were chosen so as to provide the maximum amount of relevant information.

Thus, data were actually collected for only a portion (27) of the total number (81) of treatment combinations of interest, and therefore only 27/81 = 1/3 of the matrix was filled with data. Even so, several useful conclusions could be drawn concerning ways of constructing a radiation detector device.

In using a fractional replication design, it must be assumed that certain interactions between the independent variables are zero, or negligible. Whether these assumptions can be met depends on the particular variables being investigated, and should be carefully considered in each specific instance before using such a design. A statistician can indicate specifically which interactions must be assumed to be zero (or negligible), but it is a question for the investigator to decide whether such assumptions can in fact be made, based on previous engineering knowledge. That is, these interaction assumptions required for use of a fractional replications design cannot be checked statistically or empirically, and must be evaluated on logical or engineering knowledge grounds.

Downloaded from https://www.everyspec.com

TABLE 6.5. COMBINATIONS OF PARAMETERS INVESTIGATED

		Supplier 1 A = 0 Material Lifetimes			Supplier 2 A = 1 Material Lifetimes			Supplier 3 A = 2 Material Lifetimes		
		B = 0	B = 1	B = 2	B = 0	B = 1	B = 2	B = 0	B = 1	B = 2
Processing Treatment 1 C = 0	Base Widths D = 0	0000	----	----	----	----	1200	----	2100	----
	Base Widths D = 1	----	0101	----	1001	----	----	----	----	2201
	Base Widths D = 2	----	----	0202	----	1102	----	2002	----	----
Processing Treatment 2 C = 1	Base Widths D = 0	----	0110	----	1010	----	----	----	----	2210
	Base Widths D = 1	----	----	0211	----	1111	----	2011	----	----
	Base Widths D = 2	0012	----	----	----	----	1212	----	2112	----
Processing Treatment 3 C = 2	Base Widths D = 0	----	----	0220	----	1120	----	2020	----	----
	Base Widths D = 1	0021	----	----	----	----	1221	----	2121	----
	Base Widths D = 2	----	0122	----	1022	----	----	----	----	2222

Downloaded from https://www.everyspec.com

6.5.2 Surface Fitting

As contrasted with curve fitting, where there is only one independent variable in the equation, surface fitting is concerned with the mathematical relationship between a given dependent variable and two or more independent variables, i.e., $y = F(x_1, x_2, \ldots x_n)$. In particular, the concern is with the form of equation that might be used to describe the way in which the true (population) average value of some dependent variable changes as a function of two or more independent variables, and also the estimation of the values of constants in any such equation. Again, as was true for curve fitting, the values of these constants may have theoretical meaning, or they may be used only empirically in the equation to provide a prediction of the true (population) average value for various values of the independent variables.

Curve fitting and surface fitting are, at a conceptual level, much the same so far as fitting a function and assessing goodness of fit are concerned. However, methodologically, surface fitting is much more complex than curve fitting, and often appropriate techniques are not available. Below, a rather brief description of surface fitting is given, restricted to the case of only two independent variables. Extensions to the case of more than two independent variables are easily made, at least conceptually.

6.5.2.1. Fitting an Equation

For a given body of data, where the effects of two independent variables on a given parameter are being investigated, a wide variety of equations might be fit to the data. In fitting any one of these equations, the same least-squares criterion is used, just as for the case of only one independent variable. For example, for one reason or another, the linear equation $y = a + b x_1 + c x_2$ might be fit to the data previously shown in Table 6.4 and Figure 6.9, where y = average rate, x_1 = temperature, and x_2 = voltage. Use of the least-squares criterion would provide values for a, b, and c such that the sum of the quantities $(y_m - y_c)^2$ is a minimum, where y_m is an observed average rate and y_c is the output of the equation. In this case, there are 16 values of y_m, i.e., 16 average rates, one in each cell of the matrix. Corresponding to each of these 16 average rates are 16 values of y_c. Using the least-squares criterion, values of a, b, and c are obtained that minimize the sum of squares of the 16 discrepancies between y_m and y_c. Standard techniques are available for obtaining least-squares constants in a linear equation, and the calculations are efficiently and quickly carried out by digital-computer techniques.

After the equation is fitted, y is then taken as an estimate of the population (true) average rate for any given values of temperature and voltage within the range of these variables studied, assuming the linear equation is judged to "fit" adequately (goodness of fit will be discussed subsequently).

Another equation that might be desired to be fit to the data in Table 6.4 is the Eyring rate equation, given by

$$R(T,S) = \left[ATe^{-B/kT}\right]_e^{(C + D/kT)S} , \qquad (6.4)$$

Downloaded from https://www.everyspec.com

where R(T,S) is the average rate of change, T some thermal stress (in this case, ambient temperature), S some nonthermal stress (in this case, collector voltage), k is Boltzmann's constant, and A, B, C, and D are constants to be empirically determined from the data. Again, the objective is to obtain values for A, B, C, and D that minimize the sum of the squared discrepancies between the observed average rate and the average rates calculated from the equation. Obtaining least-squares values for constants in an equation of this form is no routine problem. However, approximate least-squares values can be obtained by considering the logarithmic form of the equation,

$$\ln\left[\frac{R(T,S)}{T}\right] = \ln A - \frac{B}{k}\left(\frac{1}{T}\right) + C(S) + \frac{D}{k}\left(\frac{S}{T}\right), \qquad (6.5)$$

where it is seen that $\ln\left[\frac{R(T,S)}{T}\right]$ is linear in $\frac{1}{T}$, S, and $\frac{S}{T}$. Consequently, computerized techniques for solving for the constants in linear equations might be used to solve for A, B, C, and D. This would provide approximate least-squares values for A, B, C, and D, but not true least-squares values, because of the same considerations as pointed out previously (Section 6.4.2.2.1). How close the approximation is would depend on the characteristics of the data. Thomas[9] gives a graphical technique for solving for the constants in the Eyring model, which also yields approximate least-squares values.

6.5.2.2 Testing Goodness of Fit

The problem of "goodness of fit" of an equation fitted to the data is essentially the same for surface fitting as for curve fitting. That is, the problem is to decide whether the population (true) averages can be described (or approximated) by the type of equation selected, i.e., to decide whether the departures of the observed averages from the averages predicted by the equation can be largely attributed to various sources of error (sampling error, measurement error, etc.). If the departures of the observed data from the model can be attributed (or largely attributed) to error, then the equation is said to "fit". Otherwise, the equation does not fit. For example, if the Eyring equation were to be fitted to the data in Table 6.4, 16 rates could be calculated from the equation, one rate for each cell of the matrix. Corresponding to each of these 16 calculated rates is an observed rate of change (the observed rates shown in Table 6.4), which, in each case, would differ more or less from the rate calculated from the Eyring equation. The question then is: Can the 16 discrepancies between the rates calculated from the Eyring model and the observed rates be attributed to error? If so, the model fits. If not, the model would be rejected.

Analysis-of-variance tests, as previously described for testing the fit of a curve involving only one independent variable, can probably be extended to the case where several independent variables are in the equation. Certainly, the multiple t test method could be used, and would probably be better in many situations, since the assumptions required are less stringent.*

*In particular, as mentioned before, the assumption of equality of variance need not be made, if a t test technique is used.

Downloaded from https://www.everyspec.com

6.5.2.3 The Accuracy of Constants in a Fitted Equation

The accuracy of the calculated constants in a fitted equation, where there is more than one independent variable, is a complex problem, and cannot be gone into here. Again, as was true in curve fitting, no goodness-of-fit test provides direct information on this question – such tests only help one decide whether or not the form of the equation is a plausible one. For example, the calculated values of A, B, C, and D in the Eyring model are in error more or less, i.e., deviate from population (true) values, even if a goodness-of-fit test indicates that the Eyring model is plausible.

Formal statistical methods are available only for assessing the accuracy of constants in an equation where y is linear in the x's, or where y is linear in various functions of x's, e.g., in a polynomial equation such as $y = a + b\,x_1 + c\,x_1^2 + d\,x_2 + e\,x_2^2$. The reader is referred to Ezekiel[6] and Williams[5] for techniques for assessing the accuracy of constants in such equations.

6.6 Error Reduction

In experimentation, apart from increasing sample sizes, there are two major ways to reduce the likely amount of error in the data: (1) by experimental control and (2) by statistical control. These two techniques and their associated experimental designs are described briefly below.

6.6.1 Error Reduction by Experimental Control

In this general method, an effort is made to reduce error by introducing into the experimental design one or more control variables that are believed to be related to the observations and measurements being made on components. Generally, the control variable(s) itself is not of direct concern, being incorporated into the design only so that the effects of other independent variables can be observed with greater precision.

For example, the concern may be with average rate of change of some parameter, as a function of various levels of collector voltage, e.g., a study where only Column 2 of Table 6.4 is being studied, ambient temperature and other stresses being held constant at some values. In order to conduct such a study, four different samples of components are required for each level of collector voltage. In order to obtain greater generality in any inferences made, the investigator wants to draw these four samples from, say, three different lots, since components in any one particular lot may not yield experimental results that are true of the components in general from many lots. However, the investigator also wants to somehow control for the added "scatter" that will likely be produced in the data as a result of observing components from several lots, since it is suspected that lots may differ in their characteristics, as a result of variations in the manufacturing process.

In such a situation, the experimental design shown in Table 6.6 could be employed. The investigator wishes to age 15 components under each voltage level, and to observe, say, average time rate of change of a given parameter (averaged over the 15 components)

Downloaded from https://www.everyspec.com

as a function of voltage. In order to control for lot variations, 5 components are sampled from each lot for each voltage level. Thus, in observing any effects of voltage, say on average rate of change of some parameter, any such effects are not contaminated by possible differences between lots. The alternative design would simply be to select 15 components at random from the 3 lots, for each given voltage level. However, in this case, just by chance one voltage level may contain a large proportion of components from, say, Lot 3, and another voltage level may contain a large proportion of components from, say, Lot 1, and thus observation of voltage effects may be less precise because of differences between lots.

TABLE 6.6. AN EXPERIMENTAL DESIGN TO CONTROL FOR LOT DIFFERENCES

		Collector Voltage, V_C			
		$V_c = 0$	$V_c = 10$	$V_c = 20$	$V_c = 30$
Lot Number	Lot 1	5	5	5	5
	Lot 2	5	5	5	5
	Lot 3	5	5	5	5
		15	15	15	15

The above example represents a quite simple design for experimentally controlling for error by introducing a control variable into the experiment. The control variable introduced may be any variable that is expected to influence (be correlated with) the parameter measurements made. Several more complex designs are available for the purpose of controlling error. The class of designs called randomized blocks designs[7,1] is essentially of this type. In any event, the fundamental objective is always that of incorporating into the design one or more additional control variables, which generally are not of direct concern, but added only in order to increase the precision of the effects of the independent variables of concern and to increase precision of significance tests associated with these effects.

6.6.2 Error Reduction by Statistical Control

As contrasted with error reduction by experimental control, in error reduction by statistical control a variable (or variables) is not literally incorporated into the experimental design before any data are collected. Rather, after the data are collected, statistical corrections are made for various error sources.

The major technique used here is analysis of co-variance, which is an extension of analysis of variance, with the unique purpose of increasing the precision of the experiment by statistical control. This is accomplished by statistically correcting for one (or more) sources of uncontrolled variations that produce "scatter" in the data. Then, after such corrections for error are made, the analysis is continued in a traditional way, e.g., by making various statistical significance tests.

Downloaded from https://www.everyspec.com

As an example, individual component differences in degradation rate may in part be accounted for on the basis of different initial values (the parameter value at t = 0). That is, components with a high degradation rate may exhibit a relatively high initial value, and components with a low degradation rate may exhibit a relatively low initial parameter value. In this case, by means of analysis of co-variance techniques, variation in degradation rate due to differing initial values might be "partialled out" of the data, making any comparisons of average degradation rate as a function of some stress condition more precise.

Whether or not such a technique should be used depends on the situation and whether or not certain assumptions can be met. Only a person knowledgeable in statistics, upon examining the particular experimental data at hand, can decide whether the technique would be appropriate and useful.

REFERENCES

(1) Winer, B. J., Statistical Principles in Experimental Design, McGraw-Hill Book Company, Inc. (1962).

(2) Bradley, J. V., Distribution-Free Statistical Tests, WADD Technical Report 60-661, Project No. 7184, Task No. 71581, Wright Air Development Division, Wright-Patterson Air Force Base, Ohio (August, 1960).

(3) Siegel, S., Nonparametric Statistics for the Behavioral Sciences, McGraw-Hill Book Company, Inc., New York (1956).

(4) Lewis, D., Quantitative Methods in Psychology, McGraw-Hill Book Company, Inc., New York (1960).

(5) Williams, E. J., Regression Analysis, John Wiley & Sons, Inc., New York (1959).

(6) Ezekiel, M., and Fox, K. A., Methods of Correlation and Regression Analysis, John Wiley & Sons, Inc., New York (1959).

(7) Cochran, W. G., and Cox, G. M., Experimental Designs, John Wiley & Sons, Inc., New York (1957).

(8) Gorton, H. C., Mengali, O. J., Swartz, J. M., Mace, A. E., and Peet, C. S., "Experimental and Research Work in Neutron Dosimetry: Phase I", Final Summary Report submitted to U. S. Army Signal Research and Development Laboratory by Battelle Memorial Institute, Contract No. DA 36-039SC-78924 (March 31, 1961).

(9) Thomas, R. E., and Gorton, H. C., "Research Toward a Physics of Aging of Electronic Component Parts", Physics of Failure in Electronics, 2, edited by Goldberg and Vaccaro, Cato Show Printing Company, 25-60 (1964).

Downloaded from https://www.everyspec.com

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1 ORIGINATING ACTIVITY *(Corporate author)*
BATTELLE MEMORIAL INSTITUTE
505 King Avenue
Columbus, Ohio 43201

2a. REPORT SECURITY CLASSIFICATION
UNCLASSIFIED

2b. GROUP

3. REPORT TITLE
Reliability Physics Notebook

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
Final Report

5. AUTHOR(S) *(Last name, first name, initial)*
Edited by: Vaccaro, Joseph
Gorton, H. Clay

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
October 1965	288	223

8a. CONTRACT OR GRANT NO.
AF30(602)-3504

b. PROJECT NO.
5519

c.
Task No. 551906

d.

9a. ORIGINATOR'S REPORT NUMBER(S)

9b OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
RADC-TR-65-330

10. AVAILABILITY/LIMITATION NOTICES
Releasable to DDC and CFSTI.

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY
Reliability Branch, Engineering Division
Rome Air Development Center (EMERP)
Griffiss Air Force Base, New York 13442

13. ABSTRACT

The results of physical and chemical investigations on basic mechanisms and processes in materials which cause or contribute to degradation, aging, and failure of electronic devices are summarized. Data is presented, in practical form wherever possible, for use by design and test engineers concerned with the assessment, prediction, and improvement of the reliability of solid state electronic devices. Techniques and procedures for obtaining data on specific part types and for applying such data to accelerated testing, screening, and reliability prediction programs are discussed. Format of the report is designed to facilitate periodic revision and inclusion of results of future investigations in basic failure mechanisms in electronic materials.

DD FORM 1 JAN 64 1473

UNCLASSIFIED
Security Classification

Downloaded from https://www.everyspec.com

UNCLASSIFIED

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Reliability Solid State Physics						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through _____."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through _____."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through _____."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS), (S), (C), or (U)*

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.

Security Classification