Chapter 2. Reliability Overview

R. Shaw

Reliability has been defined as the probability that an item will perform a required function under stated conditions for a stated period of time. Hence, reliability can be modeled as a probability distribution. For most semiconductor devices, a cumulative failure distribution between 0% and 100%, as shown in Figure 2-1, will be representative of its behavior over a period of time, *t*.

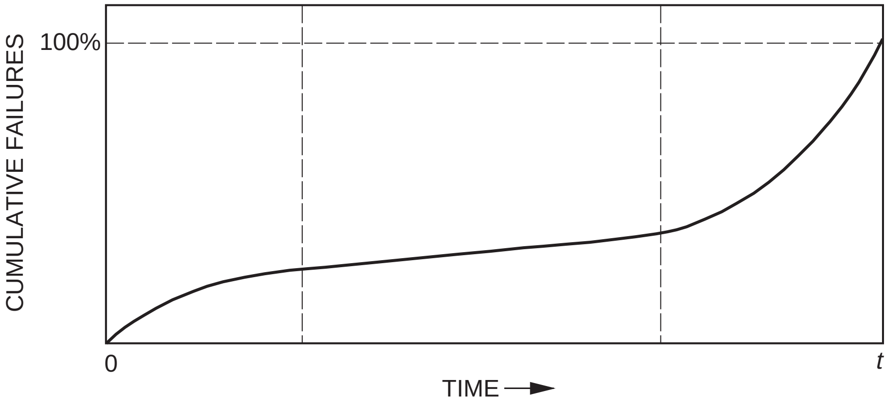

Figure 2-1. Semiconductor cumulative failure distribution.

Factors influencing the reliability of a product cover a large range of variables, including design, manufacturing, the eventual application, and the human involvement factor at each stage of production. In fact, history has shown many times that the reliability of a product from development to production follows the graph of Figure 2-2. Here the predicted or potential reliability of the product has been calculated or compared to the demonstrated reliability of a similar product. What causes the low reliability of the

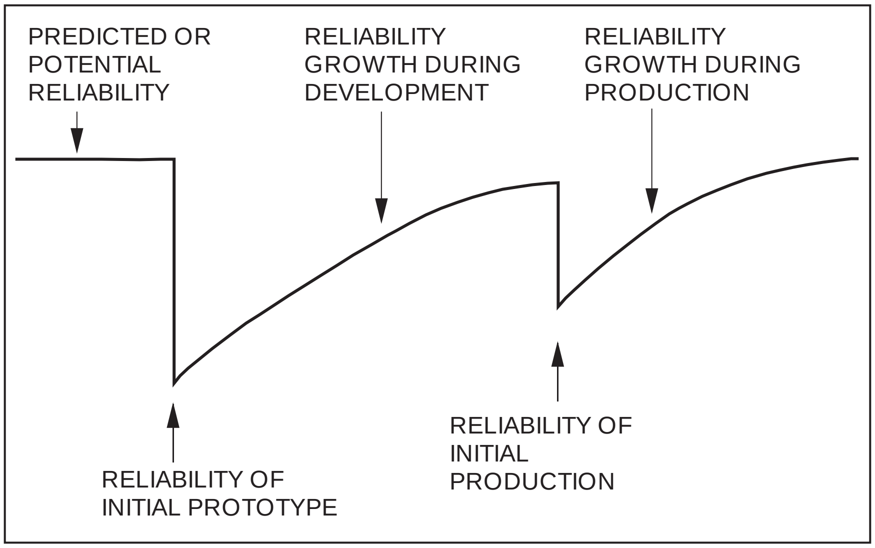

Figure 2-2. Product development cycle.

initial prototype could be a flaw in design, an unknown manufacturing process problem, the cumulative effect of several environmental stresses on the part, or a combination of several of these factors. Once the initial prototype has been produced, the reliability of the part improves as development progresses and failure mechanisms are determined and overcome. With transition of the product to a manufacturing production line, reliability usually regresses. This reduction in reliability could be caused by the change from a development or research fabrication process to a production line fabrication process. The manufacturing production line environment can be very different from the environment of the research and development pilot line. Also, at this point in the product's life cycle, the human-involvement factor is usually at its most drastic transition, causing many variances in the fabrication process.

These reliability shortcomings from design to production can be minimized by incorporating, as early in the development cycle as possible, statistical process control methodologies in the fabrication of the part and by performing life-test measurements. Implementing these techniques forces reliability growth to occur in conjunction with product development. For the manufacturer, this means a quicker time-to-market cycle of a reliable product that will not require costly warranty repair or replacement. For the user, it means that the state-of-the-art product can be confidently incorporated in his advanced system.

I. Failure

The definition of a failure is important to any analysis on semiconductor device reliability. A failure could be classified into two groups:

(1) Degradation failures, where an important parameter of a component drifts so far from its original value that the component no longer functions properly.

(2) Catastrophic failures—the end of component life; i.e., complete destruction of the component.

Part failure at any given point in time takes place when the combined effect of the stresses imposed on the part exceeds the part strength. Typical factors that influence the failure rate of semiconductors are; temperature, voltage level and polarity, complexity, base material, handling and electrostatic discharge, and humidity.

A. Physical Failure Mechanisms

Although both passive and active components of GaAs MMICs are subject to reliability problems, the active elements (such as a FET) are often the limiting factor. Ohmic contacts on FETs can be a reliability limiting factor, gradually degrading in contact resistance as diffusion acts to destroy the ohmic alloy, but the major limitations have been found to be related to the FET channel. The exact nature of these channel defects may vary, but the effects are consistent with a reduced channel thickness, as though the gate were "sinking" into the material.

Another major failure mechanism in semiconductors, both silicon and GaAs, is metal migration. Metal migration is the physical movement of metal in a conductor caused by current flow. Electron scattering from metallic atoms literally pushes these atoms in the direction of electron flow. Metal can be depleted from one part of the conductor and accumulate at a nearby part. At the depletion site, the cross-sectional area of the conductor is reduced. This increases the current density, which increases the effect even more and can lead to burnout at the thin portion of the conductor. In addition, a buildup of metallization at accumulation points can lead to shorts in metal above the buildup, such as air bridges or capacitor plates. Electromigration is the main reason the current density in metallic elements on MMICs is limited to approximately 2×10^5 A/cm^2. The "fusing" (burnout) current density is much greater.

Although the above discussion centered on FET channel failures and electromigration as major failure mechanisms in GaAs MMICs, lack of careful attention to other elements can result in severe reliability problems. Accelerated-life testing is needed to identify and remove such limitations. With these precautions, the median lifetimes of GaAs MMICs can exceed 1×10^6 h at normal operating temperatures.

Chapter 4, "Basic Failure Modes and Mechanisms," provides further information and discussion on this topic.

B. Radiation Failure Mechanisms

The ability of GaAs devices to withstand radiation is important in both space and military applications. Objects in Earth orbit are subjected to radiation from the radiation belts surrounding the planet. The cumulative dose absorbed over time can be considerable, and, of course, shielding in space applications must be minimal for the obvious weight and cost considerations. Many military applications require the ability to withstand intense radiation caused by nuclear explosions. The amount of radiation generated over a short period by nuclear events can be very high. In summary, there is interest in the ability of GaAs devices to withstand both long-term cumulative radiation and high dose rates over short periods. GaAs devices generally have greater radiation tolerance than do silicon devices, and this is one of their advantages in radiation environments.

Chapter 10, "Radiation Effects in MMIC Devices," provides further information and discussion on this topic.

II. Quantifying Reliability

Quantifying reliability is achieved from the concept of reliability as a probability distribution. The probability of a component surviving to a time *t* is the reliability, *R(t),* of the component, and is expressed as

$$R(t) = \frac{\text{number surviving at instant } t}{\text{number at time } t = 0}$$

The failure rate can be expressed as *f(t),* where

$$f(t) = \frac{\text{number failing per unit time at instant } t}{\text{number surviving at instant } t}$$

The failure rate can therefore be defined as the probability of failure in unit time of a component that is still working satisfactorily. For constant failure rate *f*, *R*(*t*) is given by

$$R(t) = \exp(-ft)$$

$R(t)$ is therefore an exponentially varying function of time, as shown in Figure 2-3.

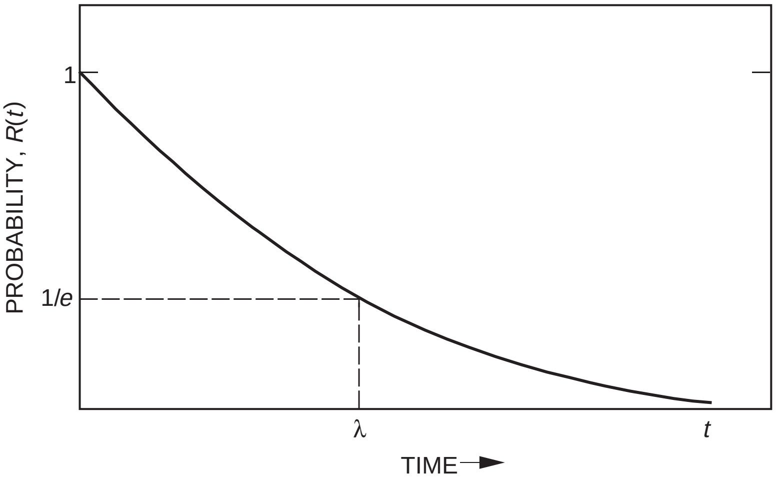

Figure 2-3. Probability of survival to time *t*, for a constant failure rate.

The failure rate, $f(t)$, is given as the number of units failing per unit time. In practice, the number of components failing per second is a fraction of a percent; to obtain more manageable values the units are scaled. Therefore, $f(t)$ may be expressed as the percent (%) failure per 1×10^6 h or as the number of devices failing in 1×10^9 h. The latter unit is known as the FIT and is commonly used as the unit of reliability:

$$1\ \text{FIT} = 1\ \text{failure}/1 \times 10^9\ \text{device h}$$

The mean number of failures in a given time is defined by the mean time between failures (MTBF) and is another commonly used method of quantifying component reliability. Assuming the failures occur randomly at a constant failure rate, the MTBF is given by

$$MTBF = 1/f$$

This may also be written as the probability of success or zero failures:

$$P(s) = e^{(-t/MTBF)}$$

where

$P(s)$ = probability of success
t = time

Figure 2-4 shows $P(s)$ versus time as normalized to *MTBF*. From this plot it can be seen that after 1/2 *MTBF*, the probability that there will be no failures is 60% and 37% after 1 *MTBF*.

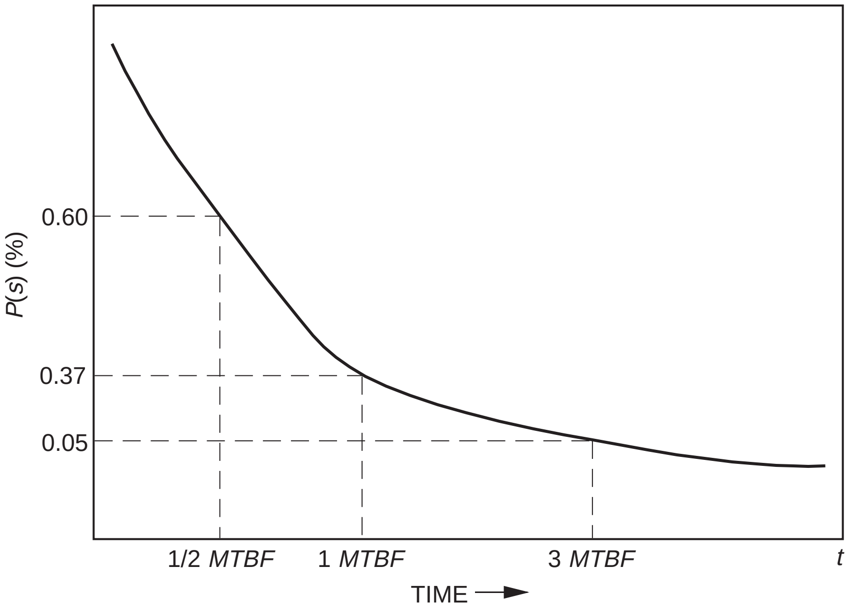

Figure 2-4. Probability of success normalized to the MTBF.

When modeling failures, confidence limits are put on the distributions indicating the extent to which the data are representative of a batch of components. For example, a large sample (> 1×10^5 devices) resulting in 1×10^3 failures in 1×10^{12} device hours would indicate a failure rate (*f*) of 1 FIT. This value of *f* would have a much higher confidence limit than that of one device operated continuously for 1×10^9 h, after which time it fails.

A common graphical interpretation of the failure rate is shown in Figure 2-5. This model is known as the "bathtub" curve and was initially developed to model the failure rates of mechanical equipment. However, it has now been adopted by the semiconductor industry and has become an integral part of semiconductor reliability theory.

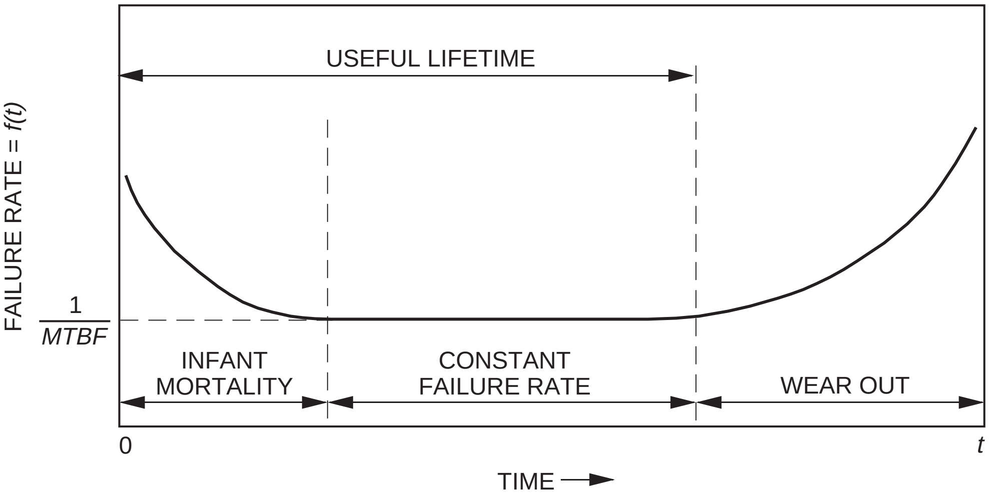

Figure 2-5. Semiconductor failure rate.

The bathtub curve in its simplest form consists of the three regions shown in Figure 2-5. The failure rate is theorized to be high at the start, dropping off as the weaker devices fail early. The failure rate then approaches a constant as the components enter their useful lifetime. Failures in this period can be attributed to random overload of the components. Finally, wear-out occurs and the curve increases sharply.

III. GaAs Device Reliability

GaAs device reliability involves probability statistics, time, and a definition of failure. Given a failure criterion, the most direct way to determine reliability is to submit a large number of samples to actual use conditions and monitor their performance against the failure criteria over time. Since most applications require device lifetimes of many years, this approach is not practical. To acquire MMIC reliability data in a reasonable amount of time, most people have used accelerated-life tests at high temperatures. By exposing the devices to elevated temperatures, it is possible to reduce the time to failure of a component, thereby enabling data to be obtained in a shorter time than would otherwise be required. Such a technique is known as "accelerated testing" and is widely used throughout the semiconductor industry. The rate at which many chemical processes take place is governed by the Arrhenius equation:

$$r = A \exp\left(\frac{-E_a}{kT}\right)$$

where

$r =$ rate of the process

$A =$ a proportional multiplier, which can be a function of temperature $\left(A = A(t)\right)$

$E_a =$ a constant known as the activation energy for a given process

$k =$ Boltzman's constant, 8.6×10^{-5} (eV / K)

This equation has been adopted by the semiconductor industry as a guideline by which the operation of devices in varying temperature conditions can be monitored. Experimental data obtained from life tests at elevated temperatures are processed via the Arrhenius equation to obtain a model of device behavior at normal operating temperatures. Rearranging the Arrhenius equation allows the temperature dependence of component failure to be modeled as follows:

$$\ln \frac{t_2}{t_1} = \frac{E_a}{k}\left(\frac{1}{T_2} - \frac{1}{T_1}\right)$$

where

$t_{1,2} =$ time to failure

$E_a =$ activation energy in electron volts (eV)

$T =$ absolute temperature in K

To properly analyze life-test data requires the adoption of a mathematical failure distribution. Several are commonly used, including the normal, lognormal, exponential, and Weibull distributions. Most of the test operators have adopted the lognormal distribution because it most closely fits the measured reliability data from life-tested GaAs semiconductor devices. The lognormal graph is a plot of normal cumulative-percent-failure versus log time. If the life-test data fit a straight line on this graph, the data fit the lognormal distribution. The intersection of this line with 50% cumulative failure indicates the median lifetime. Median life is the time it takes for half of the devices to fail. Figure 2-6 shows a typical Arrhenius plot.

To accurately predict lifetimes at normal operation temperatures, at least three different high-temperature life tests must be performed. The median life from each of the three tests is transferred to an Arrhenius plot and fit with a line. The slope of the line is the activation energy. Median life at any temperature can then be determined. Median life should not be confused with mean time to failure (MTTF). *MTTF* is the reciprocal of the instantaneous failure rate. *MTTF* is not constant with time due to the lognormal failure distribution. One must specify an operation time to calculate the exact *MTTF*. However, a close approximation to the average *MTTF* is calculated as follows:

$$MTTF = T_{op} * \exp\ (sigma^2\ /\ 2)$$

where

T_{op} = median life at the desired operating temperature

sigma = the lognormal standard deviation

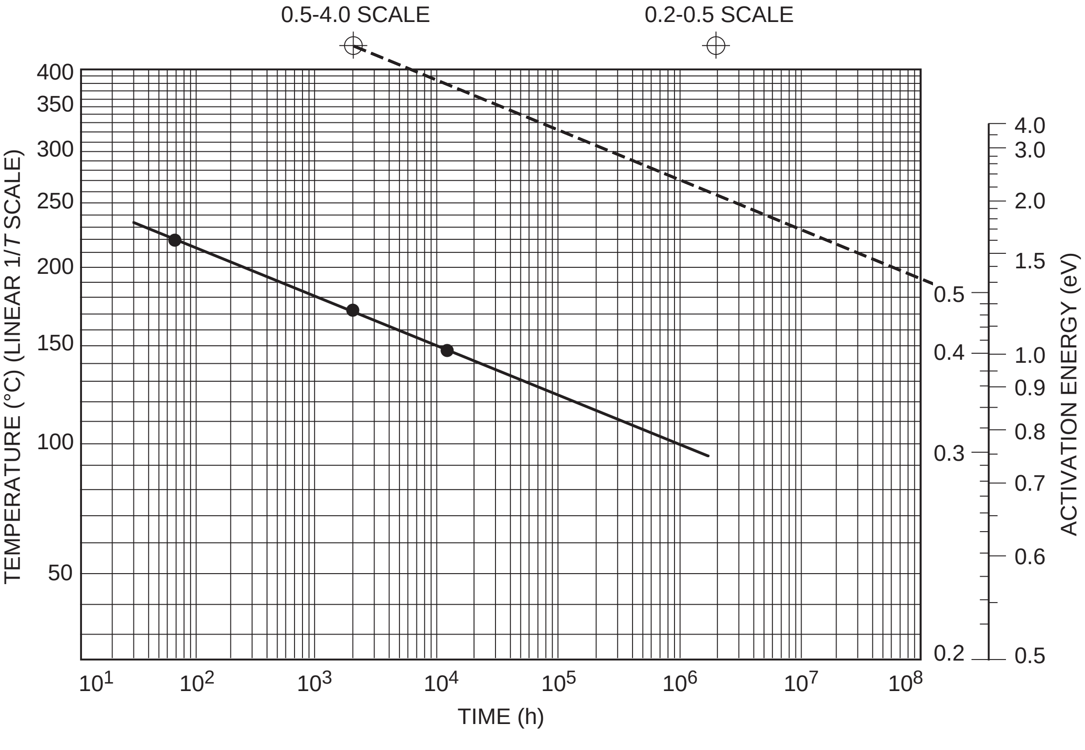

Figure 2-6. Arrhenius plot. (Courtesy of Artech House.)

Ideally, accelerated life tests should be conducted with very large sample sizes. However, this is not always practical or economical. The sample size determines the confidence in the lifetime predictions. The smaller the sample size, the less confidence we have in the prediction. Confidence limits are defined in terms of percentage. For example, an upper and lower 90% confidence limit would indicate that repeating the life test 10 times, 9 out of 10 tests would predict a median life between the two limits. Confidence limits can be calculated for median life with the following equation:

$$\text{upper limit} = T_{test} * \exp\left[\left(t(df,\ alpha) * sigma / N\right)\right]$$

$$\text{lower limit} = T_{test} * \exp\left[\left(-t(df,\ alpha) * sigma / N\right)\right]$$

where

T_{test} = median life at test temperature

$t(df,\ alpha)$ = value from students' t distribution

df = degrees of freedom $(N-1)$

$alpha$ = (1% confidence) / 2

N = sample size

It is apparent that knowledge of temperature is fundamental in obtaining accurate reliability data from accelerated temperature testing. A GaAs device or MMIC with active elements will generally have areas, such as FETs, that are far hotter than other areas. Thin-film resistors can also be significantly hotter than surrounding portions of the chip. The chemical or physical changes that lead to failure usually occur in these hotter regions. Therefore, one needs to know the temperature of these regions to obtain accurate determinations of activation energy. Of course, *MTTF* can be determined as a function of any convenient temperature, such as the base-plate temperature. However, even in this case, comparison data from differing institutions are facilitated if the temperature at the failure site, such as the FET, is used.

GaAs is a relatively poor thermal conductor; thermal conductivity of GaAs is less than one-third that of Si at room temperature. Further, the active parts of GaAs devices, such as the gate channel regions of FETs, are also very small. These two factors mean that active areas on GaAs devices can be appreciably hotter than nearby regions of the device, and significantly hotter than the ambient or base-plate temperature. The thermal conductivity of GaAs decreases with increasing temperature. This means that as the ambient or base-plate temperature increases, the temperature differences within the chip also increase. The buildup of heat at active devices is characterized by the thermal resistance of the device. The thermal resistance is defined as the temperature difference between the hottest spot and some reference spot, usually the ambient or base-plate temperature, divided by the power dissipated in the device. Therefore, thermal resistance is expressed in ^{o}C/W. Note that thermal resistance will normally vary with device size and will certainly vary as the thickness of the die.

Since most GaAs device failures occur in the FET channel, all life-test data are referenced to the channel temperature. The importance of accurately determining the channel temperature of each device submitted to life test cannot be overstressed. Variables affecting the channel temperature include ambient temperature, device thermal impedance, package and mounting materials, power dissipation, and RF levels. Extensive reliability life tests on numerous GaAs components have been performed since the early 1980s. Typical measured activation energies range from 1.2 eV to 1.9 eV.

Additional Reading

High-Power GaAs FET Amplifiers, J. L. B. Walker, Editor, Artech House, Inc., Norwood, MA, 1993.

Jensen, F., and Niels E., *Burn-In, An Engineering Approach to the Design and Analysis of Burn-In Procedures*, John Wiley & Sons, New York.

Amerasekera, E. A., *Failure Mechanisms in Semiconductor Devices,* John Wiley & Sons, New York, 1987.