JPL Publication 96-25

GaAs MMIC Reliability Assurance Guideline forSpace Applications

Sammy Kayali
Jet Propulsion Laboratory
George Ponchak
NASA Lewis Research Center
Roland Shaw
Shason Microwave Corporation
Editors

December 15, 1996

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

The research described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

Preface

This document was conceived in response to comments by various industry representatives lamenting the lack of an industry-accepted method for MMIC qualification. A low-level effort to address this problem was initiated in the summer of 1992 by individuals at the Jet Propulsion Laboratory, NASA Lewis Research Center, and NASA Johnson Space Center. These efforts were combined in July 1993 to form the MMIC Reliability Assurance Working Group, which gained the support of NASA Headquarters, Code Q.

The original concept was an official government-sponsored MMIC qualification specification describing all the required test and evaluation procedures performed by the manufacturer. This approach was presented to industry representatives at the first MMIC Qualification Workshop held at NASA Lewis Research Center in Cleveland, Ohio, in September 1993. At this meeting, various users and suppliers of MMIC devices expressed their strong desire to avoid government specifications and requested a document that would be an educational tool. The format was envisioned to be a source book of GaAs MMIC reliability and design methodology techniques useful in developing a qualification plan for the production and use of GaAs MMIC in space applications. The workshop attendees concluded that the title of the document should be GaAs MMIC Reliability Assurance Guideline for Space Applications. The guide was developed to be a practical application of industry-accepted reliability assurance practices used for the specification, manufacture, qualification, and procurement of GaAs-based MMICs.

The text contains background material on and discussion of the tests, screens, and evaluations normally conducted on MMIC devices prior to approval for use in high-reliability applications. The information is focused on the needs of the engineer, the program-level manager, and the purchaser, with the emphasis on the common approaches to GaAs MMIC reliability and qualification methodologies used and accepted in the industry.

Background information is provided on the materials, design methodology, test techniques, environment effects, common failure mechanisms, and fabrication processes—information needed to structure an effective qualification plan for the specific application required. Using this information as a common reference point, the user and the manufacturer can discuss trade-offs and determine the value-added tests necessary to realize a cost-effective qualification plan.

The guide begins with an introduction of GaAs usage and brief summary of MMIC development history. This is followed by a reliability overview and a summary description of reliability theory. These chapters give the reader an understanding of the usage of GaAs devices in various applications and provide the background necessary to understand reliability test results and the implication of failure.

GaAs material properties and common device structures used in MMIC designs are discussed in Chapter 3. This chapter also provides general descriptions of the common processes and the various general-purpose MMIC functions and circuits. Chapter 4 provides descriptions of the common failure modes and mechanisms

affecting GaAs-based device; this information can be of great importance in developing characterization and qualification plans.

Device modeling and MMIC design methodology are discussed in Chapters 5 and 6, respectively. These chapters provide general information needed to understand the various aspects of MMIC design. The text also addresses the reliability aspects of MMIC design and provides a generalized design methodology useful to both the user and manufacturer.

Chapter 7 discusses MMIC testability and provides examples of test implementations. General test structures and process monitors employed at various stages in the manufacture of MMIC devices are also presented. Qualification methodologies are discussed in Chapter 8, along with approaches to the different aspects and levels of device development and qualification. The significance of package reliability is given in Chapter 9, together with brief descriptions of the common packaging materials and their related effects.

Finally, Chapter 10 addresses radiation effects on GaAs MMICs and discusses the radiation environments experienced during space flight and their effects on device performance.

The information contained in this document has been collected from users and manufacturers through direct interaction and collaboration. For example, the approaches to process and product acceptance, as presented in Chapter 8, were collected and compiled to present a simpler way of addressing the subject. However, this information is presented only as a suggested approach and should be modified to accommodate the different methodologies practiced by the manufacturers.

The NASA MMIC Reliability Assurance Working Group:

Sammy Kayali
George Ponchak
Roland Shaw

October 1996

Abstract

This guide is a reference for understanding the various aspects of monolithic microwave integrated circuits (MMIC). There are special emphases on the reliability aspects of MMIC devices. GaAs material properties and common device structures along with the applicable failure mechanisms are addressed in detail. MMIC design and qualification methodologies provide the reader with the means of developing suitable qualification plans. Radiation effects on GaAs devices and packaging effects on MMIC device reliability are discussed with supporting references.

Contributors

Lisa Aucoin	Raytheon Company
Charles Barnes	Jet Propulsion Laboratory
Chung-Hsu Chen	TRW
Y. C. Chou	University of California, Irvine
Alan N. Downey	NASA Lewis Research Center
Robert Ferro	Aerospace Corporation
Anthony Immorlica	Lockheed Martin Company
Wen Yan Jiang	MPB Technologies, Inc.
Sammy Kayali	Jet Propulsion Laboratory
George Ponchak	NASA Lewis Research Center
Edward Rezek	TRW
Robert R. Romanofsky	NASA Lewis Research Center
Luis Selva	Jet Propulsion Laboratory
Roland Shaw	Shason Microwave Corporation
Rainee Simons	NASA Lewis Research Center
Tien Trinh	TRW

Contents

Tables

Figures

Chapter 1. Introduction

R. Shaw

This chapter establishes a common reference for the varied backgrounds of the readers. It discusses reliability and quality assurance in general and reviews the effects of new technology on the failure-rate distribution of the product. It also gives the reader an overview of why gallium-arsenide (GaAs) is used and a brief summary of the development of the monolithic microwave integrated circuit.

I. Why GaAs is Used

Perhaps the primary benefit of GaAs comes from its electron-dynamic properties. In equivalently doped n-type GaAs and silicon, the effective mass of the electric charge carriers in GaAs is far less than that in silicon. This means that the electrons in GaAs are accelerated to higher velocities and therefore transverse the transistor channel in less time. This improvement in electron mobility is the fundamental property that enables higher frequencies of operation and faster switching speeds.

While the principal reason for making transistors out of GaAs is greater speed in performance, which is realized either as a higher maximum frequency of operation or higher logic switching speeds, the physical and chemical properties of GaAs make its use in transistor fabrication difficult. Most of the early development in solid-state electronic devices centered on silicon- and germanium-based materials because of the relative ease with which the material could be processed. Silicon and germanium are elemental semiconductor materials, whereas GaAs is a binary compound. This is the root fact that caused many technical obstacles in the use of GaAs. Other properties not in GaAs' favor for early solid-state device development included a lower thermal conductivity and a higher coefficient of thermal expansion than silicon and germanium. However, as new market applications demanded higher performances that could be achieved only with the superior electron dynamics of GaAs, these obstacles have been overcome.

The markets that drove the breakthroughs in material-growth and device-fabrication techniques of GaAs semiconductors were the defense and space industries. These industries required systems with higher frequency circuits for radars, secure communications, and sensors. Many federal agencies put in place programs to develop GaAs devices as primary products in their systems. The maturity of GaAs led to the emergence of new commercial markets, such as wireless local area networks (WLANs), personal communication systems (PCSs), direct broadcast satellite (DBS) transmission and reception by the consumer, global positioning systems (GPSs), and global cellular communication. These commercial markets required the insertion of GaAs technology to meet system performances not attainable with silicon and germanium. Some of the advances achieved with GaAs technology included the use of higher frequencies to avoid spectrum crowding, new digital transmission techniques that require linear amplifiers at higher RF power levels, and lower voltage/lower current amplifiers to maximize the operating and standby times of equipment that had to be powered by batteries. In some instances, GaAs is the system "enabler," without which there would be no product or service to sell. Although these emerging markets offer advanced services and products to the consumer, several limitations to their acceptance over silicon-based systems exist. One drawback is that the failure mechanisms and reliability of silicon are better understood than those of GaAs. Another drawback is the cost and availability of GaAs

when compared to silicon. The use of silicon in lower frequency analog circuits and in very large scale integration (VLSI) technology has developed proven practices and a strong production base for the semiconductor industry. This manufacturing maturity equates to a lower cost for silicon-based rather than GaAs-based technology. However, when the cost to manufacture is compared to performance, the value added to the system by the GaAs technology in most cases more than pays for the increased fabrication cost. As the WLAN, PCS, DBS, GPS, and cellular markets grow, the cost to manufacture GaAs will decrease, and the issue of using GaAs rather than silicon will hinge on the ability of GaAs to satisfy the technical needs of the marketplace.

II. Hybrid and Monolithic Integrated Circuits

From 1930 to 1960, microwave or high-frequency technology consisted of circuits manufactured using waveguide: rectangular hollow metal pipes that "guided" the electromagnetic energy to its destination. The design was usually experimental and the production was generally expensive and long. At that time, the microwave engineer was known as a "plumber" and his tool of trade was a hammer. Around 1960, the development of semiconductors in "planar" geometries and the production of cheap, low-loss dielectric materials were the beginnings of the microwave integrated circuit (MIC). This technology was later called hybrid microwave integrated circuitry because the active devices (such as diodes and transistors) and some of the passive elements (resistors, capacitors, and inductors) were discrete components mounted to a dielectric slab or substrate. The MIC utilized metal transmission lines that were photolithographically etched onto the substrate to guide the electromagnetic energy to various components of the circuit. The performance approached the design prediction better than the waveguide predecessor, but many perturbations in the line geometries and inconsistent material properties caused much of the final circuit layout to be experimentally determined. Other factors that made hybrid-circuit production difficult were the labor-intensive processes of assembly and electrical performance testing. The assembly process required mounting each individual discrete device on the substrate, and, because of variations in component placement, the electrical test operation required labor to tune the circuit performance. The attachment of devices to the substrate and the tuning techniques required to make them perform became an art form and a hard process to control. Eventually, at higher and higher frequencies, these processes became the limiting constraints to performance, cost, yield, and reliability.

The idea of a "monolithically" integrated circuit—where the active and passive components are formed on the substrate—eliminated many of the problems with hybrid integrated circuits. The monolithic microwave integrated circuit (MMIC) uses an insulating crystalline material as both the dielectric and the active layer material. For many new applications, GaAs has become the material of choice because of its ability to perform at high frequencies. It also has a high-resistivity semi-insulating property that reduces cross talk between devices. This permits the integration of active (radio-frequency) devices, control (logic) devices, transmission lines, and passive elements on a single substrate.

Unlike the hybrid MICs, a GaAs MMIC's performance cannot be easily "tuned" by adjusting lumped or distributed elements. Once the circuit is processed, its performance is, for the most part, set. Therefore, the design of the MMIC must be based on accurate physical and electrical models for both the passive and active elements, including effects due to manufacturing process tolerances. This design process uses powerful interactive software programs for the synthesis, analysis, and layout of linear

and nonlinear circuits. Development of this software capability has matured since the 1970s, and improvements continue to be made as the technology matures. Many manufacturers have "libraries" of existing device models, which have allowed the MMIC designer to realize the desired performance without having to experimentally characterize the device.

In comparison to the other forms of microwave technology previously discussed, GaAs MMICs offer the following advantages:

(1) Size and weight reduction.

(2) Cost reduction for medium- to large-scale production volumes.

(3) Enhanced system performance from the inclusion of several functions (e.g., RF and logic) on a single circuit.

(4) Enhanced reproducibility from uniform processing and integration of all parts of the circuit.

(5) Enhanced reliability from integration and process-control improvements.

(6) Wider frequency-bandwidth performance from the reduction of parasitics in discrete device packaging.

(7) Design performance realized without several iterations—the result of processing and material repeatability, and computer-aided design enhancement.

III. Reliability and Quality Assurance

For any application, the user of the part wants the assurance that the part will continue to function correctly over a given time and under certain environmental conditions. Part failure at any given time takes place when the combined effect of the stresses imposed on the part exceeds the part strength. These statements allude to the time dependency of both part reliability and user expectation. For example, an expendable system might have a useful life of 1 minute while a satellite system must have a predicted life of several years. Each user has a different expectation of part reliability and a different level of commitment to pay for the assurance that the part will meet the expectation. Traditionally, the procurement of highly reliable (hi-rel) parts meant that the user of the component specified to the manufacturer additional requirements to be met in the fabrication of the part. These specifications were usually in terms of recording fabrication process steps, performing additional visual inspections, and incorporating additional screens and burn-in tests. The user of the hi-rel part was expected and usually willing to pay the cost for this increased reliability and quality assurance.

To understand what the user bought with this additional testing requires an investigation of the nature of part failures and their causes or failure mechanisms. It is well documented in the reliability discipline that most products experience a disproportionate number of failures in the early period of their service. This phenomenon, typically referred to as "infant mortality," can arise with any stress applied to the device (e.g., temperature, environmental, and voltage stress). Usually the infant failures occur because of a manufacturing process irregularity that decreased the product's strength in proportion to the strength of the stress imposed. Once this population of infant failures passes, the remaining units have a failure distribution centered around an expected value for a given stress. The long-term failure distribution is

usually determined by the chemical and physical properties associated with the technology, design, material used in the product, and, most importantly, the total strength of the environmental stresses imposed.

As an example, consider the elasticity of a rubber band. As the rubber band is stretched, it will eventually snap. The strength of the force required to snap the band can be recorded. If this process is repeated on several similar rubber bands, there will be a slight variation of the strength required to break each one. This distribution in failure is caused by slight variations in the manufacture of the different rubber bands. The mean value of strength for breakage is determined by the technology, design, material used in the band, and, to a greater degree, the environment (e.g., temperature and humidity) to which the band is exposed. In addition, you would find a few rubber bands that would snap under a very small amount of stress. These unusually weak bands when analyzed would probably show a defect caused in the manufacturing process, such as a thinner band, or a hole in the material, or a slight tear. These defective units are the infant-mortality population. If the fabrication processes are monitored and kept in control, the number of infant failures is reduced.

This bimodal distribution model is applicable to many types of components, including semiconductors, and, with this model, the value of the additional cost associated with a hi-rel part specification can be understood. Traditionally, a user procuring a hi-rel part would specify that the manufacturer perform stress tests to delete the infant-failure population from the delivered units. In addition, the user of the hi-rel part would specify that the manufacturer control the fabrication processes to reduce the total size of the infant-mortality population. With regard to the mean-failure, the hi-rel user usually specified that the manufacturer calculate a predicted mean-time failure-rate figure of merit. Generally, with a mature technology and manufacturing process such as that used with silicon, this calculation was a meaningful estimate based on the complexity of the circuit and the normal environmental conditions imposed on the circuit during operation. This traditional method of hi-rel procurement using individual part specifications was effective in achieving a part that was more reliable than a commercial product. The additional cost incurred to achieve this level of reliability and quality assurance was considered justified and required to achieve a confidence that a system failure would not occur during the mission.

However, today the large commercial markets of WLAN, PCS, DBS, GPS, and cellular telephony demand the quality and reliability of the hi-rel user but at consumer prices. These high-volume markets have impacted the business philosophy of the semiconductor manufacturer. Many of the manufacturers now fabricate their standard commercial product line utilizing statistical process control for repeatability and uniformity. This has greatly reduced the infant-mortality population without having to impose the hi-rel part specification. Hi-rel users can take advantage of this industry change to decrease the cost of part procurement without adversely affecting the reliability and quality assurance of system performance.

Inasmuch as semiconductor manufacturers have reduced the infant mortality population by improving repeatability in fabricating the devices, the long-term failure mechanisms of GaAs cannot be assumed to be predictable based on silicon technology. The hi-rel user must understand that many of the failure mechanisms associated with silicon devices do not apply to GaAs MMICs, and new device structures bring new failure mechanisms. Many of the traditional assumptions for mean-time failure rate predictions do not hold for these new devices. Thus, today's hi-rel user must be more aware of measurement-based predictions of long-term failure rate over calculation-based predictions. This usually impacts the procurement of the hi-rel part by including a

measurement demonstrating the long-term reliability of the technology to be used. Typically this can be done by some method of accelerated-life test.

This guidebook proposes a hi-rel qualification methodology that does not utilize the product specification philosophy. Rather, the philosophy proposed envisions a methodology that includes a process qualification and a product qualification. The process qualification involves the verification of statistical process control to insure consistent fabrication from device to device. The process qualification works with the manufacturer's knowledge of how to produce a reliable part within the standard processes of the production line. This enables a lower cost and a shorter delivery time. The product qualification is a validation of the circuit to perform to a minimum performance under stress and environmental conditions. It usually includes a measurement demonstrating the failure rate of the part. With this two-part qualification plan, the technology, the fabrication, and the part are verified to meet the expected level of quality and reliability.

Additional Reading

Deyhimy, I., "Gallium Arsenide Joins the Giants," IEEE Spectrum, Vol. 32, No. 2, pp. 33–40, February 1995.

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.

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.*

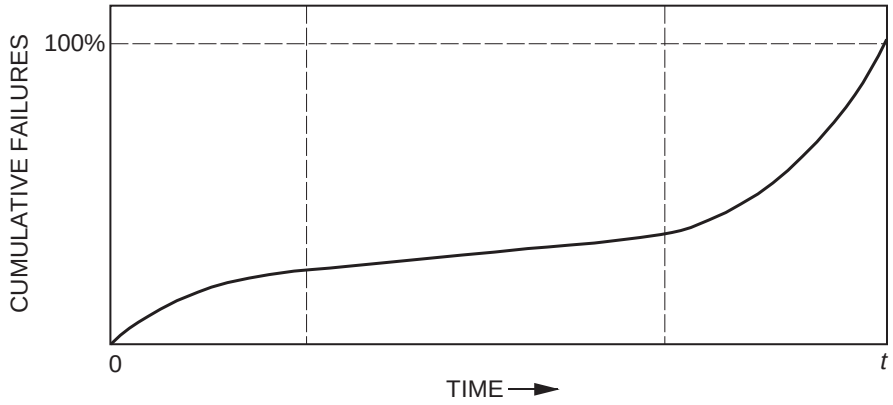

Figure 2-1. Semiconductor cumulative failure distribution.

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

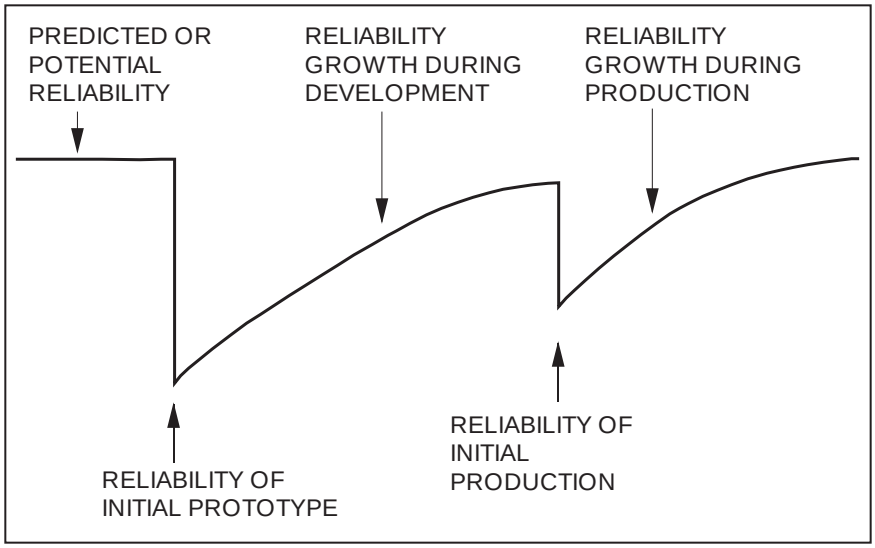

Figure 2-2. Product development cycle.

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.

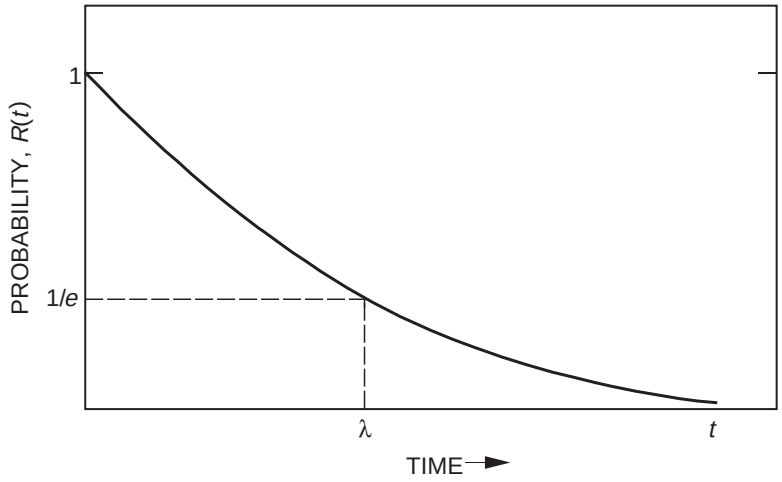

Figure 2-3. Probability of survival to time *t*, for a constant failure rate.

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*.

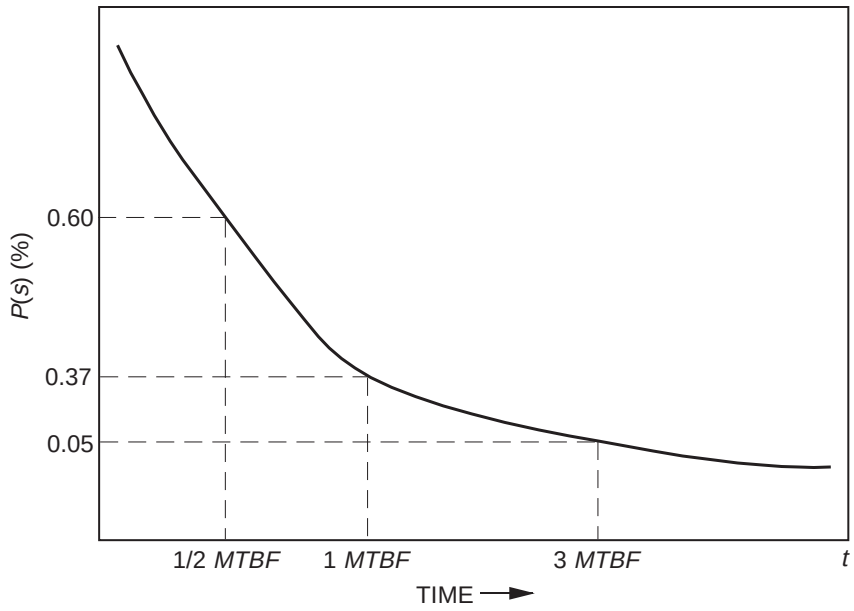

Figure 2-4. Probability of success normalized to the MTBF.

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 ($> 1 \times 10^5$ 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.

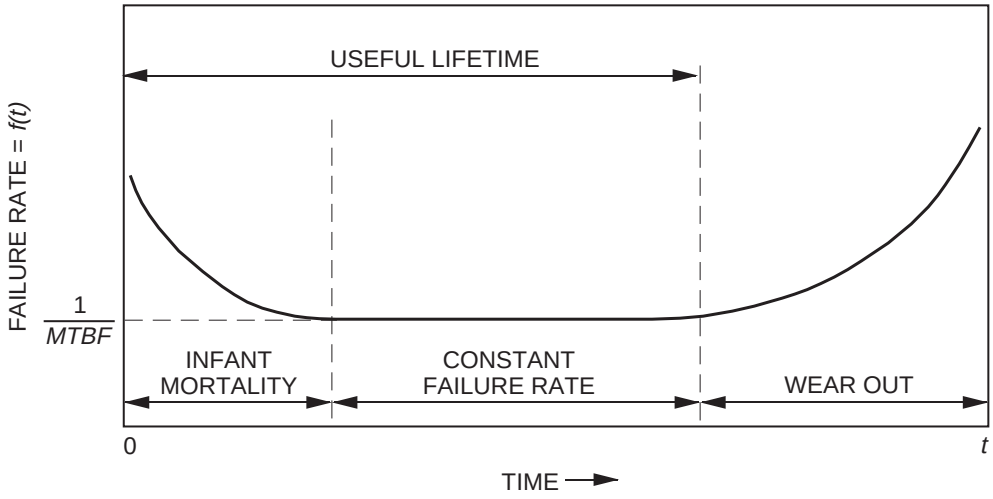

Figure 2-5. Semiconductor failure rate.

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 Weilbull 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

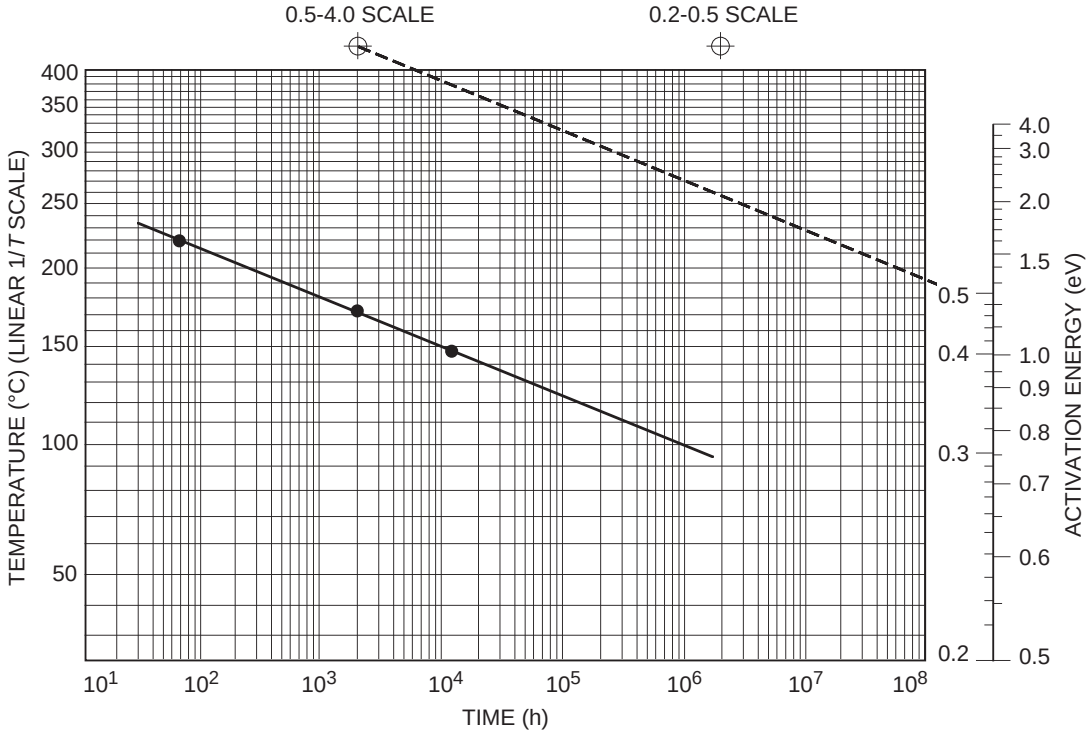

Figure 2-6. Arrhenius plot. (Courtesy of Artech House.)

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.

Chapter 3. GaAs Properties, Device Structures, and Circuits

L. Aucoin, Y. C. Chou, A. N. Downey, R. Ferro, S. Kayali, G. E. Ponchak, and R. R. Romanofsky

I. GaAs Material Properties

S. Kayali

GaAs is a III–V compound semiconductor composed of the element gallium (Ga) from column III and the element arsenic (As) from column V of the periodic table of the elements. GaAs was first created by Goldschmidt and reported in 1929, but the first reported electronic properties of III–V compounds as semiconductors did not appear until 1952 [1].

The GaAs crystal is composed of two sublattices, each face centered cubic (fcc) and offset with respect to each other by half the diagonal of the fcc cube. This crystal configuration is known as cubic sphalerite or zinc blende. Figure 3-1 shows a unit cube for GaAs and Table 3-1 provides a listing of some of the general material characteristics and properties.

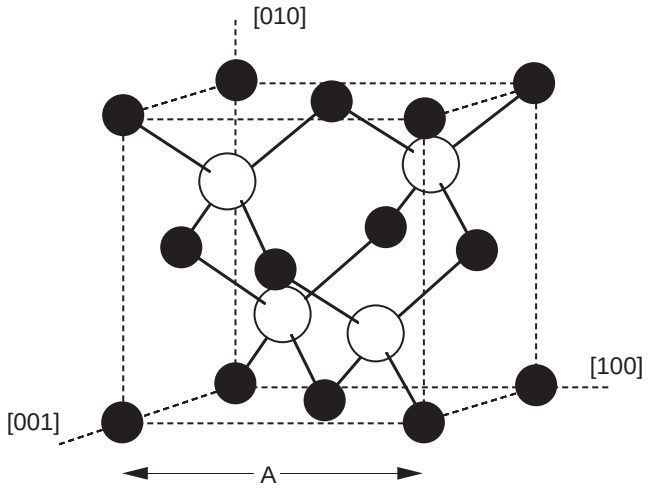

Figure 3-1. Unit cube of GaAs crystal lattice.

A. Energy Band Structure

As a result of the laws of quantum mechanics, electrons in isolated atoms can have only certain discrete energy values. As these isolated atoms are brought together to form a crystal, the electrons become restricted not to single energy levels, but rather to ranges of allowed energies, or bands called the valance and conduction bands (Figure 3-2). These two bands are separated by an energy band gap, which is a very important characteristic of the semiconductor material. At zero kelvin, all the electrons are confined to the valance band and the material is a perfect insulator. Above zero kelvin, some electrons have sufficient thermal energy to make a transition to the conduction band where they are free to move and conduct current through the crystal. The probability of an electron having enough energy to make the transition is given by the Fermi distribution function. The Fermi level shown on Figure 3-2 is the energy level at which the probability function is equal to one half. For pure semiconductors, the Fermi level is approximately in the center of the band gap. Note, though, that no electron actually has an energy of E_F, since they are not permitted to exist at energies in the band gap. The amount of energy required for an electron to move from the valance band to the

Table 3-1. Room-temperature properties of GaAs.

Property	Parameter
Crystal structure	Zinc blende
Lattice constant	5.65 Å
Density	5.32 g/cm^3
Atomic density	4.5×10^{22} atoms/cm^3
Molecular weight	144.64
Bulk modulus	7.55×10^{11} dyn/cm^2
Sheer modulus	3.26×10^{11} dyn/cm^2
Coefficient of thermal expansion	5.8×10^{-6} K^{-1}
Specific heat	0.327 J/g-K
Lattice thermal conductivity	0.55 W/cm-°C
Dielectric constant	12.85
Band gap	1.42 eV
Threshold field	3.3 kV/cm
Peak drift velocity	2.1×10^7 cm/s
Electron mobility (undoped)	8500 cm^2/V-s
Hole mobility (undoped)	400 cm^2/V-s
Melting point	1238°C

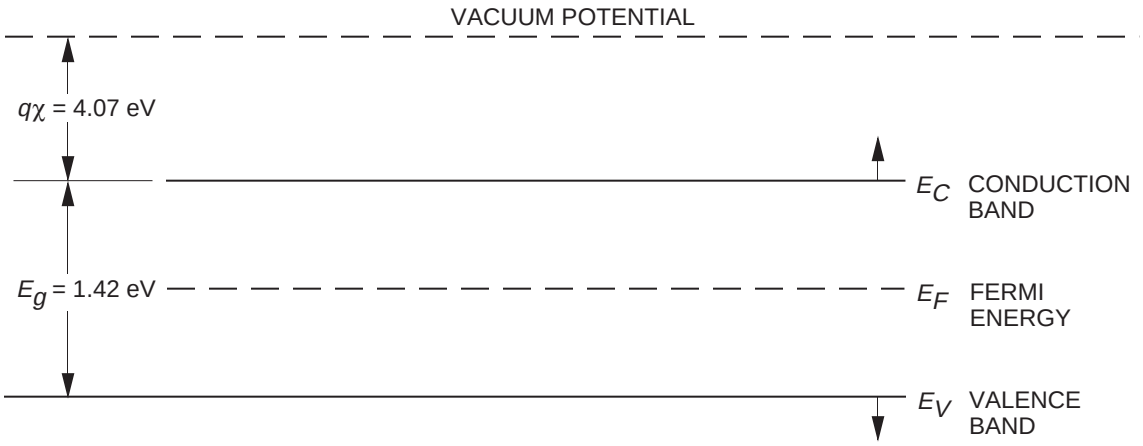

Figure 3-2. Energy band diagram for GaAs.

conduction band (energy band gap) depends on the temperature, the semiconductor material, and the material's purity and doping profile. For undoped GaAs, the energy band gap at room temperature is 1.42 eV. The energy band diagram is usually referenced to a potential called the vacuum potential. The electron affinity, $q\chi$, is the energy required to remove an electron from the bottom of the conduction band to the vacuum potential. For GaAs, $q\chi$ is approximately 4.07 eV [2,3].

GaAs is a direct band gap semiconductor, which means that the minimum of the conduction band is directly over the maximum of the valance band (Figure 3-3). Transitions between the valance band and the conduction band require only a change in energy, and no change in momentum, unlike indirect band-gap semiconductors such as silicon (Si). This property makes GaAs a very useful material for the manufacture of light emitting diodes and semiconductor lasers, since a photon is emitted when an electron changes energy levels from the conduction band to the valance band.

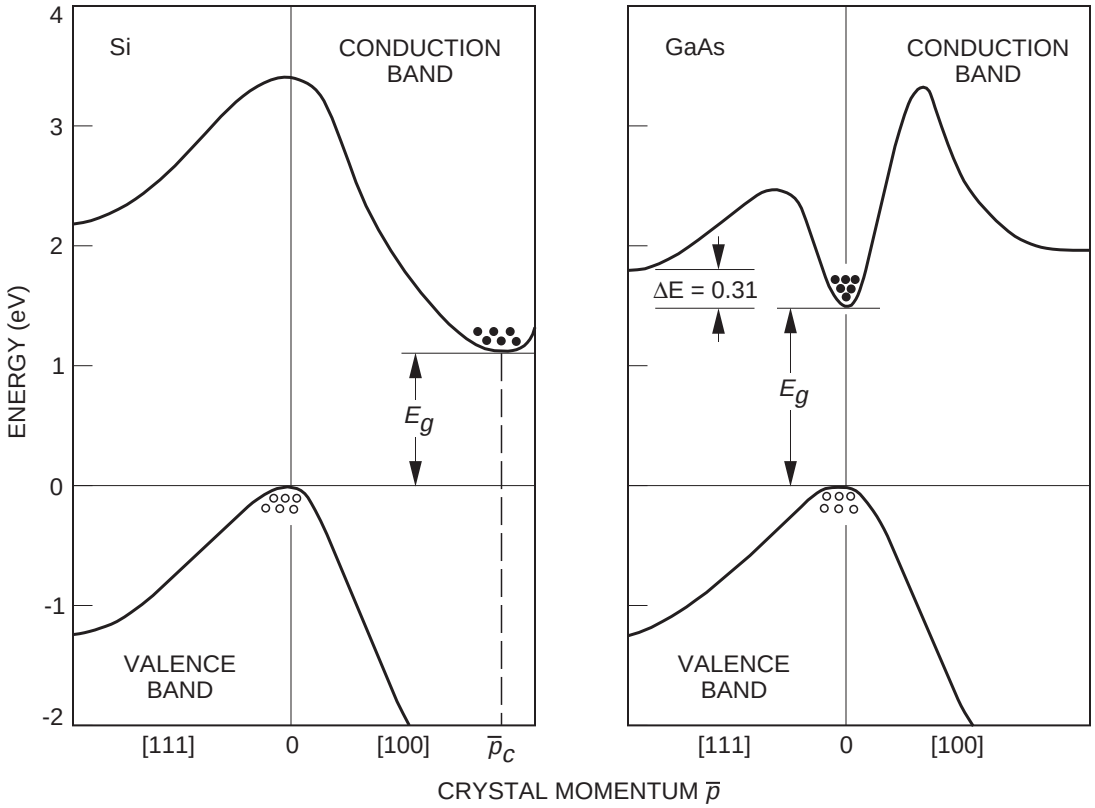

Figure 3-3. Energy band structure of Si and GaAs.

Alternatively, an incident photon can excite an electron from the valence band to the conduction band, allowing GaAs to be used in photo detectors.

B. Mobility and Drift Velocity

GaAs has several advantages over silicon for operation in the microwave region—primarily, higher mobility and saturated drift velocity and the capability to produce devices on a semi-insulating substrate.

In a semiconductor, when a carrier (an electron) is subjected to an electric field, it will experience a force ($\mathbf{F} = -q\mathbf{E}$) and will be accelerated along the field. During the time between collisions with other carrier ions and the semiconductor lattice, the carrier will achieve a velocity that is a function of the electric field strength. This velocity is defined as the drift velocity (v). From the conservation of momentum, it can be shown that the drift velocity (v) is proportional to the applied electric field (Figure 3-4) and can be expressed as

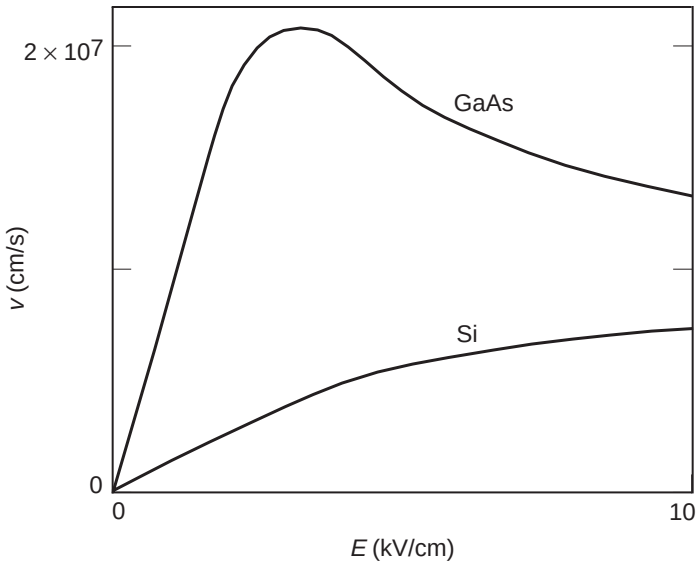

Figure 3-4. Drift velocity of electrons in GaAs and Si as a function of the electric field.

$$v = -\left(\frac{q\tau_c}{m^*}\right)\mathbf{E} \tag{3-1}$$

The proportionality factor depends on the mean free time between collisions (τ_c) and the electron effective mass (m^*). The proportionality factor is called the electron mobility (μ) in units of $cm^2/V\text{-}s$.

Mobility is an important parameter for carrier transport because it describes how strongly the motion of an electron is influenced by an applied electric field. From the equation above, it is evident that mobility is related directly to the mean free time between collisions, which in turn is determined primarily by lattice scattering and impurity scattering. Lattice scattering , which is a result of thermal vibrations of the lattice, increases with temperature and becomes dominant at high temperatures; therefore, the mobility decreases with increasing temperature. Impurity scattering on the other hand, which is a result of the movement of a carrier past an ionized dopant impurity, becomes less significant at higher temperatures [2].

Although the peak mobility of GaAs in the linear region can be as much as six times greater than that of silicon (Si) at typical field strengths, the advantage of GaAs may be only as much as a factor of two [4]. This still translates to the fact that GaAs devices can work at significantly higher frequencies than Si. The exact increase in the speed of operation depends on factors such as the circuit capacitance and the electric field regime in which the device operates.

C. Semi-Insulating GaAs

The importance of semi-insulating GaAs is based on the fact that devices made of it by direct ion implantation are self-isolating, so that it is ideally suited to integrated circuit fabrication. Moreover, the semi-insulating substrate provides greatly reduced

parasitic capacitances, thus faster devices, and allows for integration and the implementation of monolithic microwave integrated circuits (MMIC).

Semi-insulating GaAs must meet the following requirements to provide semiconductor quality material:

(1) Thermal stability during epitaxial growth or anneal of ion-implanted active layer.

(2) Absence of undesirable substrate active layer interface effects, such as back-gating and light sensitivity.

(3) No degradation of active layer properties by outdiffusion of impurities from substrate during thermal processing.

(4) Lowest possible density of crystalline defects, such as dislocations, stacking faults, and precipitates.

To achieve some of these requirements, buffer layer technology was developed. A buffer layer is a relatively thick, high-resistivity epitaxial layer grown on the semi-insulating substrate. Another epitaxial layer is then grown on the buffer layer and used for the active layer. The buffer layer provides a physical barrier for undesirable substrate impurities and imperfections.

GaAs bulk resistivity can range from 10^{-6} Ω-cm to about 10^{22} Ω-cm, with the practical range being 10^{-3} Ω-cm to 10^{8} Ω-cm. This high resistivity is about six orders of magnitude greater than that of silicon and provides excellent isolation and substrate insulation. Undoped GaAs can be made semi-insulating by the addition of either oxygen or chromium to the melt. The resistivity of the semiconductor can be controlled by counter doping with a deep-level impurity that has a conductivity type opposite to that of the impurities introduced during growth.

D. Crystal Defects

No semiconductor crystalline material is perfect, and GaAs crystals, in spite of the efforts to control crystal growth, contain a number of crystal defects, dislocations, and impurities. These defects can have either desirable or undesirable effects on the electronic properties of GaAs. The natures of these defects and the observed effects are determined by the method of their incorporation into the material and the general growth conditions.

1. Point Defects

Localized defects of atomic dimensions, called point defects, can occur in an otherwise perfect crystal lattice. These point defects can include vacancies, interstitials, misplaced atoms, intentionally introduced dopant impurities, and impurities introduced inadvertently during the material growth process. The study of point defects is important because of the effect these defects have on the electronic properties of the material and the strong relationship between diffusion and the number and type of defects in the crystalline material. The electrical properties of a semiconductor can be manipulated by the deliberate insertion of chemical defects (impurities) into the material during the growth and processing steps. However, intrinsic defects present in the material also play an important role in the electronic behavior of GaAs.

Many intrinsic defects are observed in GaAs. The concentration and effect of these defects are determined by the manner in which the material is grown. Intrinsic defects in GaAs include both arsenic and gallium vacancies, their concentration being determined by the overpressure of arsenic during processing. The effect of these vacancy defects has been observed to be neutral [5], deep donor-like, and deep acceptor-like [6].

EL2, an important defect in GaAs, is present in material grown from an arsenic-rich melt. This defect is donor-like in character and is located at the middle of the energy gap [7]. It is thermally very stable and can withstand processing temperatures up to 900°C, and acts as an electron trap. The importance of this defect lies in its ability to convert p-type GaAs to semi-insulating material, and its thermal stability.

2. Dislocations

A dislocation is a one-dimensional array of point defects in an otherwise perfect crystal. It occurs when the crystal is subjected to stresses in excess of the elastic limit of the material. Dislocations interact with chemical and other point defects. This interaction exists between the localized impurity atoms and the strain field in the vicinity of the dislocations. The presence of a dislocation is usually associated with an enhanced rate of impurity diffusion leading to the formation of diffusion pipes. This effect translates to the introduction of trapping states in the band gap, altering the etching properties of the wafer, and, most importantly, altering the electrical properties of the devices. Studies have shown detrimental effects of dislocations and dislocation densities on the source drain current and threshold voltage of field-effect transistors FETs [8,9], carrier concentration, and sheet resistance [10].

Dislocations generally are introduced as a result of a temperature gradient present during crystal growth. Modern crystal growth methods can routinely produce 7.6-cm (3-in.) wafers with dislocation densities of 10^4 to 10^5 cm^{-2} for the Liquid Encapsulated Czochralski (LEC) and 8000 to 25,000 cm^{-2} for Horizontal Bridgeman (HB) techniques.

3. Impurities in GaAs

Chemical point defects (doping impurities) can be introduced to the crystalline material either deliberately or inadvertently as contamination during processing. In general, substitutional impurities are electronically active, whereas many contaminants are interstitial in nature and are electronically inactive. Dopants are classified as either donors or acceptors. A donor has one more electron than the atom it is replacing in the crystal. This extra electron is easily removed or donated to the conduction current. An acceptor, on the other hand, has one less electron than the atom it is replacing. Thus, an acceptor can easily capture an electron and prevent it from adding to the conduction current. Regardless of the type or character of the impurity, the electrical properties of the semiconductor are altered.

Figure 3-5 shows the energy band diagram of Figure 3-2 with the addition of impurities. Shallow donor or acceptor impurities have energy levels within $3kT$ of the conduction and valance band, respectively. Since the energy required for an electron to transition from these impurity energy levels to the nearest band edge is very small, they are typically fully ionized at room temperature. The Fermi level shifts from the band center towards the impurity levels to reflect this. In other words, for donor impurities, the Fermi level shifts towards the conduction band, and V_{CF} decreases as the donor doping concentration increases. A similar description can be made of acceptor impurities. It is

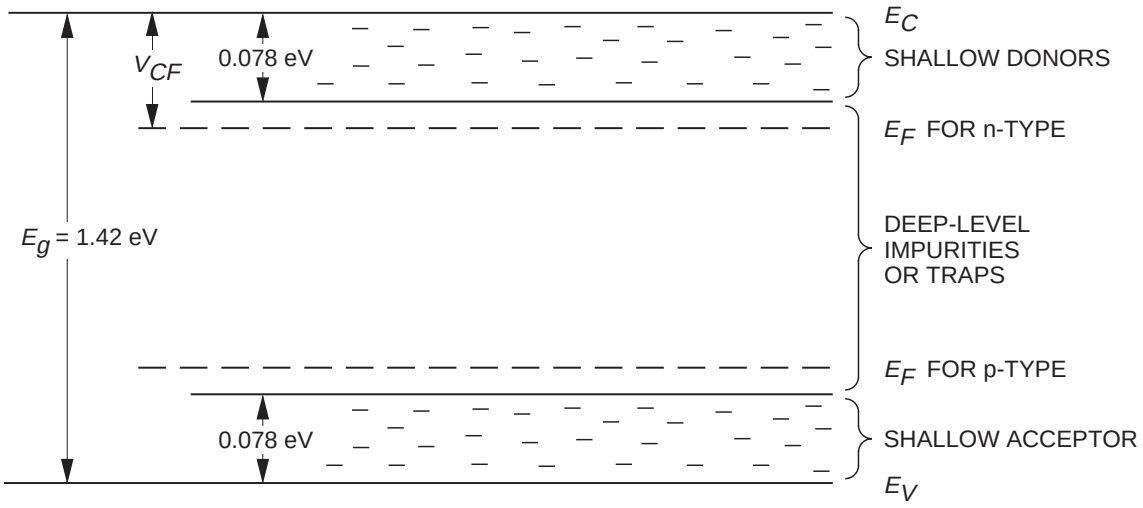

Figure 3-5. Energy band diagram of GaAs with impurities.

these shallow impurities that are used for doping purposes. Impurities with energies in the center of the band gap are called deep impurities. Deep impurities generally degrade device performance by reducing the carrier lifetime.

Both impurity types, deep and shallow, are present in GaAs in the form of complexes with gallium or arsenic. One of the most common is silicon. This group IV element can be used to give either p-type GaAs by incorporating it at low temperatures, or n-type GaAs by processing it at high temperatures. Another group IV element, carbon, is also used extensively to provide p-type GaAs. Chromium (Cr) behaves as an acceptor, with an impurity level close to the center of the energy gap. This property makes it very useful for counterdoping n-type GaAs to make it semi-insulating. Other elements such as copper, oxygen, selenium, and tin are also used in GaAs processing to provide the desired n- or p-like behavior.

E. Thermal Characteristics

GaAs has a thermal conductivity of 0.55 W/cm-°C, which is about one-third that of silicon and one-tenth that of copper. As a consequence, the power handling capacity and therefore the packing density of a GaAs integrated circuit is limited by the thermal resistance of the substrate. The reliability of GaAs devices is directly related to the thermal characteristics of the device design, the mounting technique used for the die, and the materials used for that interface.

The thermal conductivity of GaAs is related to the temperature of the material over a wide temperature range and varies approximately as $1/T$, where T is the temperature in kelvin. However, thermal conductivity can be considered linear over a very short temperature range [11].

The power handling capabilities , reliability, and performance of semiconductor devices are directly related to the junction temperature of the device during operation. While GaAs has a higher thermal resistivity than silicon, this is somewhat offset by the higher band gap of GaAs, allowing higher operating temperatures. Nevertheless, thermal considerations are extremely important in device design, packaging, and application.

References

[1] H. Welker, Z. Naturforsch, Vol. A, No. 7, p. 744, 1952.

[2] S. M. Sze, *Semiconductor Devices Physics and Technology,* J. Wiley, New York, 1985.

[3] R. F. Pierret, *Semiconductor Fundamentals*, Addison-Wesley Publishing, New York, 1989.

[4] R. Williams, *Modern GaAs Processing Methods*, Artech House, Dedham, MA, 1990.

[5] H. C. Casey, "Diffusion in III-V Compound Semiconductors," *Atomic Diffusion in Semiconductors*, D. Shaw, Editor, Plenum, New York, 1973.

[6] S. Y. Chiang, "Properties of Vacancy Defects in GaAs Single Crystals," *J. Appl. Phys*., Vol. 45, p. 2986, 1975.

[7] S. Markram-Ebied, "Nature of EL2: The Main Native Midgap Electron Trap in VPE and Bulk GaAs," in *Semi-insulating III-V Materials*, D. Look, Editor, Shiva Publishing Ltd., England, 1984.

[8] Y. Nanishi, *Jpn. J. Appl. Phys*., Vol. 22, 1983.

[9] S. Miyazawa, *Appl. Phys. Lett.*, Vol. 44, 1984.

[10] T. Honda, *Jpn. J. Appl. Phys*., Vol. 22, 1983.

[11] S. Wemple, "Thermal Design of Power GaAs," in *GaAs FET Principles and Technology*, J. DiLorenzo, Editor, Artech House, Dedham, MA., 1982.

Additional Reading

Blakemore, J., "Semiconducting and Other Major Properties of Gallium Arsenide," *J. Appl. Phys.* Vol. 53, No. 10, October 1982.

II. Metal–Semiconductor Junctions

G. E. Ponchak

The earliest solid-state device was reported in 1874. It consisted of a wire tip pressed into a lead-sulfide crystal. This simple metal–semiconductor junction was the first solid-state device and became known as a whisker contact rectifier. Although whisker contact rectifiers are rarely used anymore, the metal–semiconductor junction is the most important solid-state component in microwave integrated circuits. A few examples of circuit elements that include metal–semiconductor junctions are Schottky diodes, varactor diodes, metal–semiconductor field-effect transistors (MESFETs), high-electron-mobility transistors (HEMTs), and heterojunction bipolar transistors (HBTs).

Using modern semiconductor fabrication processes, the metal–semiconductor junction is very easy to create. Metal is selectively deposited onto an n-GaAs region and an alloying bake is performed if it is required. In other words, fabrication of this junction requires only one mask level and possibly a bake. Besides its ease of fabrication, the junction is very versatile. By varying the type of metal or the semiconductor doping level, the junction can be made into a rectifying or a nonrectifying junction. Rectifying junctions preferentially permit current to flow in one direction versus the other. For example, electrons may flow easier from the metal into the semiconductor than the opposite. Therefore, a rectifying junction acts as a gate keeper to stop current from flowing in the reverse direction. The rectifying junction is commonly called a Schottky contact or a Schottky barrier junction. The nonrectifying junction or ohmic contact permits current to flow across the junction in both directions with very low resistance.

Metal–semiconductor junctions represent the essential and basic building blocks of GaAs-based devices. Therefore, it is essential to get an understanding of the metal–semiconductor junction structure and operation, and the reliability issues related to them. It will become clear throughout this text that a large volume of data has been collected on the reliability issues and failure mechanisms related to metal–semiconductor junctions. This section will introduce the reader to the metal–semiconductor junction and its characteristics, and it will present an introduction to the related failure mechanisms and reliability concerns. Chapter 4 will provide a more detailed discussion of metal–semiconductor-related failure mechanisms.

A. Junction Physics

Figure 3-6 shows a schematic of a metal–semiconductor junction formed on an n-type GaAs substrate with an external bias supply connected to the metal. Although the schematic is simple, it is also an accurate representation of the junction. To understand the junction dynamics, it is necessary to examine the energy-band diagram of the junction. It helps to first study the energy band diagram for a metal and an n-type semiconductor separated from each other such that neither material is influenced by the other. Figure 3-7(a) shows such a case. As discussed in Section 3-I, a finite number of electrons exist in the conduction band of the semiconductor, and the number of these free electrons is dependent on the temperature and doping concentration or purity of the material. Likewise, there are a number of free electrons in the metal, and the number of free electrons is dependent on the metal and the temperature. The only new parameter introduced in Figure 3-7(a) is the metal work function, ϕ_m. The work function is the energy required to remove an electron from the Fermi level of the metal to a vacuum potential. Most of the metals commonly used in GaAs circuits and devices have work functions between 4 and 5.5 eV.

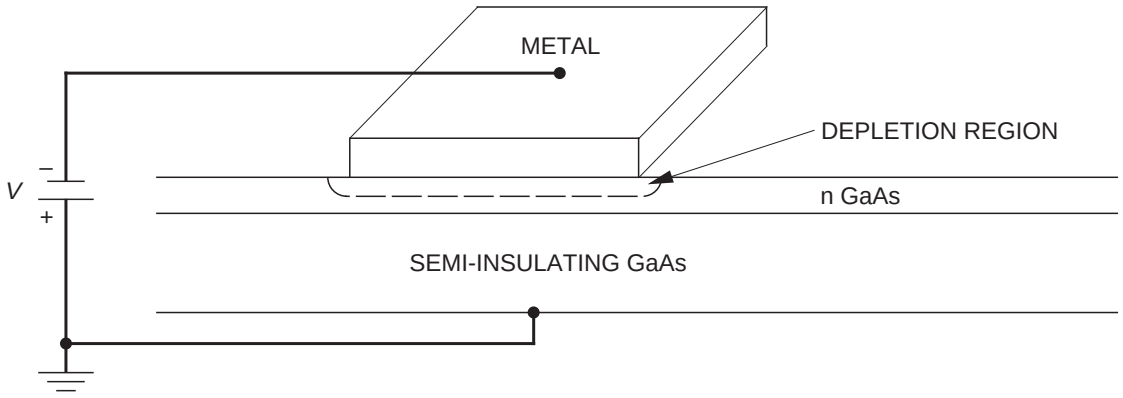

Figure 3-6. Schematic and cross section of metal–GaAs junction.

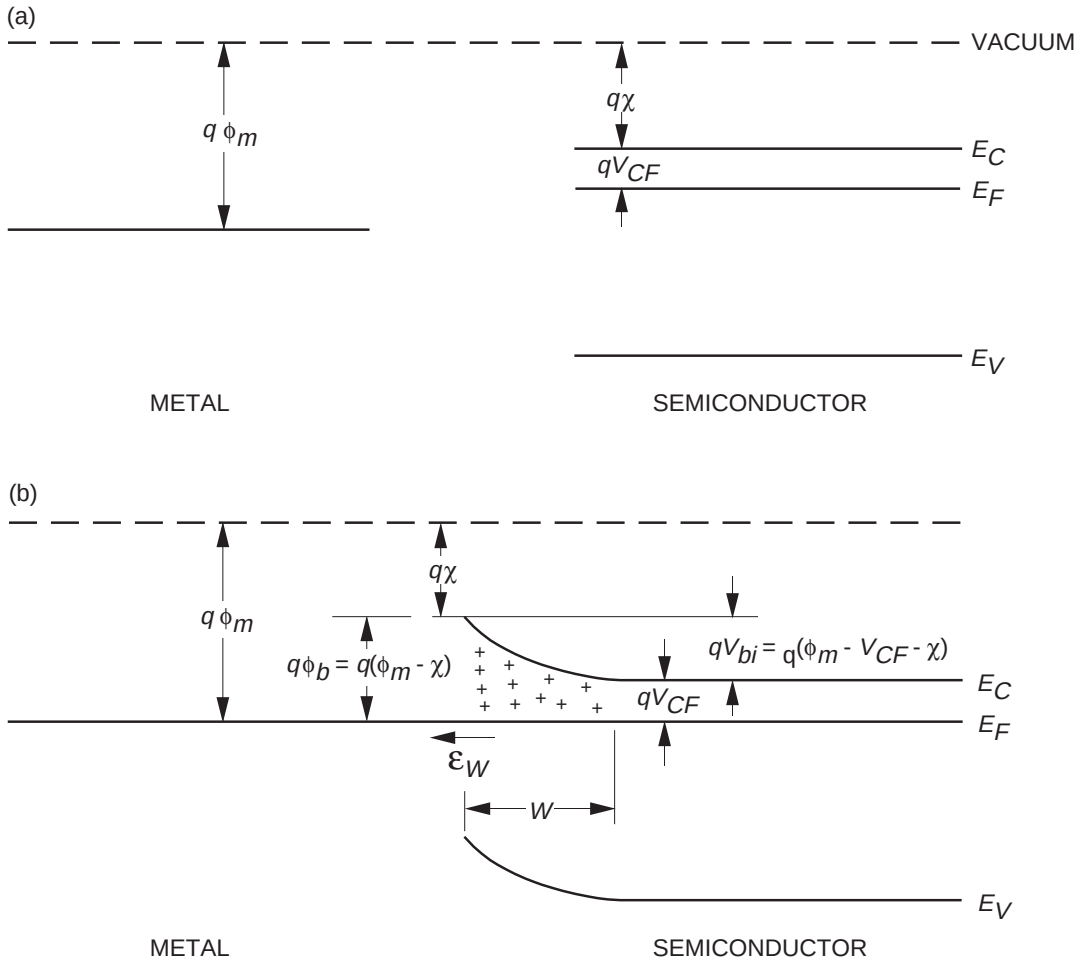

Figure 3-7. Energy band diagram of metal and semiconductor (a) separate from each other and (b) in intimate contact.

If the semiconductor Fermi level is greater than the metal Fermi level, $\chi + V_{CF} < \phi_m$, as is shown in Figure 3-7(a), then when the metal and semiconductor are put in intimate contact, electrons will diffuse from the semiconductor to the metal. As electrons

are depleted from the semiconductor, a net positive charge is created in the semiconductor at the junction. This positive charge will exert a force on the electrons that opposes the diffusion current. Equilibrium is established when these two forces are equal. Figure 3-7(b) shows the contact in equilibrium. Notice that the semiconductor energy bands bend in response to the forces just described. It is within this region, called the depletion region, that all of the junction's electrical properties are established. The amount of band bending is called the built-in potential, V_{bi} . For an electron to cross from the semiconductor to the metal, it must overcome V_{bi}, whereas an electron moving from the metal to the semiconductor must overcome the barrier potential, ϕ_b. To a first approximation, the barrier height is independent of the semiconductor properties, whereas V_{bi} is dependent on the doping level.

If an external potential is applied across the junction, the added electric field will disturb the equilibrium conditions. Consider first a positive external potential (see Figure 3-6). This will create an electric field across the junction that is opposite to the electric field caused by the depleted GaAs atoms. The result is that the diffusion current will not be sufficiently opposed, and current will flow across the junction. This is shown schematically in Figure 3-8(a). Note the reduction in the barrier for electrons flowing

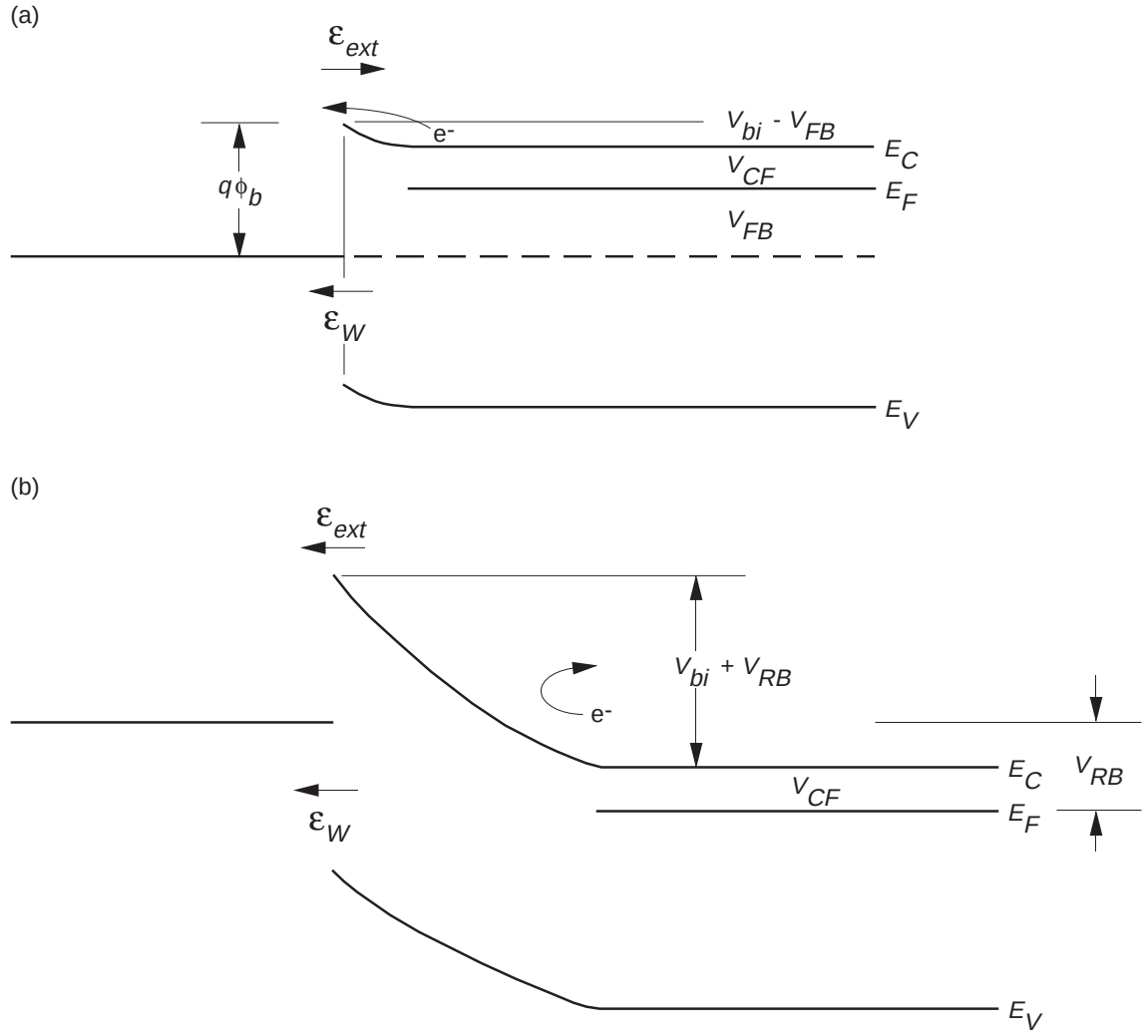

Figure 3-8. Energy band diagram of metal–semiconductor junction under (a) forward bias and (b) reverse bias.

from the semiconductor to the metal, but not for electrons flowing from the metal to the semiconductor. If a negative voltage is applied to the metal, the external field will reinforce the electric field caused by the depleted carriers, increase the band bending at the junction, and prevent the diffusion current from flowing. (See Figure 3-8(b).)

The preceding description assumed ideal material conditions. Specifically, it was assumed that the semiconductor lattice structure was uniform and perfect, even at the surface of the material. In practical cases, this is not possible. The atoms on the exposed surface do not have the required neighboring atom to complete all of the covalent bonds. Therefore, these surface atoms may either give up an electron and become a positively charged donor ion, or accept an electron and become a negatively charged acceptor ion. Surface states and their associated charge cause the energy bands of the semiconductor to bend even before the metal is introduced, as shown in Figure 3-9. Furthermore, when the metal is brought into contact with the semiconductor, the surface states may be able to accommodate all of the charge movement required to equalize the free electrons between the two materials. When this occurs, the barrier potential is no longer dependent on the metal work function. Also, no additional band bending of the semiconductor occurs because of the metal–semiconductor contact. In other words, the junction characteristics are not dependent on the metal interface. Surface states can create severe reliability problems for GaAs devices since they are generally planar devices that use only the upper few thousand angstroms of the substrate. Therefore, besides altering the built-in voltage of the contact, surface states may also provide leakage paths for current [1].

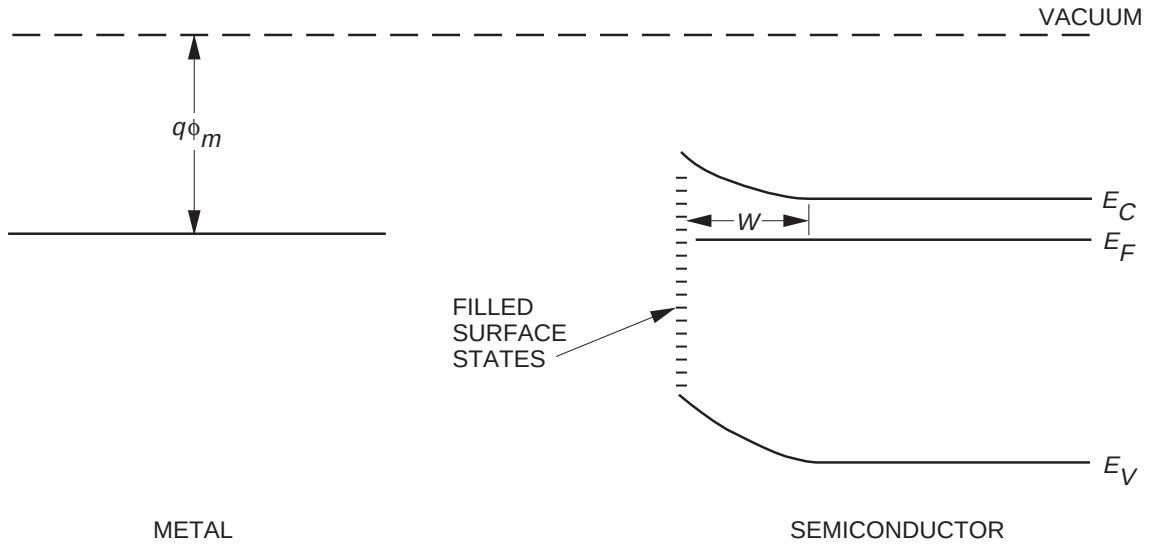

Figure 3-9. Energy band diagram of metal and semiconductor separate from each other when semiconductor surface states exist.

B. Junction Characteristics

Now that the critical parameters have been introduced, their dependence on the semiconductor and metal properties can be examined. First, consider the depletion width. Under abrupt barrier approximations, which are valid for junctions between metals and semiconductors, the width is given by

$$W = \sqrt{\frac{2\varepsilon_r \varepsilon_0}{qN_d}\left(V_{bi} - V - \frac{kT}{q}\right)} \tag{3-2}$$

where N_d is the donor doping concentration, k is Boltzmann's constant, and q is the charge of an electron. The term kT/q, often referred to as V_T, is approximately 0.026 V at room temperature whereas V_{bi} is approximately 1 V. From Equation (3-2), it is seen that the depletion width is smaller for highly doped semiconductors, and that the depletion width varies inversely with the applied bias. Based on the preceding discussion relating to Figure 3-8, it is noted that a positive bias increases current flow and decreases the depletion width. The opposite occurs for a negative bias.

Depletion widths can be quite large. As an example, consider two GaAs substrates at room temperature with an aluminum contact. Let the first have a typical MESFET channel doping of $N_d = 10^{17}/\text{cm}^2$ and the second have a typical ohmic contact doping of $N_d = 10^{19}/\text{cm}^2$. With no external bias supplied, the depletion widths for these two samples are approximately 0.048 μm and 0.006 μm, respectively. Although these appear to be very small quantities, it will become apparent throughout the rest of this chapter that these depletion widths are in fact large compared to the device dimensions required for microwave circuits.

Another critical parameter is the electric field across the depletion region. The concern is that the maximum electric field that occurs at the metal–semiconductor interface must be kept smaller than the breakdown field of GaAs, approximately 4×10^5 V/cm. If $E_m > 4 \times 10^5$ V/cm, electrons have enough kinetic energy to create electron/hole pairs during electron/atom collisions at a faster rate than the free charges can recombine. These new electrons also are accelerated by the electric field and create more electron/hole pairs. This runaway process is called “avalanche breakdown.” The result of avalanche breakdown is often a catastrophic junction failure. The maximum electric field is given by

$$E_m = \frac{qN_d}{\varepsilon_r \varepsilon_0} W = \sqrt{\frac{2qN_d}{\varepsilon_r \varepsilon_0}\left(V_{bi} - V - kT/q\right)} \tag{3-3}$$

The field is stronger when large doping concentrations are used or if a large reverse bias is applied across the junction.

The charge storage in the depletion region also creates a capacitance across the junction, which is given by

$$C = \frac{\varepsilon_r \varepsilon_0}{W} = \sqrt{\frac{q\varepsilon_r \varepsilon_0 N_d}{2\left(V_{bi} - V - \frac{kT}{q}\right)}} \tag{3-4}$$

Note two things about Equation (3-4). First, the capacitance is a function of the applied voltage. Therefore, the junction behaves as a voltage-controlled capacitance. It is this feature of the junction that is exploited in varactor diodes, which are commonly used in phase shifters and voltage controlled oscillators (VCOs). The second thing to note is that the capacitance is dependent on the doping concentration. Therefore, by varying the doping profile across the junction, the capacitance-voltage curve can be varied. Alternatively, if the doping concentration is altered during the life of the diode,

the capacitance will change, and a frequency shift in the VCO or a phase change from the phase shifter will occur.

Although an understanding of the depletion width and its associated capacitance are critical for the gate design of a field effect transistor, it is the current flow through the junction that the circuit designer is ultimately concerned with. In general, current flow through the junction is due to several mechanisms. It is necessary to examine only two of these for the purposes of this text. The first is the transport of electrons over the potential barrier, usually called thermionic emission. Thermionic emission current assumes that only electrons with energies greater than the energy of the potential barrier add to the current flow (see Figure 3-8). Several methods of analysis have been proposed to determine the current density, and although each uses different assumptions and boundary conditions, they all result in an equation of the form:

$$J = J_0 * e^{-q\phi_b/kT} * \left[e^{qV/kT} - 1\right] \tag{3-5}$$

J_0 increases with the doping concentration, N_d, and temperature. Note that J is exponentially dependent on the barrier potential, temperature, and the applied voltage. It is this strong dependence on the applied voltage that makes the junction a good rectifier. Furthermore, the dependence on temperature makes this current mechanism dominant at higher temperatures. When Schottky diodes are characterized, the measured current does not fit Equation (3-5) exactly but rather

$$J = J_0 * e^{-q\phi_b/kT} * \left[e^{qV/nkT} - 1\right] \tag{3-6}$$

where the parameter n is called the "ideality factor." An ideal diode would have $n = 1$, but for actual diodes, $n > 1$. A change in the ideality factor over the life of the diode is an indication that the metal–semiconductor interface is changing.

The second current mechanism that needs to be described is due to quantum mechanical tunneling through the potential barrier. Recall from quantum mechanics that the position of a particle is not absolute, but described by a distribution function. Therefore, although the majority of electrons will be confined by a potential barrier, there is a probability that some of the electrons will exist in the region of the potential barrier. Furthermore, if the potential barrier is thin enough, there is a probability that some electrons will travel through the barrier. This current component is referred to as tunneling current. Figure 3-10 shows the band diagrams for two cases where tunneling current is dominant. The first is a contact between a metal and a highly doped semiconductor. In this case, the depletion width, or the barrier width, is small; recall the example given earlier for the depletion width as a function of doping concentration. The second example shown in Figure 3-10(b) is an extremely reversed biased junction. The tunneling current may be given by

$$J \propto e^{\frac{-4k}{h}\sqrt{\varepsilon_r \varepsilon_0 m^*}\frac{\phi_b}{\sqrt{N_d}}} \tag{3-7}$$

which shows that the tunneling current will increase exponentially with the ratio

$$\sqrt{N_d} \,/\, \phi_b$$

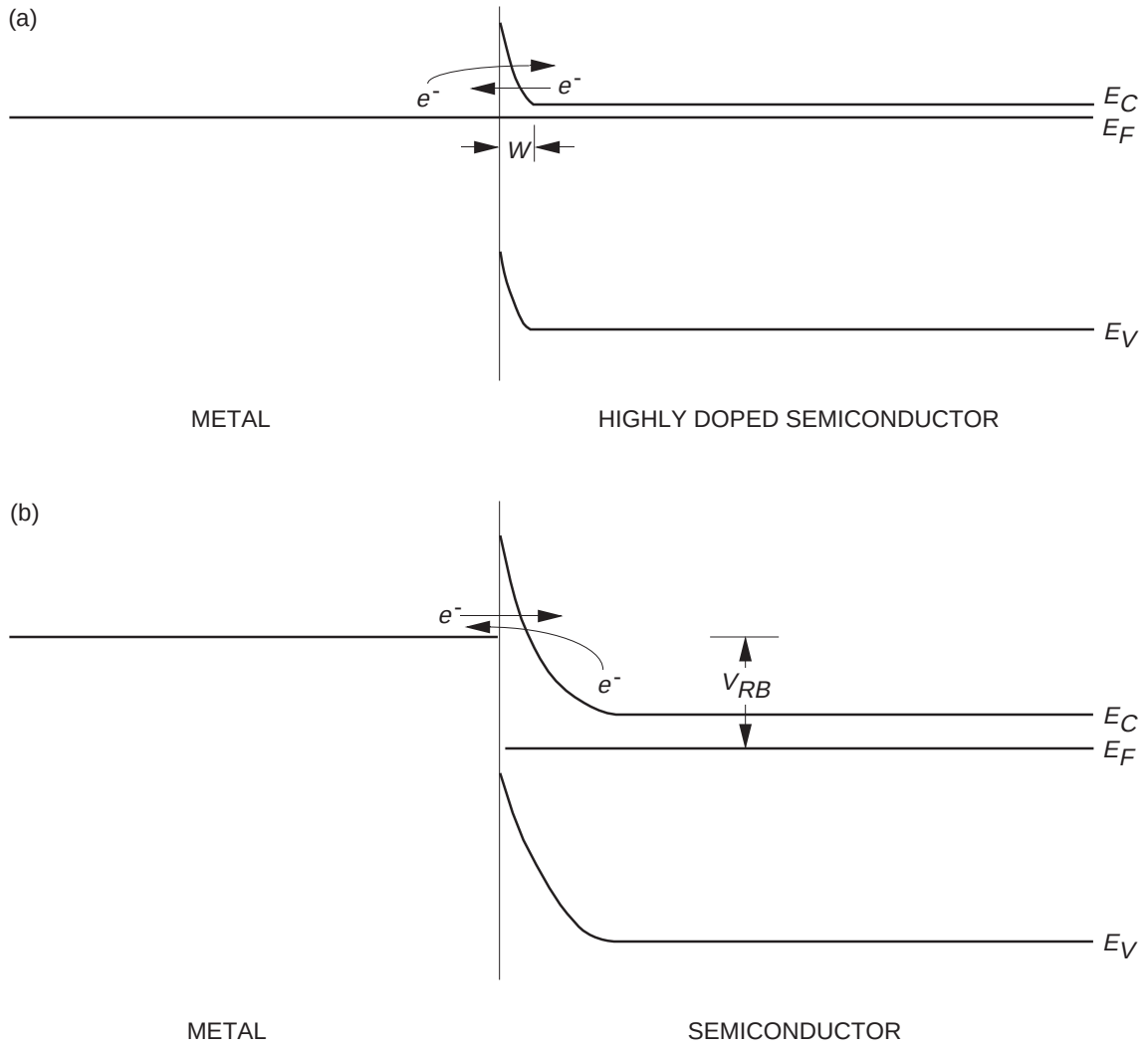

Figure 3-10. Energy band diagram of (a) metal-n^+–semiconductor junction and (b) metal–semiconductor junction under reverse bias.

For doping concentrations greater than 10^{17} cm^{-3} and for low temperatures, the tunneling current can be dominant. Since J is independent of V, this junction makes good ohmic contacts.

C. Device Structures

The practical implementation of a planar diode is shown in Figure 3-11(a). The diode is fabricated either on a molecular-beam-epitaxy- (MBE-) grown n layer or by ion implantation of an n region in the semi-insulating GaAs substrate. This is followed by the deposition of an ohmic contact metal, normally AuGe, and an ohmic contact alloying bake. Lastly, the Schottky contact metal is deposited. A simple equivalent circuit for the diode is shown in Figure 3-11(b). R_{ohm} refers to the ohmic contact junction resistance and R_{chan} refers to the resistance between the two metal contacts. Although both resistances are parasitic and ideally would be eliminated, practical limitations do not permit this.

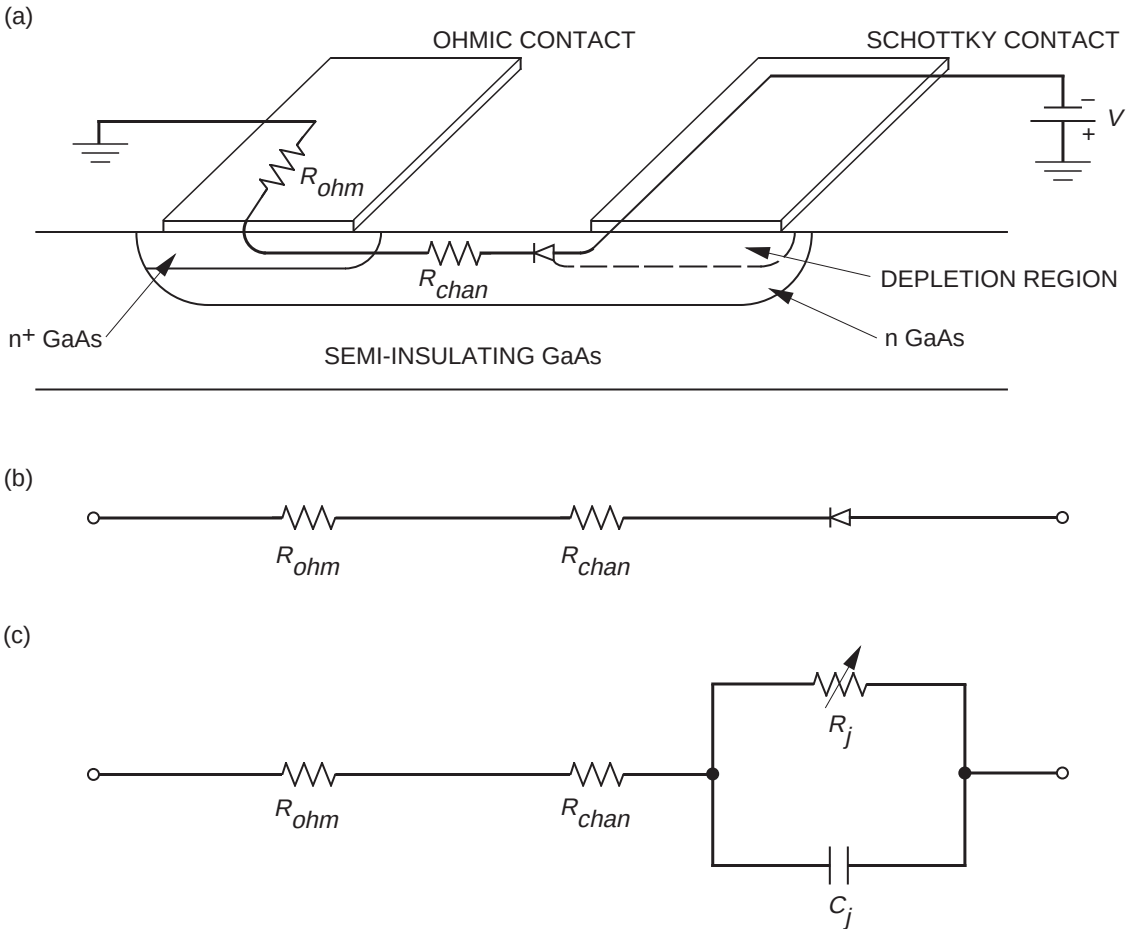

Figure 3-11. GaAs planar diode: (a) schematic, (b) simple equivalent circuit, and (c) equivalent circuit for a planar Schottky diode.

The diode electrical specifications will normally determine the doping concentration of the n region. Therefore, the ohmic contact resistance cannot be altered unless an n^+ region is formed upon which the ohmic contact can be made. R_{chan} can be reduced if the distance between the two contacts is reduced. Modern lithography permits the contacts to be separated by as little as 0.2 μm, although the contact separation is typically on the order of 1 μm. Unfortunately, the electric field between the two contacts increases as the spacing is reduced. If the electric field is increased too much because of the RF power or the dc bias, metal shorts may develop between the contacts leading to device failure. Therefore, limitations on the power handling capability of the diode are normally imposed.

Lastly, consider the diode itself. It has already been shown that the depletion region creates a capacitance called the junction capacitance, C_j. In addition, there is also a junction resistance, R_j, which is in parallel to C_j. R_j accounts for the current flow through the depletion region and can be derived from Equation (3-6) as

$$R_j = \frac{nkT}{qJA} \tag{3-8}$$

where A is the diode area. Therefore, the equivalent circuit for a planar Schottky diode is shown in Figure 3-11(c). Notice that R_j is shown as a variable resistance due to its dependence on J, which in turn is dependent on the applied voltage.

The important figure of merit for Schottky diodes is the forward current cutoff frequency:

$$f_c = \frac{1}{2\pi R_F C_F} \tag{3-9}$$

where R_F is the total series resistance and C_F is the junction capacitance at a slight forward bias. Schottky diodes have been fabricated with cutoff frequencies greater than 1 THz. In general, a diode can be used at frequencies less than $f_c/10$. Therefore, it is desirable to have a small R_j and C_j as well as a small R_{ohm} and R_{chan}.

To minimize R_j, the diode area must be increased, but to minimize C_j, the diode area must be decreased. Furthermore, a parasitic capacitance due to the fringing fields along the edges of the Schottky contact exists. This parasitic capacitance is proportional to the diode periphery and the number of corners on the contact. Since the ratio of the contact periphery to area increases as the area of the contact decreases, it is not practical to reduce C_j solely by decreasing the area.

D. Reliability

Reducing the diode area has been discussed as a method of reducing C_j. Besides the disadvantages of having a small diode area already discussed, there are other disadvantages. First, the fringing fields around the periphery of the diode will be greater and will lead to increased leakage current. Second, the fringing fields can be larger than the electric field predicted by Equation (3-3), especially around the corners of the contact. Therefore, the reverse breakdown voltage will be smaller. Third, the current through the diode is equal to $J*A$. Therefore, the current density must be increased as the area is decreased to maintain reasonable current through the device. If the current density is increased too much, failures due to electromigration may occur. Finally, the increased current density and the reduced junction area may cause the junction temperature to increase. Since GaAs is a relatively poor thermal conductor, thermal-related failure mechanisms may increase as well. To get around these problems, it is better to maximize the area and to minimize C_j by decreasing N_d. Note that reducing N_d requires n^+ regions for the ohmic contacts and increases the risk of ionic contamination failures.

The critical points to remember about the junction as it relates to device reliability are

(1) The sensitivity of its electrical characteristics to the semiconductor doping concentration.

(2) The interface barrier potential.

(3) The junction temperature.

Small changes in any of these parameters can greatly change the junction impedance and therefore the current that flows through the junction. While the circuit designer can control the junction temperature through proper packaging and heat sinking, unfortunately the barrier potential and doping concentration may change unpredictably over the life of the junction—especially at higher operating temperatures or if metal–

semiconductor interactions occur. These failure mechanisms will be fully described in Chapter 4.

Reference

[1] K. A. Christianson, "Contribution of Surface and Interface States to the Reverse Bias Aging of GaAs Schottky Barriers," 1*990 GaAs REL Workshop Digest*, Oct. 7, 1990.

Additional Reading

Goyal, R., *Monolithic Microwave Integrated Circuits: Technology & Design,* Artech House, Norwood, MA, 1989.

Sze, S. M., *Physics of Semiconductor Devices,* John Wiley & Sons, New York, 1981.

Warner, R. M., Jr., and B. L. Grung, *Transistors—Fundamentals for the Integrated-Circuit Engineer*, John Wiley & Sons, New York, 1983.

Yang, E. S., *Fundamentals of Semiconductor Devices,* McGraw-Hill Book Company, New York, 1978.

III. Metal–Semiconductor Field-Effect Transistors (MESFETs)

G. E. Ponchak

GaAs metal–semiconductor field-effect transistors (MESFETs) are the most commonly used and important active devices in microwave circuits. In fact, until the late 1980s, almost all microwave integrated circuits used GaAs MESFETs. Although more complicated devices with better performance for some applications have been introduced, the MESFET is still the dominant active device for power amplifiers and switching circuits in the microwave spectrum.

The basic MESFET is shown schematically in Figure 3-12. The base material on which the transistor is fabricated is a semi-insulating GaAs substrate. A buffer layer is epitaxially grown over the semi-insulating substrate to isolate defects in the substrate from the transistor. The channel or the conducting layer is a thin, lightly doped (n) conducting layer of semiconducting material epitaxially grown over the buffer layer. Since the electron mobility is approximately 20 times greater than the hole mobility for GaAs, the conducting channel is always n type for microwave transistors. Finally, a highly doped (n^+) layer is grown on the surface to aid in the fabrication of low-resistance ohmic contacts to the transistor. This layer is etched away in the channel region. Alternatively, ion implantation may be used to create the n channel and the highly doped ohmic contact regions directly in the semi-insulating substrate. Two ohmic contacts, the source and drain, are fabricated on the highly doped layer to provide access to the external circuit. Between the two ohmic contacts, a rectifying or Schottky contact is fabricated. Typically, the ohmic contacts are Au–Ge based and the Schottky contact is Ti–Pt–Au.

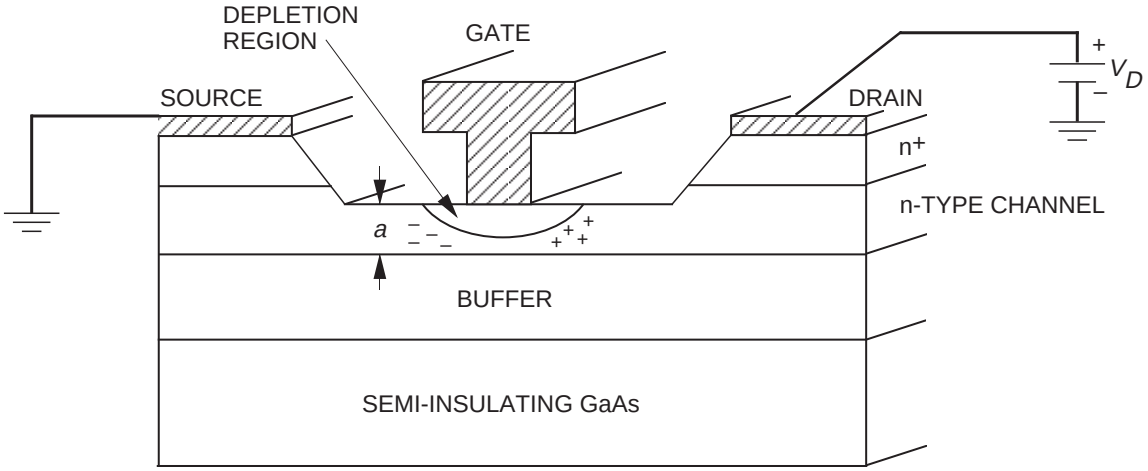

Figure 3-12. Schematic and cross section of a MESFET.

A. Device Physics

The basic operation of the MESFET is easily understood by first considering the I–V characteristics of the device without the gate contact, as shown in Figure 3-13. If a small voltage is applied between the source and drain, a current will flow between the two contacts. As the voltage is increased, the current increases linearly with an associated resistance that is the sum of the two ohmic resistances, R_S and R_D, and the channel resistance, R_{DS}.

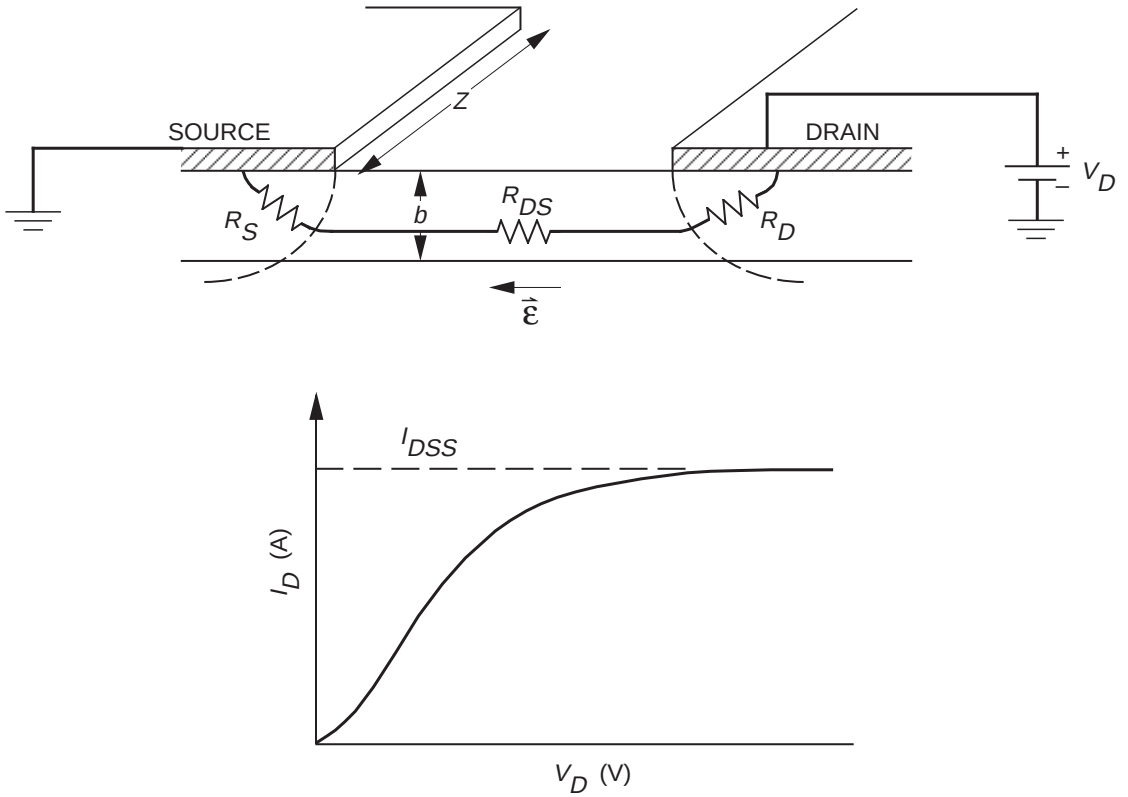

Figure 3-13. Schematic and I–V characteristics for an ungated MESFET.

$$I_D = \frac{V_D}{R_D + R_S + R_{DS}} \tag{3-10}$$

If the voltage is increased further, the applied electric field will become greater than the electric field required for saturation of electron velocity, as shown in Figure 3-4. Under large bias conditions, an alternative expression for I_D is useful; this expression relates the current directly to the channel parameters:

$$I_D = Q(x)v(x) = Zb(x)qn(x)v(x) \tag{3-11}$$

This expression omits the parasitic resistances, R_S and R_D. The parameters in Equation (3-11) are Z, the width of the channel; $b(x)$, the effective channel depth; q, the electron charge; $n(x)$, the electron density; and $v(x)$, the electron velocity, which is related to the electric field across the channel. Note that if $v(x)$ saturates, I_D will also saturate. This saturation current is called I_{DSS}.

Now consider the effect of the gate electrode placed over the channel but without any gate bias, $V_G = 0$. As presented in Section 3-II, a depletion region formed under the gate electrode reduces the effective channel depth, $b(x)$, and therefore increases the resistance to current flow under the gate. The depletion region depth is dependent on the voltage drop across the Schottky junction. Since the current flowing through the channel is equivalent to a current flow through a distributed resistor, there is a larger voltage drop across the drain end of the channel than at the source end. This results in the depletion region depth being greater on the drain side of the channel. The nonuniform channel

depth has two effects on the device operation. First, there is an accumulation of electrons on the source side and a depletion of electrons on the drain side of the depletion region. This dipole of charge creates a feedback capacitance between the drain and the channel; this capacitance is typically called C_{DC}. The second effect is that the electric field due to the dipole adds to the applied electric field causing the saturation conditions to occur at a lower V_D. By applying a bias to the gate junction, the depletion depth and therefore the resistance of the current flow between the source and drain and the saturation current can be controlled. If a large enough negative gate bias is applied, the depletion region depth will equal the channel depth, or the channel will be pinched off. This gate bias is called the pinch-off voltage and is given by

$$V_P = \left(\frac{qN_d}{2\varepsilon_0 \varepsilon_r} \right) a^2 \tag{3-12}$$

Under pinch-off conditions, the drain current drops to a very small value. Therefore, the transistor can act as a voltage-controlled resistor or a switch.

The most important feature of MESFETs is that they may be used to increase the power level of a microwave signal, or that they provide gain. Because the drain current can be made to vary greatly by introducing small variations in the gate potential, the MESFET can be modeled as a voltage-controlled current source. The transconductance of the MESFET is defined as

$$g_m = \left. \frac{-\partial I_D}{\partial V_{GS}} \right|_{-V_{DS}-constant} \tag{3-13}$$

Using short-channel approximations, it can be shown that the transconductance may be written as

$$g_m = \frac{I_S}{2V_P} \left(\frac{I_S}{I_S - I_D} \right) \tag{3-14}$$

where I_S is the maximum current that can flow if the channel were fully undepleted under saturated velocity conditions. This is the same as the saturation current discussed for the device without the gate electrode shown in Figure 3-13. Since I_S is proportional to the channel depth, a, and V_P is proportional to the square of the channel depth, g_m is inversely proportional to the channel depth. In addition, from Equation (3-10), we note that for large I_S and g_m, the parasitic resistances R_S and R_D must be minimized.

The most commonly used figures of merit for microwave transistors are the gain bandwidth product, the maximum frequency of oscillation, f_{max}, and the frequency where the unilateral power gain of the device is equal to one, f_t. Consider first the parameter f_t. If short gate length approximations are made, f_t can be related to the transit time of the electrons through the channel, t, by the expression

$$f_t = \frac{1}{2\pi\tau} = \frac{v_{sat}}{2\pi L} \tag{3-15}$$

Since v_{sat} is approximately 6×10^{10} μm/s for GaAs with doping levels typically used in the channel, the gate length must be less than 1 μm for f_t to be greater than

10 GHz. The parameter f_{max} may be approximated by

$$f_{\max} = \frac{f_t}{2}\sqrt{\frac{R_{DS}}{R_G}} \tag{3-16}$$

where R_G is the gate resistance. From the above two expressions for f_t and $f_{max,}$ it is apparent that the gate length should be made as small as possible. Both the limits of fabrication and the need to keep the electric field under the channel less than the critical field strength required for avalanche breakdown set the lower limit on L at approximately 0.1 μm. For the gate to have effective control of the channel current, the gate length L must be larger than the channel depth, a, or $L/a > 1$. This requires a channel depth on the order of 0.05 to 0.3 μm for most GaAs MESFETs. The small channel depth requires that the carrier concentration in the channel be as high as possible to maintain a high current.

B. Reliability

The small feature sizes described above may create reliability problems in microwave GaAs MESFETs. The small cross section of the gate electrode results in a high current density, especially for power transistors, which leads to electromigration failures. To reduce the gate resistance, gold is typically used over the gate refractory metal. Since gold creates deep-level traps in GaAs, which effectively reduce the carrier concentration and therefore the current of the device, barrier metals such as platinum must be used. In addition, because the channel depth is so small, any diffusion of gate metals into the GaAs creates large changes in the current that flows through the channel and decreases the pinch-off voltage. The small distance between the gate and drain electrodes also creates high electric fields, which may create an avalanche generation of electrons. These "hot electrons" may then become trapped in the surface states of the GaAs or in the passivation material that is commonly deposited over the device. The reliability problems that occur greatly depend on the device technology, as well as its application. In small-signal applications, the degradation of the ohmic contacts or interdiffusion of the gate metals with the GaAs in the channel leads to shifts in I_D, g_m, and V_P.

Although power MESFETs also suffer from parametric degradation, catastrophic failures are more common. However, advances in device technology and operation within safe limits have decreased the incidence of burnout. For power amplifiers, the MESFET must be designed for maximum power output. This is equivalent to requiring a large drain-to-source voltage and a large drain current. Unfortunately, both of these parameters may not be maximized simultaneously. To maximize I_D, a large carrier concentration or a large gate width is required; note that the channel depth may not be increased since that would degrade the frequency range of the device. The carrier concentration may not be increased without degrading the gate-to-drain breakdown voltage, which must be maximized to maximize V_{DS}. Therefore, the only alternative is to increase the gate width, Z. Unfortunately, in microwave circuit design, long line lengths do not appear as lumped elements with a uniform potential along the length, but rather as distributed transmission lines with potential nulls occurring every half wavelength. The general rule of thumb is that a line should be less than one tenth of a wavelength long to be considered a lumped element. For GaAs, this is equivalent to

$$Z \le \frac{11.3}{f}\ \text{mm} \tag{3-17}$$

where *f* is the frequency in GHz. Therefore, at X-band, 8 to 12 GHz, the maximum gate width that may be used is approximately 1 mm. If greater current is required, multiple gate fingers may be used in a parallel connection. This parallel connection of gate fingers in a tightly packed region increases the localized temperature of the circuit. Since GaAs is a poor thermal conductor, power transistors will typically operate at least 10 deg above the carrier temperature. This increased device temperature, with the higher fields and currents used in power MESFETs, often leads to catastrophic failures.

Additional Reading

Bahl, I., and P. Bhartia, *Microwave Solid State Circuit Design*, John Wiley & Sons, New York, 1988.

Sze, S. M., *Physics of Semiconductor Devices*, John Wiley & Sons, New York, 1981.

IV. HEMTs and PHEMTs

L. Aucoin

GaAs-based high-electron mobility transistors (HEMTs) and pseudomorphic HEMT (or PHEMTs) are rapidly replacing conventional MESFET technology in military and commercial applications requiring low noise figures and high gain, particularly at millimeter-wave frequencies. The application of PHEMTs for high-efficiency power amplification is gaining popularity. Other commonly used names for HEMTs include MODFET (modulation doped FET), TEGFET (two-dimensional electron gas FET) and SDHT (selectively doped heterojunction transistor).

Since HEMTs and PHEMTs are field-effect transistors, the basic principles of their operation are very similar to those of the MESFET described in Section 3-III. The main difference between HEMTs and MESFETs is the epitaxial layer structure. In the HEMT structure, compositionally different layers are grown in order to optimize and to extend the performance of the FET. For III–V semiconductors using a GaAs substrate, the common materials used are $Al_xGa_{1-x}As$ and GaAs. For most device applications, the Al_xAs mole fraction is between $0.2< x < 0.3$. The PHEMT also incorporates $In_xGa_{1-x}As$, where In_xAs is constrained to $x < 0.3$ for GaAs-based devices. These different layers form heterojunctions since each layer has a different band gap. Structures grown with the same lattice constant but different band gaps are simply referred to as lattice-matched HEMTs. Those structures grown with slightly different lattice constants are called pseudomorphic HEMTs or PHEMTs. Figure 3-14 shows the band-gap energy as a function of lattice constant for the III–V semiconductors. In this section, a simple epitaxial layer structure for HEMTs will be described with reference to MESFETs where appropriate. Some salient features of HEMT and PHEMT device physics will be highlighted. Lastly, the effects of the epitaxial structure and the device operation on reliability will be discussed.

A. Device Physics

The epitaxial structure of a basic HEMT is illustrated in Figure 3-15. Similar to the MESFET, the HEMT structure is grown on a semi-insulating GaAs substrate using molecular beam epitaxy (MBE), or less common, metal–organic chemical vapor deposition (MOCVD). Table 3-2 contains the common MESFET, HEMT, and PHEMT epitaxial structures.

The buffer layer, also typically GaAs, is epitaxially grown on the substrate in order to isolate defects from the substrate and to create a smooth surface upon which to grow the active layers of the transistor. Many PHEMT structures contain a superlattice structure to further inhibit substrate conduction. A superlattice structure is a periodic arrangement of undoped epitaxial layers used to realize a thicker epitaxial layer of a given property. For example, alternating layers of $Al_xGa_{1-x}As$ and GaAs form a typical PHEMT superlattice. The $Al_xGa_{1-x}As$ has a larger band gap than GaAs, making it superior to GaAs as a buffer. However, due to strain problems, the $Al_xGa_{1-x}As$ layer thickness is limited. To resolve this problem, the $Al_xGa_{1-x}As$ is grown to just below its thickness limit and a thin layer of GaAs is grown on top. The GaAs relieves the strain and allows another layer of $Al_xGa_{1-x}As$ to be grown. This process is typically repeated 10 to 15 times, creating a layer that is "essentially" a thick buffer of $Al_xGa_{1-x}As$.

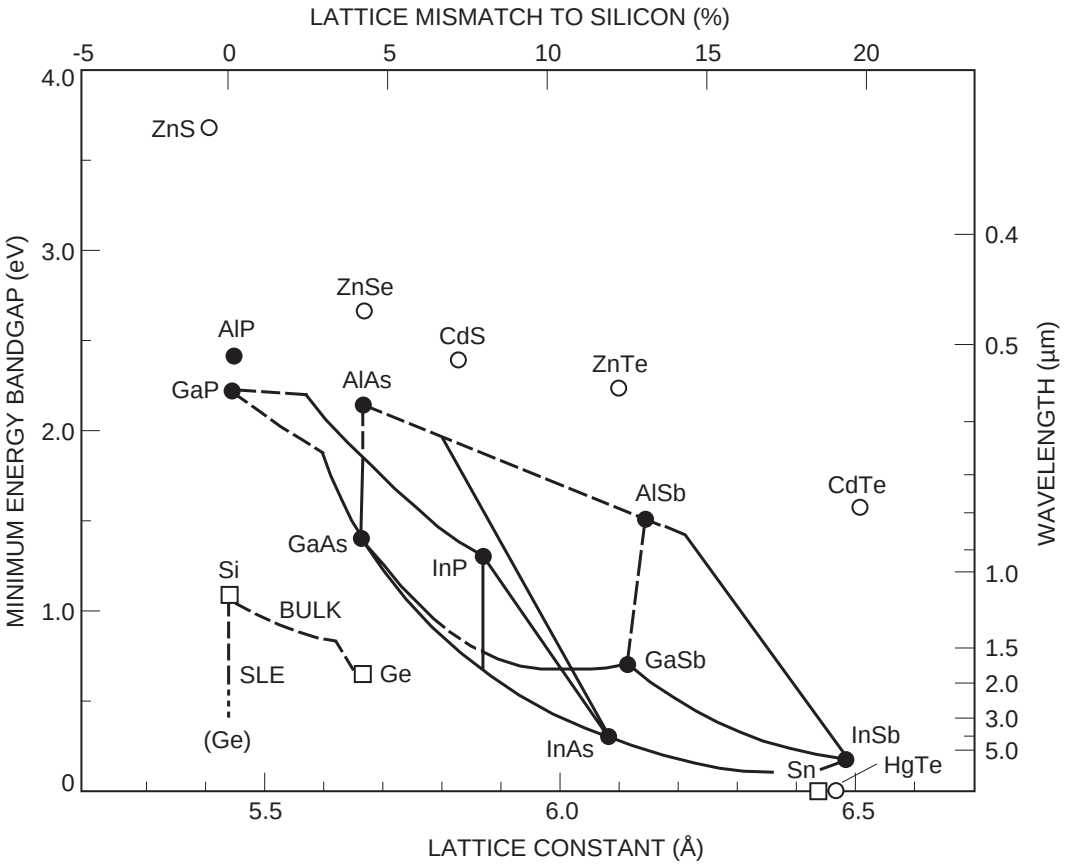

Figure 3-14. Minimum band-gap energy vs lattice constant data for III–V semiconductors. The right axis indicates the wavelengths of light that would be emitted by a laser or LED for a material of the corresponding bandgap. Connecting lines give information for alloys of the materials at the endpoints of a given line segment. Solid lines indicate a direct bandgap and dashed lines an indirect bandgap. For Ge–Si, the line denoted BULK corresponds to unstrained, lattice-mismatched growth and the line SLE to strained layer epitaxy of Ge–Si on unstrained Si. (From [1]; reprinted by permission of John Wiley & Sons, Ltd.)

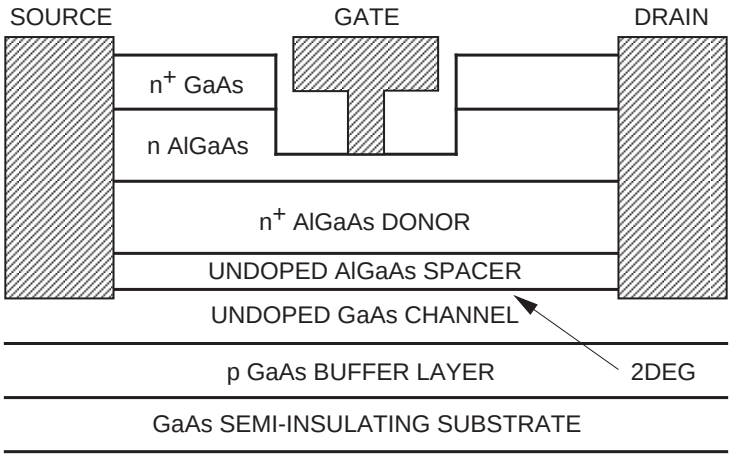

Figure 3-15. Epitaxial structure of a basic AlGaAs/GaAs HEMT.

Table 3-2. Epitaxial layer compositions for basic GaAs-based HEMT and PHEMT devices compared with those of MESFET.

Device Layer	MESFET	HEMT	PHEMT
Ohmic contact	n^+ GaAs	n^+ GaAs	n^+ GaAs
Schottky contact	n GaAs	n AlGaAs	n AlGaAs
Donor		n^+ AlGaAs or Si pulse doping	n^+ AlGaAs or Si pulse doping
Spacer		Undoped AlGaAs	Undoped AlGaAs
Channel	n^+ GaAs	Undoped GaAs	Undoped InGaAs
Buffer	p^- GaAs	p^- GaAs	p^- GaAs

In the conventional HEMT structure, the channel is grown next. In the ideal system, all of the electron conduction would take place in this channel. The most important point about the channel layer in the HEMT and PHEMT devices is the two-dimensional electron gas (2DEG) that results from the band-gap difference between $Al_xGa_{1-x}As$ and GaAs (or $Al_xGa_{1-x}As$ and $In_xGa_{1-x}As$, in the case of the PHEMT). Illustrated in Figure 3-16 is the band diagram of a generic HEMT showing the 2DEG formed by the different band gaps. The 2DEG is formed since the higher band gap of $Al_xGa_{1-x}As$ allows free electrons to diffuse from the $Al_xGa_{1-x}As$ to the lower band gap GaAs (or $In_xGa_{1-x}As$) near the interface. A potential barrier then confines the electrons to a thin sheet of charge known as the 2DEG. In contrast to the MESFET, which has a doped channel and consequently lots of ionized donors, the 2DEG has significantly less Coulomb scattering, resulting in a very high mobility device structure.

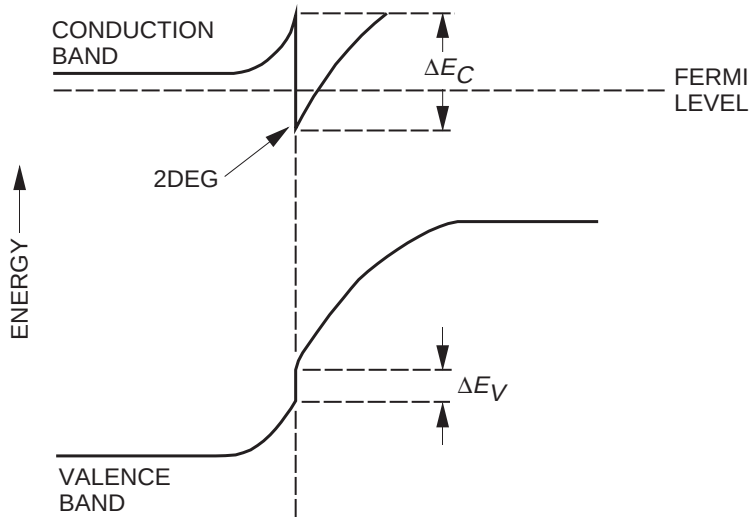

Figure 3-16. Energy band diagram of a generic AlGaAs–GaAs HEMT showing the 2DEG quantum well channel.

The remainder of the HEMT structure contains an $Al_xGa_{1-x}As$ spacer layer, a donor layer n^+ $Al_xGa_{1-x}As$, an n $Al_xGa_{1-x}As$ Schottky contact layer, and a highly doped n^+ GaAs layer. The spacer layer serves to separate the 2DEG from any ionized donors generated by the pulse doping or n^+ active layer. The drawback of the spacer layer,

however, is that the sheet carrier concentration (total amount of charge) in the channel is reduced as the spacer layer thickness is increased. The donor layer or Schottky layer is an n^+ $Al_xGa_{1-x}As$ layer and serves as the source of electrons. To avoid the possibility of electron conduction in the $Al_xGa_{1-x}As$ (which has a low electron mobility), the thickness of the Schottky must be chosen so that the depletion region of the gate overlaps the depletion at the $Al_xGa_{1-x}As$/2DEG interface for depletion mode devices. The n^+ GaAs is present to realize low-resistance ohmic contacts.

The structure described above is a basic HEMT structure. Most of the structures used today are variants of this, having been optimized for performance and applications. For instance, many PHEMTs used for power applications will incorporate two silicon pulses, the second one below the channel, to increase the total charge available.

The fabrication and basic operation of HEMT and PHEMT devices are very similar to those for the MESFET. There exist some differences, mainly related to the presence of the $Al_xGa_{1-x}As$ in the epitaxial structure. As mentioned previously, $Al_xGa_{1-x}As$ has a larger band-gap energy than GaAs and the band gap increases with the AlAs mole fraction. HEMTs require ohmic contacts directly to the 2DEG, which is made more difficult with increased AlAs mole fraction. An advantage of the AlGaAs is the higher Schottky barrier height resulting from the deposition of the gate metal on the AlGaAs. Unfortunately, the high doping in the donor layer decreases the breakdown voltage. However, power HEMT and PHEMT structures with higher breakdown voltages (>10 V) have been engineered using either double recess technology or by reducing the doping in the Schottky layer.

Under operation, HEMTs and MESFETs are biased similarly. When a negative gate bias is applied to the HEMT device, the Schottky layer becomes depleted. As the gate is biased further, the 2DEG becomes depleted. This results in the modulation of the channel (2DEG) by a negatively applied gate bias where gain and amplification occur until the channel is pinched off (i.e., fully depleted). The transconductance is given by

$$g_m = \left(\varepsilon v_{sat} W_g\right) / d$$

where ε = permittivity of $In_xGa_{1-x}As$, v_{sat} = saturated velocity of $In_xGa_{1-x}As$, W_g = unit gate width of the device, and d = distance from the gate to the 2DEG. Under high-electric-field conditions, the HEMT shows a higher saturated velocity over the MESFET . Since the conduction of electrons from the source to the drain takes place in a channel that is well confined, g_m will remain very high at low drain currents. This is somewhat contrasted with the MESFET because at low drain current, the distance d will increase because the edge of the depletion region enters the tail of the doping profile. This results in a compression of the g_m. The higher mobility of the HEMT results in lower parasitic drain and source resistances. As a result, $f_t = g_m/(2\pi C_{gs})$ and f_{max} are increased from the MESFET case for a given gate length leading to a lower noise figure and higher gain.

B. Reliability

The reliability of HEMTs and PHEMTs is affected by the epitaxial structure, device fabrication, and device geometry. One of the drawbacks of using $Al_xGa_{1-x}As$ in the material structure is the occurrence of traps, called DX centers, for Al_xAs mole fractions around $x = 0.26$. These traps are deep donor levels that can lead to reduced drain current, an increase in low frequency noise and photoconductivity, and are particularly problematic at low temperature. Further, the creation of DX centers increases with higher doping of the $Al_xGa_{1-x}As$. DX centers are avoided by keeping the

Al_xAs content below $x = 0.24$ for n-type doped $Al_xGa_{1-x}As$. A second possible reliability problem that can occur is the deconfinement of the 2DEG under high-temperature conditions. Thermally accelerated testing has shown that the Al can migrate laterally into the gate, resulting in a change in the conduction band discontinuity.

To take advantage of the high g_m, HEMTs and PHEMTs rely on small geometries for optimum performance. Paralleling the MESFET, HEMTs and PHEMTs suffer the same electromigration and metal interdiffusion reliability problems associated with the ohmic and gate metallizations under device operation. In addition, hot electron traps resulting from the generation of avalanche electrons are a problem for HEMTs and PHEMTs. The hot electrons cause a degradation in the current and in the gain and power under microwave drive as they become trapped in the passivation or the $Al_xGa_{1-x}As$ passivation interface. For power devices, catastrophic failures or "burnout" can also be an issue due to the high channel temperatures resulting from the large currents required for high power.

Reference

[1] J. C. Bean, "Materials and Technologies," in *High-Speed Semiconductor Devices*, S. M. Sze, Editor, John Wiley & Sons, New York, 1990.

Additional Reading

GaAs Integrated Circuits, J. Mun, Editor, MacMillan Publishing Company, New York, 1988.

HEMTs and HBTs: Devices, Fabrication and Circuits, F. Ali, and A. Gupta, Editors, Artech House, Boston, 1991.

Reliability of GaAs MMICs, A. Christou, Editor, John Wiley and Sons, West Sussex, England, 1992.

V. Heterojunction Bipolar Transistors

Y.C. Chou and R. Ferro

AlGaAs/GaAs heterojunction bipolar transistors (HBTs) are used for digital and analog microwave applications with frequencies as high as Ku band. HBTs can provide faster switching speeds than silicon bipolar transistors mainly because of reduced base resistance and collector-to-substrate capacitance. HBT processing requires less demanding lithography than GaAs FETs, therefore, HBTs can cost less to fabricate and can provide improved lithographic yield. This technology can also provide higher breakdown voltages and easier broad-band impedance matching than GaAs FETs.

In comparison with Si bipolar junction transistors (BJTs), HBTs show better performance in terms of emitter injection efficiency, base resistance, base-emitter capacitance, and cutoff frequency. They also offer good linearity, low phase noise and high power-added efficiency. Table 3-3 shows a comparison of typical device characteristics between AlGaAs/GaAs HBTs and Si BJTs. HBTs are used in both commercial and high-reliability applications, such as power amplifiers in mobile telephones and laser drivers.

Table 3-3. Comparison of AlGaAs/GaAs HBT and Si bipolar transistors.

Parameter	AlGaAs/GaAs HBT	Si BJT
Forward transit time, τ_F	4 ps	12 ps
Early voltage, V_a	800 V	25 V
Collector-substrate capacitance, C_{cs}	~0	~15 fF
Base resistance, R_b	70 W	200 W

For NPN BJTs, a useful figure of merit that is important in determining the current gain is the ratio,

$$\frac{\textit{electron current injected from the emitter into the base}}{\textit{hole current injected from the base into the emitter}}$$

This ratio is called the injection efficiency, and it is usually optimized in BJTs by highly doping the emitter and lightly doping the base. High injection efficiency is obtained in an HBT by using a material with a larger energy band gap for the emitter than that used for the base material. The large energy band-gap emitter blocks injection of holes from the base. Therefore, the doping concentration in the base and emitter can be adjusted over a wide range with little effect on injection efficiency. In the normal operation of a bipolar transistor, the collector-base junction is reverse biased or at least not forward biased enough to cause appreciable injection current. The collector and base material are the same in most HBTs, although some use wide band-gap collector materials to improve the collector base breakdown voltage.

It follows that AlGaAs/GaAs HBTs benefit from the following advantages:

(1) Lower forward transit time along with a much lower base resistance (due to the much higher base doping concentration), giving increased cutoff frequency f_t.

(2) Better intrinsic device linearity due to a higher beta (gain) early-voltage product.

(3) Very low collector-substrate capacitance C_{cs} in AlGaAs/GaAs HBTs due to the use of semi-insulating GaAs substrate (resistivity $\approx 10^7$ Ohm-cm).

(4) High efficiency due to the ability to turn off devices completely with a small base voltage change and the extremely small turn-on voltage variation between devices.

(5) Good wide-band impedance matching due to the resistive nature of the input and output impedances.

(6) Low cost and potential for high throughput. With the typical minimum feature size of 1 μm, there is no need for e-beam lithography.

A. Device Structure

The cross section of an example HBT is shown schematically in Figure 3-17. The material on which an HBT is fabricated is grown on a semi-insulating GaAs substrate. These epitaxial layers could be grown by molecular beam epitaxy (MBE) or the metal–organic chemical vapor deposition (MOCVD) method. A heavily doped n^+ GaAs layer with a concentration on the order of 4×10^{18} cm^{-3} is grown first for the collector contact, followed by a lightly doped n GaAs layer for the collector. The collector doping concentration is on the order of 3×10^{16} cm^{-3}. A heavily doped p^+ GaAs layer with a concentration greater than 5×10^{18} cm^{-3} is used for the base. In general, beryllium (Be) or carbon (C) is used for the base dopant. Again, a wide-band-gap AlGaAs layer is grown for the emitter. Finally, a heavily doped n^+ GaAs layer is grown to facilitate the fabrication of low-resistance ohmic contacts. An n^+ InGaAs-alloy contact layer can be also grown to provide stable, low-resistance emitter contacts. Some HBTs utilize compositional grading on both sides of the emitter-base junction to maximize electron injection efficiency and suppress hole injection from base into emitter. AuBe/Pd/Au is a typical base contact metallization and AuGe/Ni/Au is often used for the emitter and collector contacts. Finally, a multiple boron damage implant is used for device isolation.

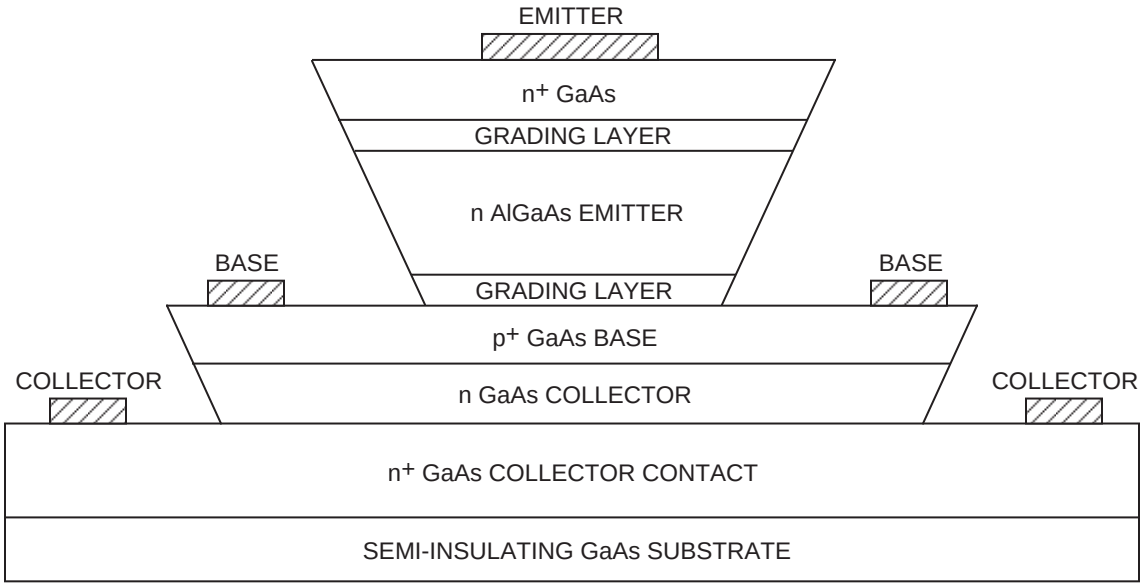

Figure 3-17. Cross section of an example HBT.

A combination of selective (emitter mesa) and nonselective wet etching processes is performed to access the HBT's individual contact. Since base resistance strongly affects HBT microwave performance, it is desirable to place the base contact as close to the emitter as possible. Although advanced lithography is able to define base contacts within 0.1 μm of the emitter, self-aligned (SA) fabrication techniques are more practical because of their low fabrication costs. Although mesa etching is the most common isolation technique for HBTs, ion implantation has been used successfully. A reliability study [1] indicated that isolation by multiple energy implants of fluorine and hydrogen produces stable isolation layers.

Recombination on the exposed extrinsic base surface is a major mechanism of current gain degradation in small geometry AlGaAs/GaAs HBTs. As shown in Figure 3-18, the use of a thin, depleted, AlGaAs ledge surrounding the emitter mesa has demonstrated improved current gain and reliability of HBT devices.

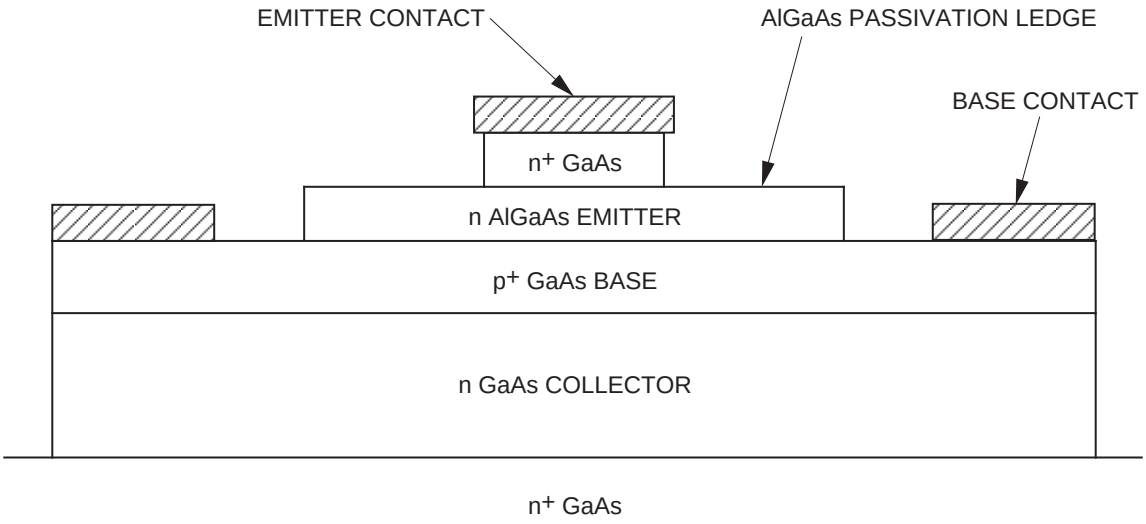

Figure 3-18. An HBT cross section showing a thin ledge of AlGaAs.

Both abrupt and compositionally graded E–B junctions can be used for AlGaAs/GaAs HBTs. Figure 3-19 depicts the energy band diagrams of these two kinds of HBTs. The AlGaAs emitter has a wider band gap than the GaAs base layer. The abrupt E–B junction has a potential spike and notch that can be smoothed out by linear compositional grading on both sides of the AlGaAs emitter over a distance of about 300 Å, thus reducing the barrier that electrons have to overcome.

B. Operating Principles

The potential barriers for hole injection (ΔV_p) and electron injection (ΔV_n) in a graded E–B junction differ by the band-gap difference (ΔE_g) between the AlGaAs emitter and the GaAs base. Therefore, we have

$$q\left(\Delta V_p - \Delta V_n\right) = \Delta E_g$$
$$\left(\Delta E_g = E_g(AlGaAs) - E_g(GaAs)\right) \tag{3-18}$$

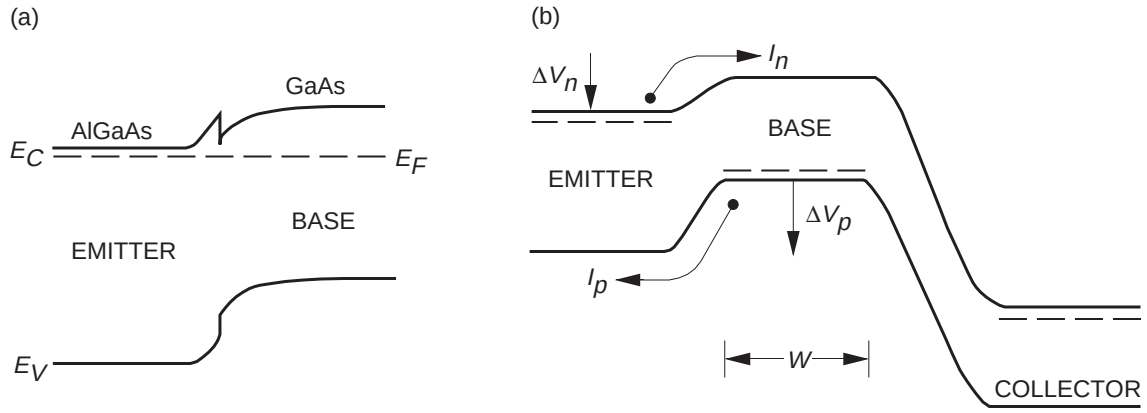

Figure 3-19. AlGaAs/GaAs HBTs: (a) abrupt E–B junction and (b) graded E–B junction.

This small band-gap difference ΔE_g affects the ratio of I_n/I_p significantly where I_n is the electron injection current from the emitter into the base and I_p is the undesired hole injection current from the base into the emitter.

I_n and I_p can be expressed by using the Boltzmann approximation,

$$I_n = qAN_E\left(D_n / W\right)e^{\left(-q\Delta Vn/kT\right)} \tag{3-19}$$

$$I_p = qAN_B\left(D_p / L_p\right)e^{\left(-q\Delta Vp/kT\right)} \tag{3-20}$$

The parameters in Equations (3-19) and (3-20) are q, the electronic charge; k, Boltzmann constant; T, temperature; A, the emitter-base junction area; D_n, the electron diffusivity in the base; W, the base width; N_E, the emitter doping concentration; D_p, the hole diffusivity in the emitter; and L_p, the hole diffusion length in the emitter. Obviously,

$$I_n / I_p = \left(D_n / D_p\right)\left(L_p / W\right)\left(N_E / N_B\right)e^{\Delta Eg/kT} \tag{3-21}$$

For $Al._3Ga._7As$/GaAs HBTs, $\Delta E_g \approx 14.6\ kT$ and $\exp\left(\Delta E_g/kT\right) \approx 2 \times 10^6$. Thus, the ΔE_g difference provides a significant improvement in I_n/I_p over the bipolar transistors case ($\Delta E_g = 0$).

This property of HBT devices allows the fabrication of a heavily doped base and a lightly doped emitter without affecting current gain too much. In practice, the base current is dominated by recombination so that common emitter current gain is typically below 100. The low emitter doping concentration decreases the E–B junction capacitance, which affects the current gain cutoff frequency and maximum frequency of oscillation. High doping in the base reduces the base sheet resistivity and base contact resistance, giving rise to the improvement of maximum frequency of oscillation.

C. Reliability

AlGaAs/GaAs HBT technology has reached a certain degree of maturity with insertion into microwave, analog, digital and low-medium-power applications. However,

additional reliability data would be needed to warrant widespread acceptance of this class of devices in high-reliability applications requiring high power and high-current-density operation.

HBTs, as other semiconductor devices, have a number of potential failure mechanisms. HBTs typically suffer from current-induced degradation at high-current-density operation. Emitter-base and collector-base leakage currents often originate at the surface of the emitter-base and collector-base junctions, respectively [2]. Thermal and recombination-aided diffusion of crystalline defects from the bulk semiconductor to the heterointerface in abrupt junction HBTs has been suggested to account for increases in base current during burn-in of AlGaAs/GaAs HBTs. Additional evidence of the role of the emitter-base heterojunction was obtained in a study that found that implant-isolated HBTs degrade more than mesa-isolated HBTs [3]. The degradation consisted of a shift in V_{be} and a decrease in h_{FE} , apparently resulting from the contact of the emitter-base junction edge to the defect-laden implant region. Passivation of the emitter junction is important in mesa-isolated HBTs. The use of depleted AlGaAs has given good results.

Beryllium-doped AlGaAs/GaAs HBTs have shown acceptable reliability in accelerated life tests and small-signal applications [4]. However, a more accelerated degradation has been observed on Be-doped HBTs as compared with C-doped HBTs under high-current-density conditions. The suspected cause of device degradation in Be-doped HBTs at high current densities is the field-aided diffusion of positively charged interstitial Be atoms from the base into the AlGaAs emitter, giving rise to V_{be} shift, current-gain decrease, and base and collector current change.

However, as with other semiconductor devices, there are a number of mechanisms by which heterojunction bipolar transistors can fail. The degradation mechanisms that have been reported in heterojunction bipolar transistors include the following:

(1) Decrease in current gain and increase in base-emitter voltage at high emitter currents.

(2) Increases in contact resistance caused by degradation of the interface between the emitter ohmic contact metallization and the emitter semiconductor. An InGaAs contact layer is helpful in solving this problem [4].

(3) Gettering of crystalline defects at the emitter-base heterojunction.

(4) Decrease in current gain and increase in base-emitter voltage for a specified collector current caused by oxidation of the emitter mesa surface in the region of the emitter base heterojunction.

Specialized epitaxy growth for Be-doped-base HBTs, strain-relaxed base layer for C-coped base HBTs [5], the use of an InGaAs emitter cap layer, and emitter ledge passivation are some techniques used to alleviate some observed degradation mechanisms.

HBTs for power applications are designed with a multifinger implementation. In a multifinger layout, the current and temperature distributions on each finger are different, leading to degradation of device power performance. One of the most undesirable phenomena is called “collector-current collapse,” which results in an abrupt decrease of collector current in the devices’ dc I–V characteristics. Figure 3-20 shows typical I–V characteristics in a power HBT with a multifinger design under collector current collapse. The collector-current collapse occurs when a particular finger (usually

center) suddenly draws most of the collector current because of its nonuniform current distribution, leading to a decrease of device current gain. Although collector-current collapse has not been observed to cause catastrophic failures on power HBTs, the output power and performance of the device are generally limited. Optimized HBT layout improves power performance and minimizes collector current collapse.

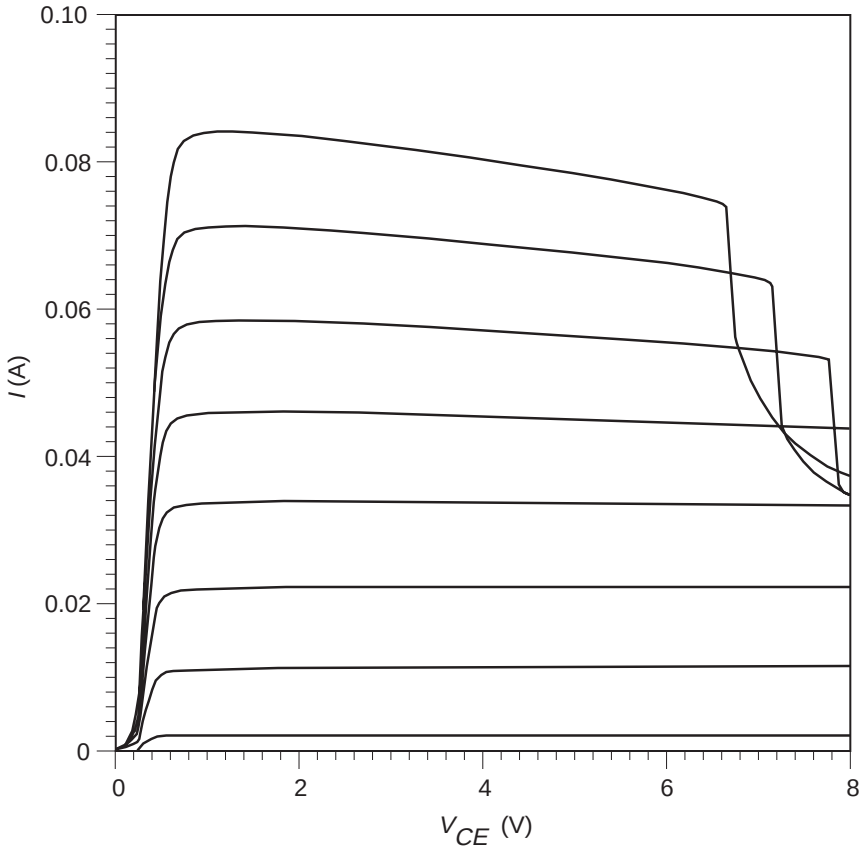

Figure 3-20. Typical I–V characteristic of a power HBT with multifinger design under collector current collapse.

References

[1] F. Ren, "Fabrication Techniques for Self-Aligned GaAs-Based HBTs and Submicron Gate Length FETs," *International Journal of Modern Physics B*, Vol. 8, No. 16, pp. 2221-43, July 20, 1994.

[2] J. Liou, K. Parab, C. Huang, B. Bayraktaroglu, D. Williamson, "Base and Collector Leakage Currents and Their Relevance to the Reliability of AlGaAs/GaAs Heterojunction Bipolar Transistors," *Proc. of 32nd Annual International Reliability Physics Symposium*, San Jose, CA, April 11–14, 1994.

[3] T. Nozu, K. Tsuda, M. Asaka, M. Obra, "Reliability Characteristics of Mesa-Etched Isolated Emitter Structure AlGaAs/GaAs HBTs with Be-Doped Base," *14th Annual GaAs IC Symposium*, Miami Beach, FL, pp. 157–60, October 4–7, 1992.

[4] F. M. Yamada, A. K. Oki, D. C. Streit, D. K. Umemoto, L. T. Tran, K. Najita, J. R. Kessler, G. Lopez, Y. Saito, and E. A. Rezck, "Reliability Performance of

Components and ICs From a Production GaAs HBT Process," 1995 GaAs Reliability Workshop, sponsored by JEDED JC-14.7 Committee on GaAs, October 29, 1995, San Diego, CA.

[5] H. Sugahara, J. Nagano, T. Nittono, K. Ogawa, "Improved Reliability of AlGaAs/GaAs Heterojunction Bipolar Transistors With a Strain-Related Base," IEEE GaAs IC Symposium Digest, San Jose, CA, pp. 115–118, October 1993.

VI. PIN Diodes

G. E. Ponchak

Although PN junctions are the workhorse of Si circuit designs, GaAs PN junctions did not develop as a viable device. The primary reason for this is the much lower hole mobility compared to the electron mobility for GaAs, whereas in Si the difference in mobility is not as great. The resulting low hole drift velocity limits the maximum frequency of GaAs p-type devices, and since GaAs is primarily used for high-frequency and high-speed circuits, GaAs PN junctions were not developed. Unfortunately, without the need or desire for GaAs PN junctions, p-type GaAs ion implantation and MBE growth were not developed and integrated into GaAs circuit production facilities. As a result, GaAs PIN (p-type–insulator–n-type) diodes were not available to MMIC designers. The unavailability of MMIC PIN diodes was unfortunate because they have fast switching speeds, high breakdown voltage, and a variable resistance with bias. These positive characteristics may be used in the design of high-power switches [1], variable attenuators [2], photo detectors, and variable-gain amplifiers [3].

The unavailability of GaAs p-type regions changed with the recent advent of the GaAs HBT MMIC. With the good performance of these HBTs, p-type ion implantation and MBE growth are now being incorporated into GaAs production facilities. By using the p^+ base layer, the n^- collector region, and the n^+ collector ohmic contact layer of the HBT as shown in Figure 3-21, MMICs with PIN diodes can now be made easily on an GaAs HBT fabrication line. Since the emitter of the HBT is not used, the diode is strictly a GaAs device.

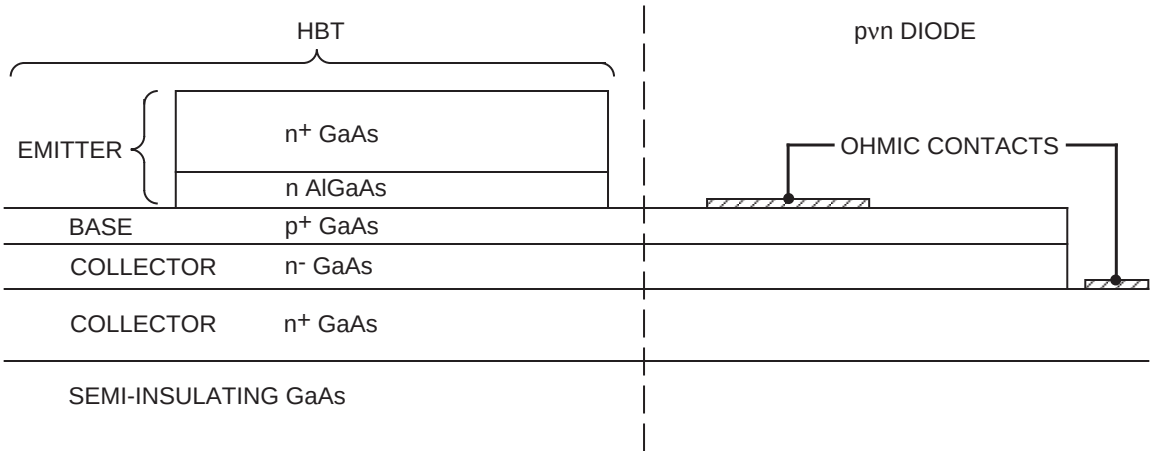

Figure 3-21. pvn fabrication from HBT structure.

A. Device Physics

Ideally, a PIN diode would have a perfect insulator between the p-type and the n-type regions. Although semi-insulating GaAs would be a good insulator, as already shown in Figure 3-21, the p and n regions of practical diodes are separated by a lightly doped n or n^- region. This n^- region is referred to as a *v* region and the resulting diode as a pvn diode. If a p^- region were used instead, the diode would be referred to as a pπn diode. Throughout this section, the pvn diode will be presented. The extension to the pπn diode is straightforward.

Figure 3-22(a) shows a pνn diode schematically with an applied voltage source connected. At the $p^+\nu$ interface, a PN junction is formed. Similar to the metal–semiconductor junction described in Section 3-II, a depletion region will be formed at the

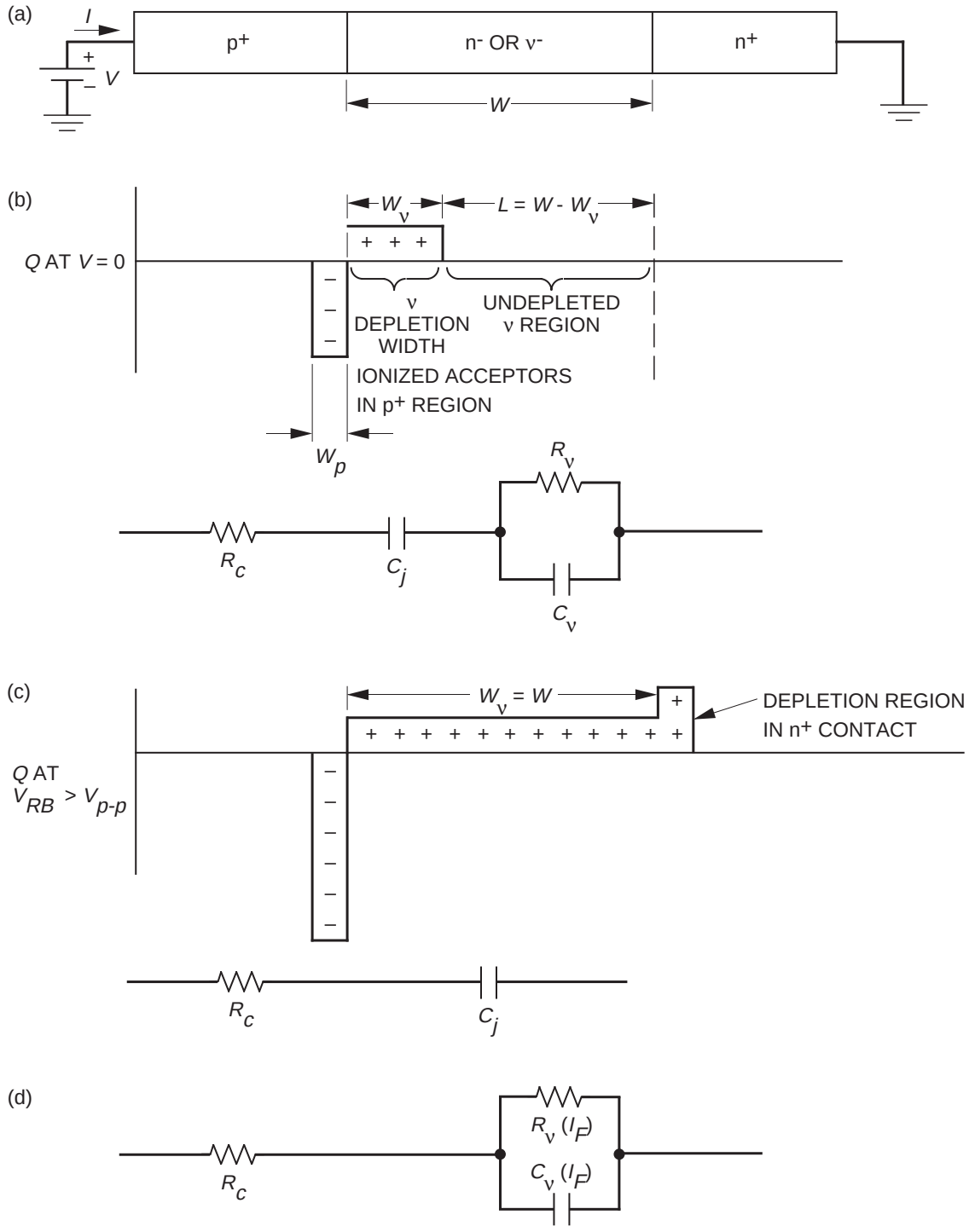

Figure 3-22. pνn diode: (a) schematic, (b) depletion region, (c) punch-through modeling for switching applications, and (d) forward-biased diode.

junction, but unlike the metal–semiconductor junction, depletion regions are formed on both sides of the junction. This is shown in Figure 3-22(b). Because the total charge in the p^+ depletion region must equal the total charge in the *v* depletion region or $W_p N_A = W_v N_D$, the depletion width in the *v* region will be greater. Usually, $W_v >> W_p$ and the total depletion region may be approximated as W_v. In this sense, the metal–semiconductor model may be used where the p^+ and the n^+ regions can be considered as metal electrodes and the depletion width and junction capacitance are determined by Equations (3-2) and (3-4) in Section 3-II. The equivalent circuit for the unbiased diode is also shown in Figure 3-22(b). R_C is the total resistance of the n^+ and p^+ ohmic contacts, $R_v = \rho L/A$ is the resistance in the *v* region, $C_v = \varepsilon A/L$ is commonly called the diffusion capacitance and accounts for charge storage in the undepleted *v* region, $C_j = \varepsilon A/W_v$ is the depletion region capacitance, *A* is the area of the diode, W_v is the depletion width, *W* is the *v* region width, and ρ is the resistivity of the *v* region. In general, R_C is constant, but C_j, C_v, and R_v are bias dependent.

If a reverse bias is applied to the junction, $V < 0$, the depletion width will increase. If a large enough reverse bias is applied, the depletion width will extend across the *v* region. The potential at which $W_v = W$ is called the punch-through potential and is commonly specified as V_{PT}. Once the punch-through potential has been reached, R_v reduces to a very small value, C_v increases to a large value, and C_j is approximately $\varepsilon A/W$. Since the parallel combination of R_v and C_v can now be approximated by a short circuit, the reversed biased diode may be approximated by R_C and C_j, both of which are constant. The reverse bias potential required for punch through is given by:

$$V_{PT} = \frac{qN_d W^2}{2\varepsilon_r \varepsilon_0} \tag{3-22}$$

It is seen that if a perfect insulator were used in the PIN diode, $V_{PT} = 0$. Although this is rarely the case, V_{PT} can still be small. For example, if a p*v*n diode were made from the HBT structure shown in Figure 3-17 of Section 3-V, V_{PT} would be approximately 5 V. The punch-through potential is important since the diode has its greatest reverse-biased impedance when $V_{RB} > V_{PT}$. Therefore, for switching applications, the diode is always driven into punch through and can be modeled as shown in Figure 3-22(c). By designing the diode so that C_j is small, the impedance of the reverse-biased diode will be large. This is required for switching applications since a large reverse-biased impedance results in good isolation. Therefore, a small *A/W* ratio is desired for switch applications.

If the diode is forward biased, the depletion width decreases and the junction capacitance increases to a large value. For most applications, the impedance resulting from C_j is small and the element can be ignored. The forward bias also causes electrons to be injected into the *v* region from the n^+ contact and holes to be injected into the *v* region from the p^+ contact. This increase in carriers causes the resistivity of the *v* region to decrease. Since the amount of charge injection is dependent on the bias potential, the forward-biased impedance is bias dependent. Using Equation (3-6) of Section 3-II, R_v may be written as [2]

$$R_v = \frac{nkT}{qI_F} \tag{3-23}$$

where I_F is the forward current and *n* is the ideality factor. Similarly, C_v may be written as [2]

$$C_v = \frac{\tau}{R_v} \tag{3-24}$$

where τ is the carrier transit time in the *v* region. If one writes the impedance for the capacitor, C_v, as

$$Z_{Cv} = \frac{R_v}{j\omega\tau} \tag{3-25}$$

it is seen that the ratio of the impedance associated with C_v to R_v can vary from values less than one to greater than one, depending on the product of ω and τ. Therefore, in general, the forward-biased diode is modeled as shown in Figure 3-22(d). For switching applications, the diode will be driven hard to decrease the forward impedance. Under strong forward drive, GaAs p*v*n diodes fabricated from HBT structures have impedances from 3 to 5 ohms.

The preceding development of the equivalent circuit models is valid for most applications. For some applications, though, a parasitic capacitance, C_p, which shunts the entire device, must be added to account for fringing fields from the two device contacts. The value of C_p depends on the diode design. For example, if the diode is designed in a circular shape, the fringing fields will be small. For diode structures designed on HBT production lines, it has been reported that $C_p > 2\ C_j$ [2]. Therefore, the diodes reverse-biased impedance may be significantly lower than predicted, and switching performance may be severely degraded.

B. Reliability

The reliability of PIN diodes fabricated in HBT production lines has not been well addressed. It is anticipated that the diode will share some of the same reliability concerns as the HBT with the exception of the emitter contact and the base-emitter junction failure mechanisms. The main reliability concern with PIN diodes is that they are typically used in high-power applications such as switching circuits. Therefore, they are often subjected to large electric fields as well as elevated temperatures.

Reliability concerns for the PIN diode arise when large forward or reverse biases or RF signals are applied. Consider first the forward-biased diode. The diode has a low impedance that permits a large current to flow across the device and an associated large I^2R loss. Since GaAs is a poor thermal conductor, the power loss causes the temperature to increase, which in turn may create thermal related failures such as metal–semiconductor diffusion. Since the p^+ region of the diode is approximately 0.1 μm thick, ohmic contact diffusion through the p^+ region is a primary concern. In addition, diffusion of ions from the p^+ and n^+ contacts into the n region will be accelerated at increased temperatures causing changes in the diode's electrical characteristics.

When the diode is reverse biased, the critical parameter that must be controlled is the electric field across the device. The electric field is given by

$$E = \frac{V_{RB}}{W} \tag{3-26}$$

If $E > 4 \times 10^5$ V/cm and W is sufficiently large to permit carrier-atom collisions, avalanche breakdown results with the possibility of catastrophic device failure. For GaAs, the maximum reverse bias that may be applied is given by

$$V_{RB} = 40W \tag{3-27}$$

where W is in microns. This limitation of V_{RB} sets the lower limit on W. Recall that a large W gives a smaller C_j and therefore better switch isolation. Therefore, it may appear that W should be made large. The disadvantage of this is that the switching speed of the diode is related to the time it takes to sweep all of the injected carriers out of the forward-biased n region, and this is directly related to the width W. The tradeoffs between fast switching speed, large power handling capabilities, and large switch isolation normally would be made by the PIN-diode designer. When the diodes are integrated with or fabricated with HBT circuits, the HBT performance must also be included in the tradeoff analysis.

References

[1] D. Teeter, et. al., "Ka-Band GaAs HBT PIN Diode Switches and Phase Shifters," *IEEE 1994 Microwave and Millimeter-Wave Monolithic Circuits Symposium Digest,* pp. 183–186.

[2] K. W. Kobayashi, et. al., "Monolithic GaAs HBT p-i-n Diode Variable Gain Amplifiers, Attenuators, and Switches," *IEEE Trans. on Microwave Theory and Tech.*, pp. 2295–2302, Dec. 1993.

[3] K. W. Kobayashi, et. al., "A Novel Baseband-1.5 GHz Monolithic HBT Variable Gain Amplifier with Pin Diode Gain Control," *IEEE 1995 Microwave and Millimeter-Wave Monolithic Circuits Symposium Digest,* pp. 181–184.

[4] J. F. White PhD, *Microwave Semiconductor Engineering,* Van Nostrand Reinhold Company, New York, 1982.

[5] H. A. Watson, *Microwave Semiconductor Devices and Their Circuit Applications,* McGraw-Hill Book Co., New York, 1969.

VII. Passive Elements

S. Kayali and G. E. Ponchak

The previous sections of this chapter have described the active elements used in MMICs, and although they are critical for circuit performance and reliability, it is the passive elements that determine the circuit's bandwidth, center frequency, and other electrical characteristics. Besides connecting the various active elements together, passive elements are used to set the bias point for the circuit and impedance match the active devices to themselves and the input and output connections of the MMIC. Passive elements are composed of lumped elements such as resistors, capacitors, and inductors and distributed elements such as transmission lines. In general, distributed elements are physically large enough that transmission line characteristics play a significant role in their function. Distributed elements have inductive, capacitive, and resistive aspects, all of which are taken into account by the transmission line analysis. The rule of thumb is that an element must be considered as a distributed element if it has dimensions greater than $\lambda/10$, where λ is the wavelength. Lumped elements on the other hand are small enough that transmission line effects do not play a significant role in their function. Nevertheless, even lumped elements are not purely inductive, resistive, or capacitive, but have aspects of all three due to parasitics. Because of these parasitic effects, lumped elements will resonate at some frequency. Failure to account for this complexity may lead to significant errors in circuit design. CAD programs such as SuperCompact and HP-EEsoF's Libra and Touchstone that are used in MMIC design contain models for most passive elements. Unfortunately, circuit designs do not always meet the modeled electrical performance goals because of unaccounted-for electromagnetic coupling effects and limitations of the models themselves.

The ability to fabricate MMICs and realize the advantages of size, ruggedness, reproducibility, and cost require that methods exist to fabricate the required passive components. Economic considerations require the MMIC chip size to be minimized as much as possible to maximize the number of chips on a wafer. Unfortunately, the passive components, especially inductive elements, tend to be large. In addition, the elements must be separated from each other to minimize electromagnetic coupling between the elements. The result is that the active elements occupy only a small portion of the MMIC area. Lastly, all passive components must be fabricated on the semi-insulating GaAs substrate.

The following is a brief description of passive elements, methods of fabrication, and the associated reliability concerns.

A. Resistors

Resistors are used for feedback circuits, setting the bias point of active devices, isolation, and terminations in power combiners and couplers. Two types of resistors are used in MMIC fabrication: thin films of lossy metals and lightly doped GaAs active layers. Figure 3-23 shows schematics of each of these two types of resistors. The resistance for both types of resistors may be given by

$$R = \frac{\rho_s L}{A}$$

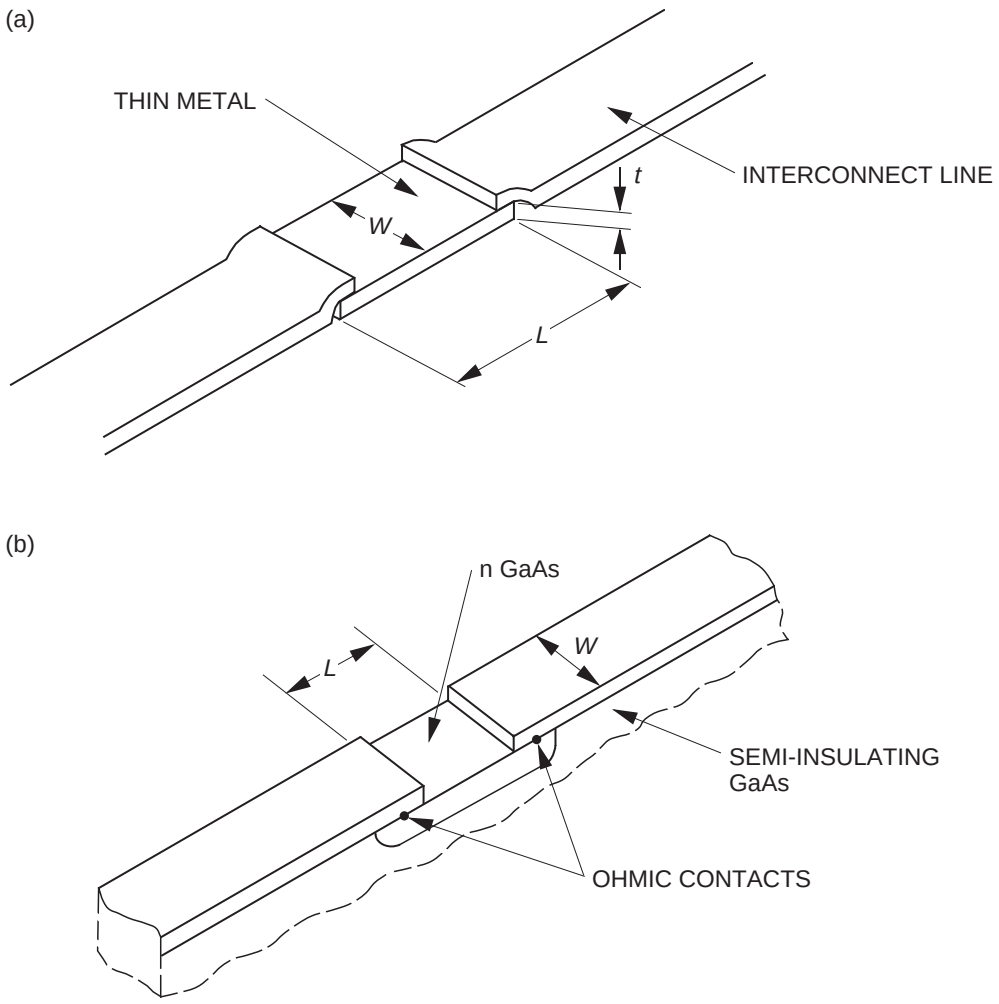

Figure 3-23. Two resistor types in MMIC fabrication: (a) thin film and (b) GaAs-based resistor incorporating an n GaAs channel and ohmic contacts.

where ρ_s is the sheet resistivity, $A = W*t$ is the cross-sectional area of the resistor, t is the resistor thickness, W is the width of the resistor, and L is the length. The efficiency of the resistor as determined by the resistance per unit length is a function of ρ_s. For metal thin-film resistors, ρ_s is a function of the metal. For GaAs based resistors, ρ_s is a function of the doping concentration.

Metal thin-film resistors are used for accurate, low-resistance applications. They are usually fabricated from TaN and NiCr, although Cr, Ti, Ta, Ta_2N, and AuGeNi alloys have also been used. Some of the advantages of thin-film resistors are a low Temperature Coefficient of Resistance (TCR), tight tolerances, small parasitics, and low sheet resistivity. The major disadvantage of thin film resistors is the added processing steps required to fabricate them, although thermal dissipation difficulties and electromigration failures are also a concern. Short resistors or those made from materials with a large resistivity have fewer parasitics, but they have a higher thermal load to dissipate. When resistors must dissipate large amounts of power, they can have the highest temperature on the MMIC and limit the MMIC reliability. Sidegating or the flow of current around the perimeter of the resistor is usually eliminated by depositing the resistor on top of an insulating film such as Si_3N_4. Electromigration failures result from the large current densities that can flow through the thin metal films. Tantalum resistors have exhibited

this problem for thin, 0.006-μm layers, with currents of 0.06 mA/μm of line width. Lastly, NiCr resistors are susceptible to degradation due to moisture. Therefore, these resistors must be passivated.

GaAs-based resistors are implemented by the use of an FET channel and ohmic contacts that are already available in the MMIC fabrication process. The total resistance of these elements is the sum of the resistance of the GaAs channel itself and the two ohmic contacts. The advantage of GaAs-based resistors is the wide range of resistivities available through changes in the doping level. GaAs-based resistors have several potential problems, however: current saturation, Gunn domain formation, and a high TCR. Above a critical electric field, the current in GaAs will saturate and the device loses its linearity. In practice, this is not a severe limitation since the length of the resistor is usually sufficient to prevent the electric field from reaching its critical value. Gunn domain formation, the initiation of microwave oscillations due to an applied static electric field, also occurs only if large electric fields are present. A more serious problem is the large positive TCR (+3000 ppm/°C). This can result in significant resistance changes over temperature. Fortunately, modeling techniques can be used to determine the resistance change that is tolerable for a given circuit and application.

The issues of current handling capacity, thermal dissipation, and distributed effects also play a role in the design and operation of GaAs based resistors. The resistor, especially one in the dc source or drain circuit, must be able to handle the current passing through it without reaching saturation. GaAs is a relatively poor thermal conductor and will not be able to remove heat rapidly enough if too much power is dissipated in too small an area. Physically large resistors on the other hand will become distributed elements and act as lossy transmission lines. In general, GaAs-based resistors may be thought of as a gateless FET. If the resistor is properly designed to operate below the current handling limit and the critical electric field limit, thermal heating should not be a problem.

B. Capacitors

Capacitance may be included in MMIC circuits in any of four basic configurations: an open-circuit transmission line, coupled lines or interdigitated capacitors, Schottky diodes, and metal-insulator-metal (MIM) capacitors. Coupled lines and open-circuit transmission lines can be used to provide fairly low capacitance values. For these two capacitor types, the capacitance is dependent on the electrical length of the transmission lines. Therefore, the capacitance is highly frequency dependent. The advantage of these capacitors is that they are easy to fabricate since they require only a single metal layer.

The most popular type of capacitor for MMICs has become the MIM capacitor because of the high capacitance per unit area that can be obtained. Therefore, smaller and less costly circuits are possible. A schematic of an MIM capacitor is shown in Figure 3-24. This thin-film capacitor is composed of two metal plates separated by a dielectric material. Typically, the dielectric material overlaps the first metal layer and the upper metal layer has a smaller area than the lower metal layer. This configuration helps to minimize fringing fields to ground and shorts between the upper and lower capacitor plates. Although the air bridge shown on Figure 3-24 is not required, it is often included to further minimize parasitic capacitance. The dielectric is typically silicon nitride (Si_3N_4, 0.1 to 0.4 μm thick) since it is already used in the MMIC fabrication process for circuit encapsulation, although SiO_2 and Ta_2O_5 are also used. Since the dielectric layer is substantially thinner than the substrate thickness, MIM capacitors exhibit significant

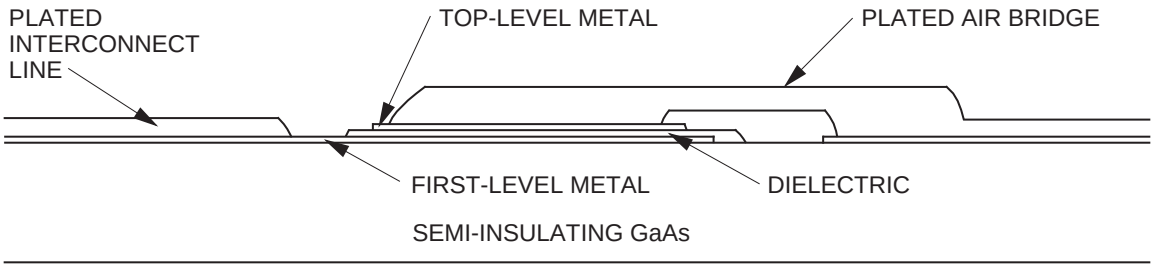

Figure 3-24. MIM capacitor using an air bridge for top-level interconnect.

fringing effects, which are a function of the perimeter. Careful experimentation to determine the magnitude of this effect for specific process parameters, such as dielectric type and thickness, is essential for any stable process. Test capacitors should also be included on the wafer for in-process verification.

The yield of MIM capacitors on a wafer plays a major role in determining the total yield for the wafer. One pinhole that ruins one capacitor also ruins the entire chip. The problem can be illustrated by considering a complex MMIC chip with ten capacitors and a capacitor yield of 95%. This case would produce a chip yield of only 60% on capacitor defects alone. The major yield limiting factor for MIM capacitors is shorts caused by pinholes in the dielectric or sharp points on the metal plates. Pinholes are very difficult to eliminate completely, but they can be minimized by good cleaning and deposition processes. Again, trade-offs require an engineering judgment based on the experimental results and the realized yields for a particular process. In addition to pinholes, the circuit design must assure that the electric field across the capacitor does not exceed the dielectric breakdown field .

C. Inductors

Inductors are necessary elements in MMICs where they function as tuning elements and RF chokes in dc bias circuits. Inductors are one of the easiest passive elements to fabricate. As a distributed element, they are realized by a section of high-impedance transmission line. These inductors are limited to inductance values below 2 to 3 nH because of the high losses associated with the long lengths of high-impedance transmission line. Lumped element inductors can be used to provide inductance up to 20 nH.

Lumped inductors are typically comprised of a transmission line in a spiral shape. A typical spiral inductor is shown in Figure 3-25. The total inductance is a result of the self-inductance of the high-impedance transmission line and the mutual inductance created by the electromagnetic coupling between the closely spaced lines. An air bridge is required to connect to the center tap of the spiral inductor. Although the spiral inductor is easy to fabricate, it is one of the most difficult devices to theoretically model because of the coupling between lines. Therefore, experimental characterization is usually required. Removable air bridges are often used in the first iteration to characterize an inductor with a different number of turns. In actual MMIC circuit designs, a model of these parts based on the measured data is used in the CAD programs.

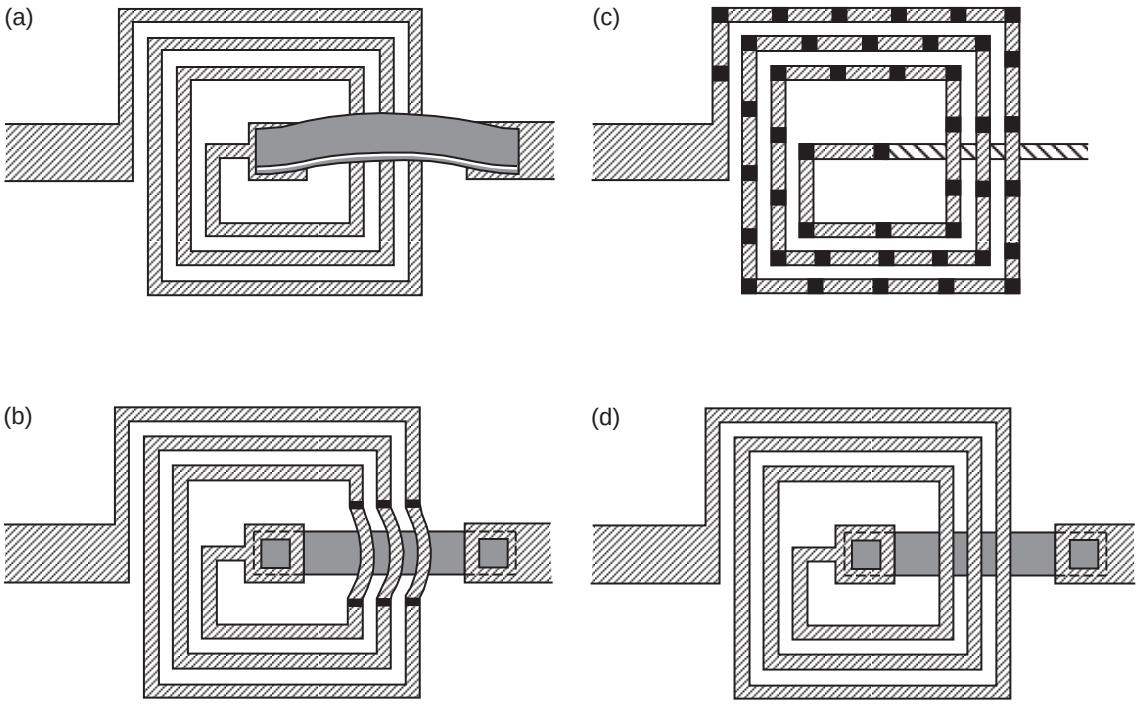

Figure 3-25. Spiral inductors: (a) as a single air bridge, (b) as air bridges over an underpass, (c) formed entirely of air bridges, and (d) using two metal levels for an underpass. (From [1].)

From a reliability point of view, the spiral inductor is a combination of transmission lines and air bridges. There are no special reliability issues associated with spiral inductors other than the air bridges and electromigration.

D. Transmission Lines

The dielectric constant of GaAs, coupled with the 100-μm-thick substrate commonly used in MMICs to facilitate heat removal, results in compact, narrow transmission lines. The critical transmission line parameters are the characteristic impedance of the line, the attenuation, and the frequency dependence of the phase velocity and impedance. The physics of transmission lines on dielectric substrates is rather complex, especially considering the realities of composite metallizations of finite thickness, discontinuities, radiation effects, and current distribution. The details of this topic are beyond the scope of this text, but adequate consideration should be given by the MMIC designer and the process engineer to understand the topic. CAD programs can provide ample support in obtaining reasonable analytical approximations, which can be used in the design stage to model the desired characteristics.

For MMIC purposes, the transmission lines are almost always microstrip, although coplanar waveguide transmission lines have been used. Both of these transmission lines are shown in Figure 3-26. The characteristic impedance of microstrip is inversely proportional to the ratio of the conductor width to the substrate thickness and also to the dielectric constant of the substrate. The conductor thickness also has a minor effect on the characteristic impedance. For GaAs substrates, the width-to-height ratio for a 50-Ω characteristic impedance is approximately 0.7. Therefore, for the typical 100-μm-thick substrate, the microstrip line width is 70 μm for a 50-Ω impedance. High-impedance lines have a smaller line width although lines thinner than 10 μm are rarely

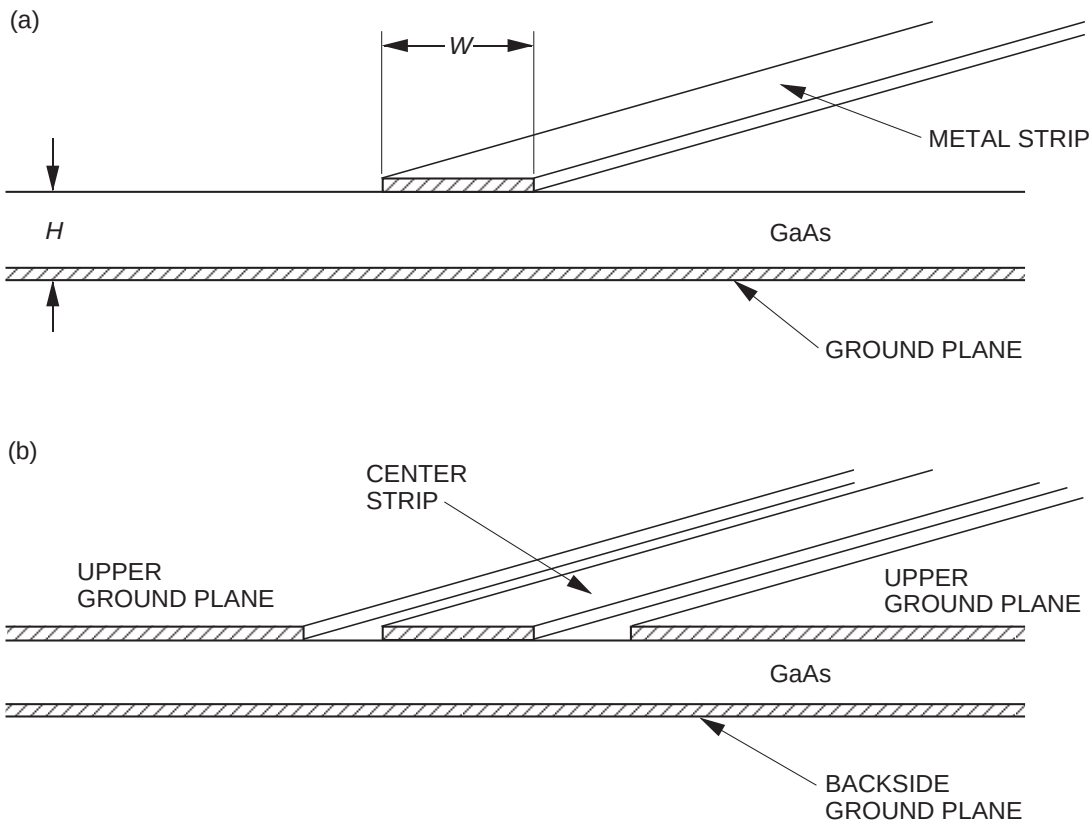

Figure 3-26. Transmission lines: (a) microstrip and (b) backside of grounded coplanar waveguide.

used because of the high conductor loss associated with the thin lines. Obviously, these are very large line dimensions compared to the gate widths discussed in Section 3-III. Electromagnetic coupling between transmission lines can be significant. Since the theoretical analysis of the coupling is difficult to perform, the transmission lines are normally separated by two to three line widths to minimize the coupling. The backside grounded coplanar waveguide has additional problems due to the excitation of parasitic modes that severely degrade circuit performance. Specifically, a slotline type of mode can propagate if the two ground planes are not held at the same potential and a parallel plate waveguide mode can propagate if the upper and lower ground planes are not held at the same potential. To accomplish these two requirements, air bridges and via holes are required.

The only reliability issues associated with transmission lines are electromigration and adhesion of metal on the substrate. Electromigration typically is a concern for the high impedance lines used in the dc bias circuit.

E. Via Holes

Both microstrip and backside grounded coplanar waveguide require via holes to either provide microstrip short circuits or to tie the upper and lower ground planes of the coplanar waveguide together. Via holes are etched through the GaAs substrate, usually from the backside of the substrate to minimize the top-side element area. This wet or dry etching process is followed by a gold sputtering step and finally a gold plating step to fill the hole with gold. A schematic of a filled via hole is shown in Figure 3-27.

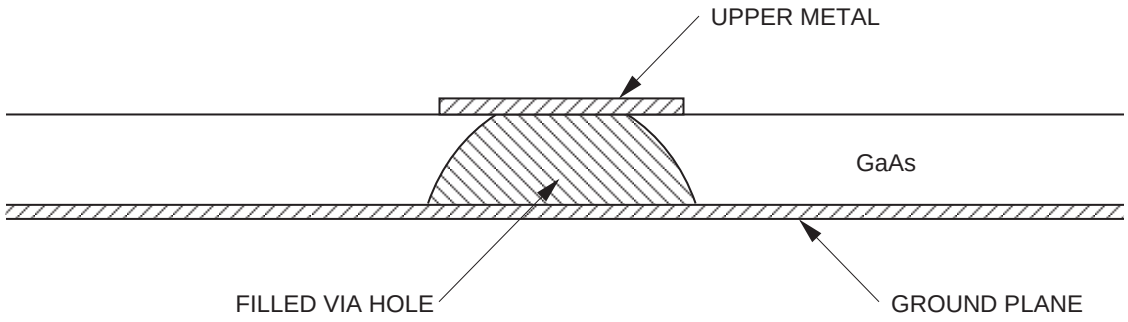

Figure 3-27. Ideal via hole in GaAs.

Several reliability issues arise from via holes. First, vias, if completely filled with gold, may cause cracking of the GaAs due to Au and GaAs thermal expansion mismatch. However, vias must contain enough gold to provide an acceptable thermal path. Since via holes are often used under the source contacts of field-effect transistors, an increase in thermal resistance may result in an increase in the junction temperature and a resulting degradation in the device reliability. Second, when via holes are placed directly under MIM capacitors, the capacitance yield is degraded because of the nonplanar via hole surface.

F. Air Bridges

Air bridges or crossovers are required for most MMIC layouts comprised of microstrip and absolutely necessary for coplanar waveguide circuits. In addition, air bridges are often incorporated into power transistors to connect the source leads of the parallel gates. An air bridge is simply a metallic bridge that permits two transmission lines to cross over each other without forming an electrical short circuit. A typical air bridge is shown in Figure 3-28. It is fabricated by depositing the first-level metal to form the transmission lines. Photoresist is then spun onto the wafer and holes opened where the bridge connections or posts are to be made. Then a second photoresist pattern is developed that defines the bridges. A gold layer is sputtered onto the wafer and the bridge interconnect metal is plated to the proper thickness.

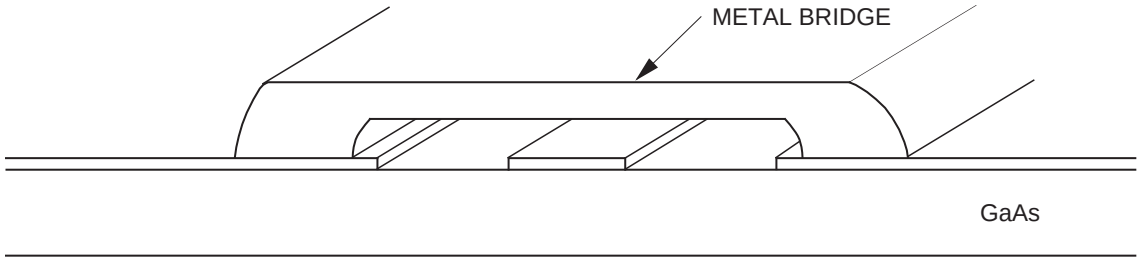

Figure 3-28. Air bridge connecting coplanar waveguide ground planes.

The reliability of air bridges depends on the bridge-metal thickness, the contact or post size, the bridge height, the interconnect metal used, and the bridge length. The primary concern is that the air bridge will sag and create a short circuit. Proper choice of the metal and the bridge length-to-width ratio and the application of dielectric coatings

under the air bridge can often minimize this effect. In addition, the bridge height and shape can be altered by simple changes in the photoresist processing steps. Typically, air bridges are on the order of several microns and ideally have an arched shape for strength. Once these factors have been established and followed as part of the design rules, the major failure modes are electromigration and cracking of the interconnect metal.

Reference

[1] M. Gillick and I.D. Robertson, "Passive Components," *MMIC Design*, I. D. Robertson, Editor, The Institution of Electrical Engineers, United Kingdom, 1995.

Additional Reading

Bahl, I., and P. Bhartia, *Microwave Solid State Circuit Design*, John Wiley & Sons, Inc., New York, 1988.

Colclaser, R. A., *Micro-Electronics: Processing And Device Design*, John Wiley & Sons, Inc., New York, 1980.

Goyal, R., *Monolithic Microwave Integrated Circuits: Technology & Design*, Artech House Inc., Norwood, 1989.

VIII. Basic Process Description

S. Kayali

A. Typical Ion-Implanted MESFET Process Flow

The starting wafers must be selected based on the specific process requirements. Low-noise processes require a different set of starting material characteristics than power processes, and each manufacturer has a defined set of wafer characteristics based on the selected process. A typical MESFET process flow is shown in Figures 3-29 and 3-30 with some optional steps shown for clarity and completeness of flow.

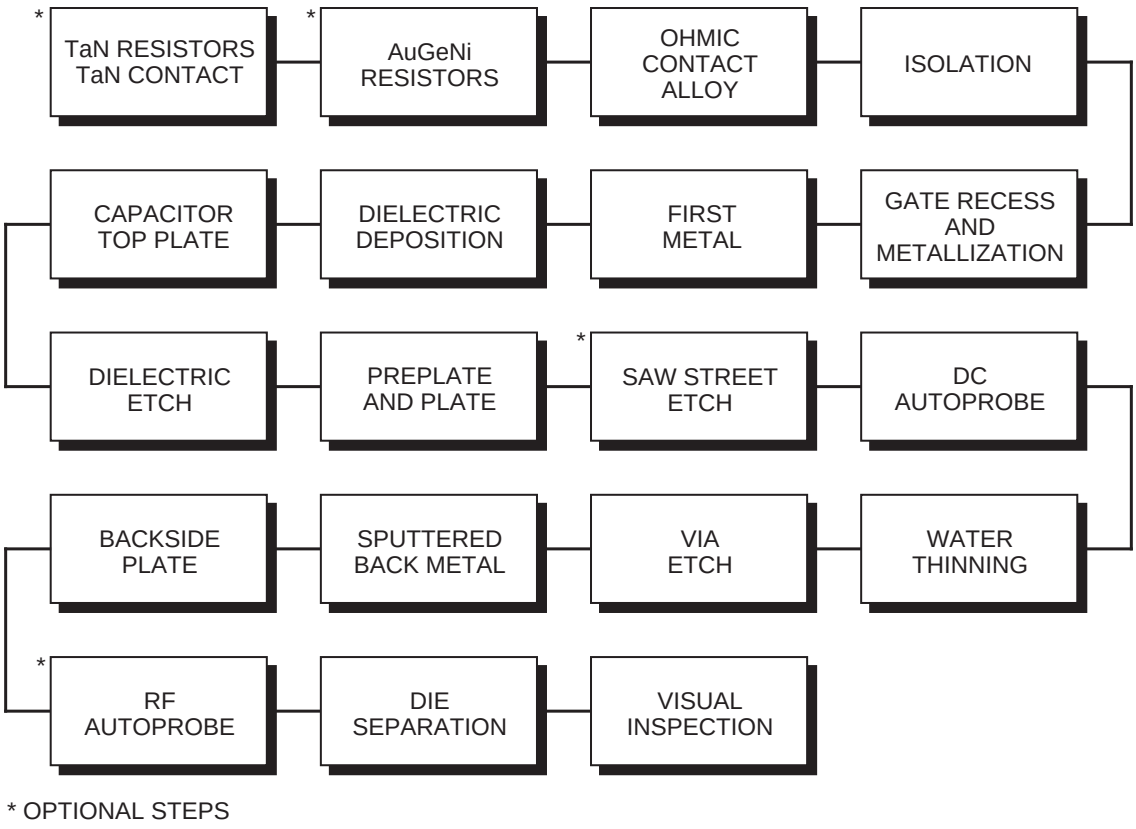

Figure 3-29. Basic sequence of process steps.

1. Resistor Deposition and Ohmic Formation

The first step normally involves the fabrication of the thin-film resistors. The AuGeNi resistor metal is evaporated, and the TaN resistor metal is sputtered and followed by a TaN contact-metal evaporation step. AuGeNi is normally used for designing low-value resistors, while TaN is used for medium-value resistors. An ohmic-contact deposition step normally follows with an alloy step, which results in a low-resistance ohmic contact to the active GaAs and also serves to stabilize the metal-film resistors.

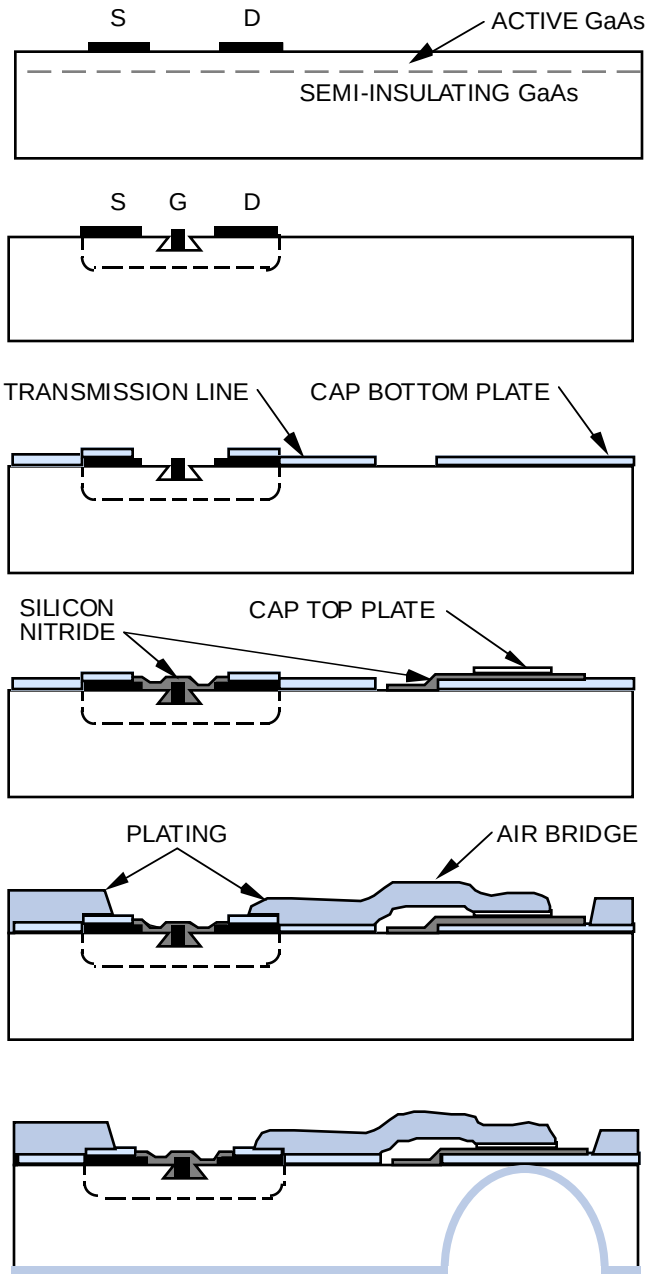

Figure 3-30. Basic process steps for MESFETs.

2. Isolation and Gate Formation

An ion implant such as boron is used to deactivate the conducting GaAs layer and form isolation patterns where desired. A direct-write e-beam can then be used to pattern the gate and gate recess in the active areas.

3. Metal and Dielectric Deposition

The first metal layer is normally an evaporated metal layer, which contacts the semi-insulating GaAs and forms the first-level interconnect. Dielectric deposition of

silicon nitride is used to protect circuit elements and provide a dielectric for capacitors. Capacitor top-plate metal is then deposited on top of the silicon-nitride dielectric. A pattern step is implemented to open the contact and define the bottom plate of the capacitor. Figure 3-31 shows a cross section of thin-film resistors and ohmic contacts on GaAs.

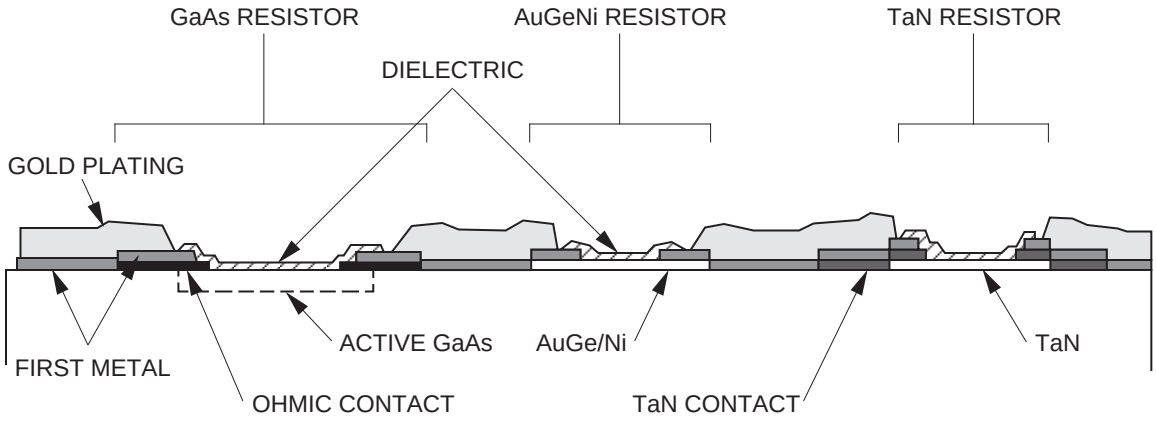

Figure 3-31. Basic process steps for GaAs, AuGeNi, and TaN resistor.

4. Plated Metal and Air Bridges

Plating is used to deposit thick layers of gold to construct air bridges, low-loss transmission lines, high-current-carrying lines, bond pads, resistor contacts, and evaporated metal step coverage. Two resist patterns are required to define the plated gold layer; the preplate and plate patterns work together to define the electroplated gold-metal level. The preplate layer defines the areas where the plating contacts underlying metal. The plate level defines the horizontal extent of the plated region. The preplate resist pattern is deposited directly on the front side of the wafer. Openings in the preplate resist are exposed and developed to define those areas where the plated metal will contact the underlying metal layer. The underlying metal is usually first metal, but plated metal can also contact capacitor top plates and other conducting layers.

After the preplate pattern is formed, a thin layer-to-metal contact is sputtered onto the entire wafer. This is the “preplate metal,” which serves to carry the electroplating current. On top of the preplate metal , a second resist pattern is formed to define the horizontal extent of any plated geometry, whether it is part of an air bridge or in contact with underlying metal. The preplate metal is removed, along with the photoresist, in all unplated areas. The preplate metal remains underneath all plated areas. Figures 3-32(a) and (b) show the air-bridge process.

5. Via Holes and Backside Processing

To allow for electrical connections between the frontside metal and the backside ground plane, via holes are formed and plated with gold. The size of the via hole depends on the substrate thickness; a circular pattern having a 50- to 60-μm diameter is normally used on 100-μm thick substrates. The plated gold layer also serves as the contact layer for die attach and a thermal path to a substrate. In processing, the via diameter at the frontside contact may vary from 12 to 160 μm on 100-μm-thick substrates. At backside, the via may be 2 to 3 times larger than the frontside pattern. Figure 3-33 shows the via hole and the process-dependent parameters.

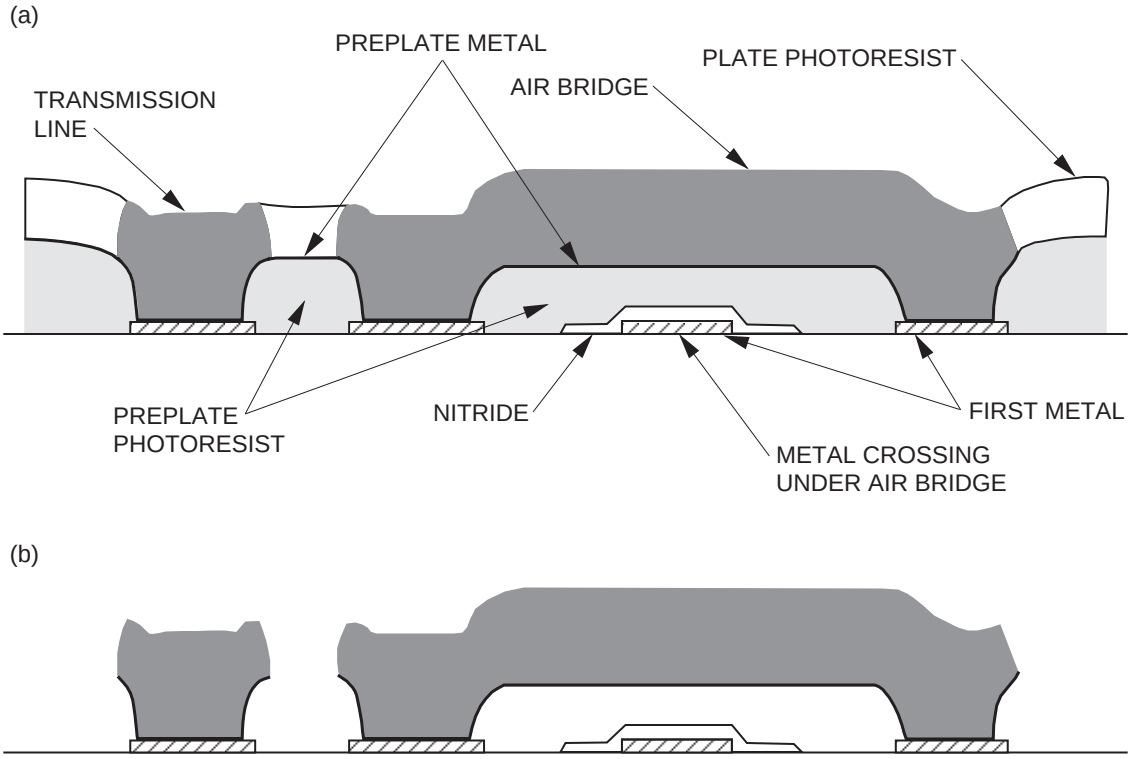

Figure 3-32. The air-bridge process: (a) with plate and preplate photoresist patterns and (b) after resist is removed.

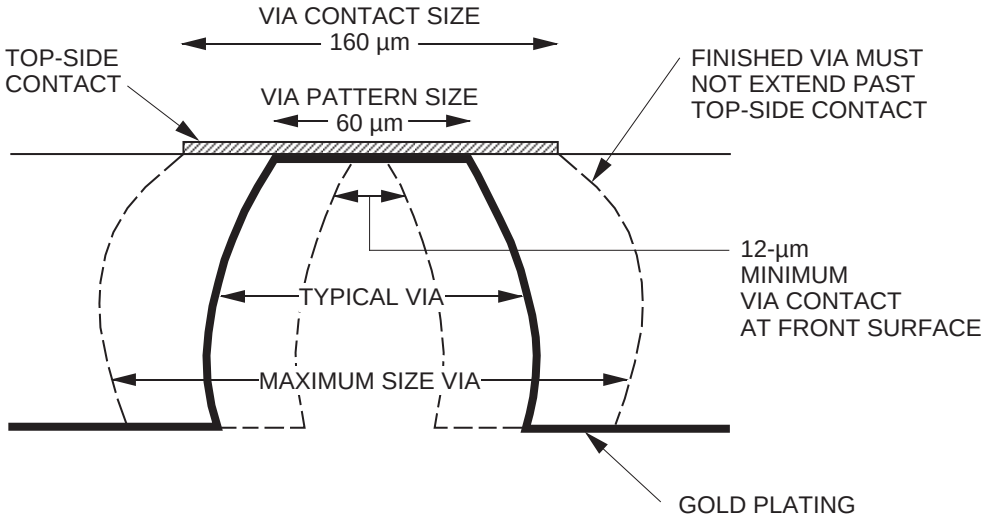

Figure 3-33. Via hole with process-variable parameters.

Reactive Ion Etch (RIE) is normally used to open the backside via holes. A layer of sputtered metal is then deposited over the entire backside of the wafer. Gold plating, approximately 6 µm thick, is then added to the sputtered metal for the die-attach capability.

The last step in the process is to physically separate the devices on the wafer. This is done by either scribing and breaking apart the devices or by using areas called saw streets, which are strips void of plated metal and are the outer boundaries of the individual MMICs. Final visual inspection is normally used in conjunction with dc probe data to select acceptable devices.

B. Typical HEMT/PHEMT Process Flow

The starting materials for HEMT-based devices require specific and stringent parameter control. Device manufacturers normally specify the applicable parameters that affect their process and are suitable to the processing flow. After the usual wafer cleaning and inspection, an epitaxial layer must be grown to provide the required material characteristics necessary for HEMT and PHEMT devices. The process starts with a GaAs buffer layer epitaxially grown to isolate defects from the substrate and provide a smooth foundation for further growth of the active layer of the transistor. A superlattice structure of undoped alternating layers of Al_xGa_{1-x} As and GaAs is them grown to further inhibit substrate conduction. The active channel is then grown using undoped GaAs in HEMT and undoped InGaAs in PHEMT. A spacer layer of undoped AlGaAs is then grown to separate the 2DEG from any ionized donors generated by the active layer. An n^+ AlGaAs donor layer is then grown to provide a source of electrons and to complete the growth of the HEMT/PHEMT epitaxial layer. Further processing steps complete the device structure and define the contact areas of the MMIC.

Differences exist in the fabrication of HEMT and PHEMT devices, but the general approach and processing flow remain essentially the same. Further details of the device structure and operation are found in Section 3-IV. The following brief description of the general processing flow is depicted in Figure 3-34.

1. Active Channel Definition and Isolation Implant

Using wafers with the epitaxial layers described above, a photoresist masking step is used to define the desired active channel area of the device. An isolation implant is then used to eliminate lateral conduction between devices. The resist layer can then be removed to expose the GaAs surface for the next process step.

2. Ohmic-Metal Formation

A photoresist masking step is used to define the desired ohmic-metal contacts. This metal is normally evaporated and then alloyed to provide the desired ohmic characteristics. Source and drain contacts along with capacitor bottom plates, resistor contacts, and, if applicable, inductors can be applied in the same process step.

3. Gate-Recess Formation

Photoresist masking and etching steps are then used to define the gate-recess areas. The etch depth and profile play major roles in the final characteristics of the device and should be carefully studied and characterized.

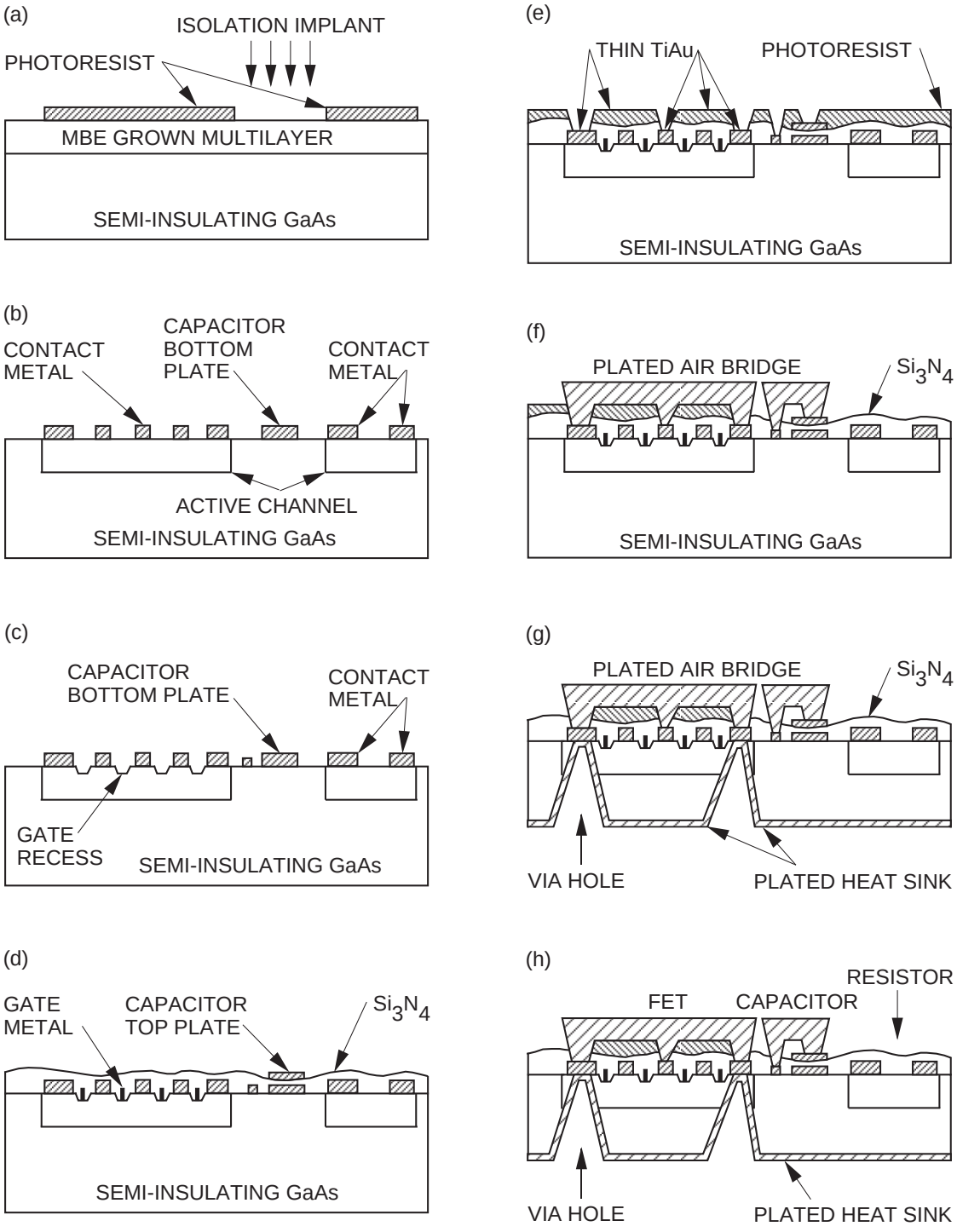

Figure 3-34. Typical HEMT/PHEMT process flow: (a) active channel definition and isolation implant, (b) ohmic-metal formation, (c) gate-recess formation, (d) gate-metal formation and nitride deposition, (e) source and contact etch, (f) air-bridge formation, (g) via-hole formation and backside processing, (h) completed typical MMIC structure.

4. Gate-Metal Formation and Nitride Deposition

Gate metal is normally sputtered or evaporated to form the desired Schottky contact. A nitride layer covering the active areas and forming the dielectric layer for capacitors is then applied. The actual method and conditions for gate-metal and nitride application are critical to the performance, characteristics, and stability of the devices being manufactured.

5. Source and Contact Etch

A photoresist masking step is normally used to pattern the source and other contact openings. A thin layer of TiAu is also applied to help in the next step of air-bridge plating. At this step, wafers can be mounted face down to perform the thinning operation. The wafers are normally thinned down to a thickness of about 25 mils.

6. Air-Bridge Formation

After cleaning the front side of the wafers, a thick layer of Au in normally plated to form the air-bridge structures. An etch to remove photoresist for the undesired locations is also performed.

7. Via-Hole Formation and Backside Processing

A photoresist masking step is used to identify the desired via-hole location. An etch after exposure of the photoresist can provide the backside contact. An Au layer is normally plated to serve as an electrical and thermal path.

8. Complete Typical MMIC Structure

The completed MMIC structure is now ready for probe test, dicing, and further packaging.

C. Typical HBT Process Flow

This section describes a typical mesa process flow used for fabrication of devices based on a non-self-aligned HBT process. Other possible processes and fabrication flows are practiced by various device manufacturers to provide higher yield and device performance.

The device fabrication sequence basically consists of etching steps to reveal the various layers in the structure and fabricating electrical contacts to each layer. Finally, devices are isolated and interconnections are made within each device as well as between devices. The following general steps as shown in Figure 3-35 are applicable to this flow:

1. Emitter Contact

The lift-off technique is generally used to achieve an ohmic (nonrectifying) contact on GaAs. Typically, AuGe/Ni/Au layers are evaporated in sequence and then alloyed to form the ohmic contact. The choices of alloying temperature and time play an

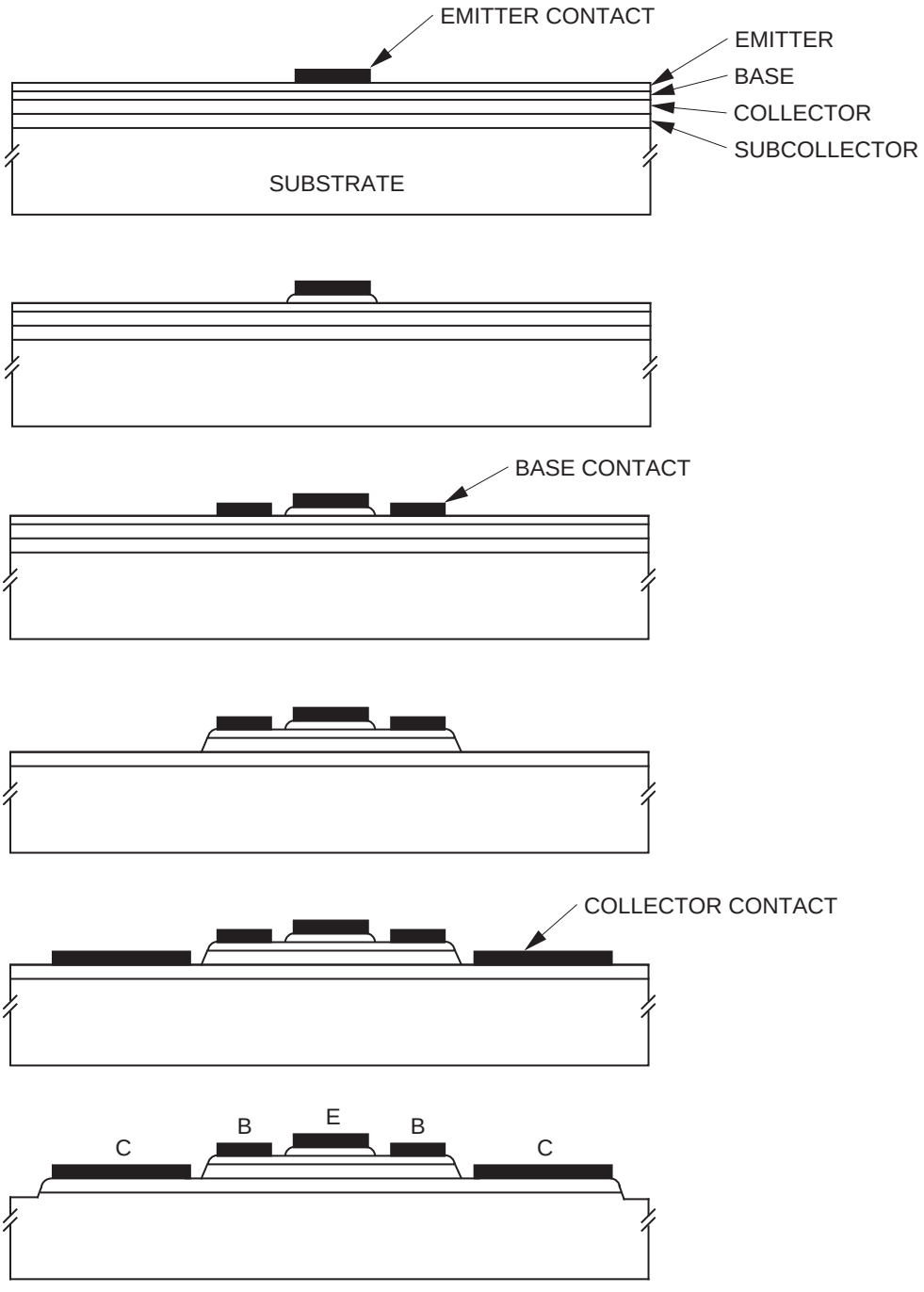

Figure 3-35. A typical HBT device processing sequence. (Courtesy of Artech House.)

important role in the quality and long-term stability of the contact. Other materials such as AuGe/Ni/Ti/Au, PdGe and Ti/W/Au can also be used for achieving this application.

2. Base Contact

A photoresist layer is generally used to mask the defined base contact area, a wet-etching or dry-etching step is then used to remove the emitter layer at the defined location, and the lift-off technique is again used to form the base contact. Alloyed contacts using materials such as AuBe, AuZn, or AuMg are common for ohmic contact formation. However, nonalloyed contacts using materials such as Ti/Pt/Au are also common.

3. Collector Contact

A process similar to that used for the emitter contact is used to achieve the collector contact. A photoresist layer masks the desired area, and the defined area is then etched using either a dry or a wet etching step. Finally, the lift-off technique forms the collector contact using the same metallization scheme as that used for the emitter contact.

4. Isolation and Interconnection

Device-to-device isolation is generally achieved by the use of wet or dry etch to remove the subcollector layer, as shown in Figure 3-35. Ion implantation can also be used to achieve the same result. Interconnection, on the other hand, can be achieved by the use of air bridges.

Other processing techniques employing self-aligned contacts (Figure 3-36) and planar structures (Figure 3-37) can provide improved device performance and higher levels of integration. Ion implantation is generally used to form conductive channels and semi-insulating layers. Planar structures can provide the added advantage of access to the various device terminals without the need of air bridges. This can provide a substantial improvement in device yield.

Additional Reading

Ghandi, S. K., *VLSI Fabrication Principles*, John Wiley & Sons, New York, 1982.

Texas Instruments GaAs Foundry Services Design Guide, Texas Instruments, 1993.

Williams, R., *Modern GaAs Processing Methods*, Artech House Inc., Boston, 1990.

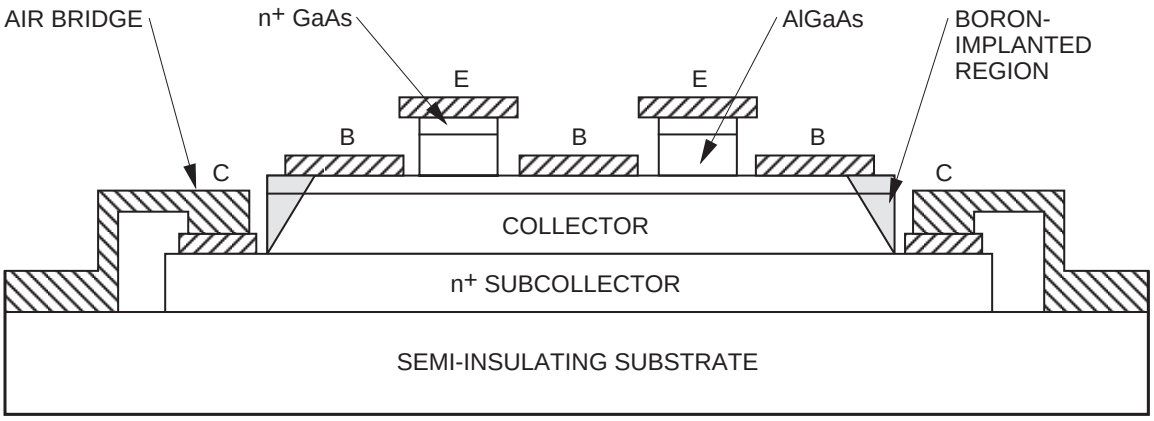

Figure 3-36. Self-aligned HBT cross-sectional view. (Courtesy of Artech House.)

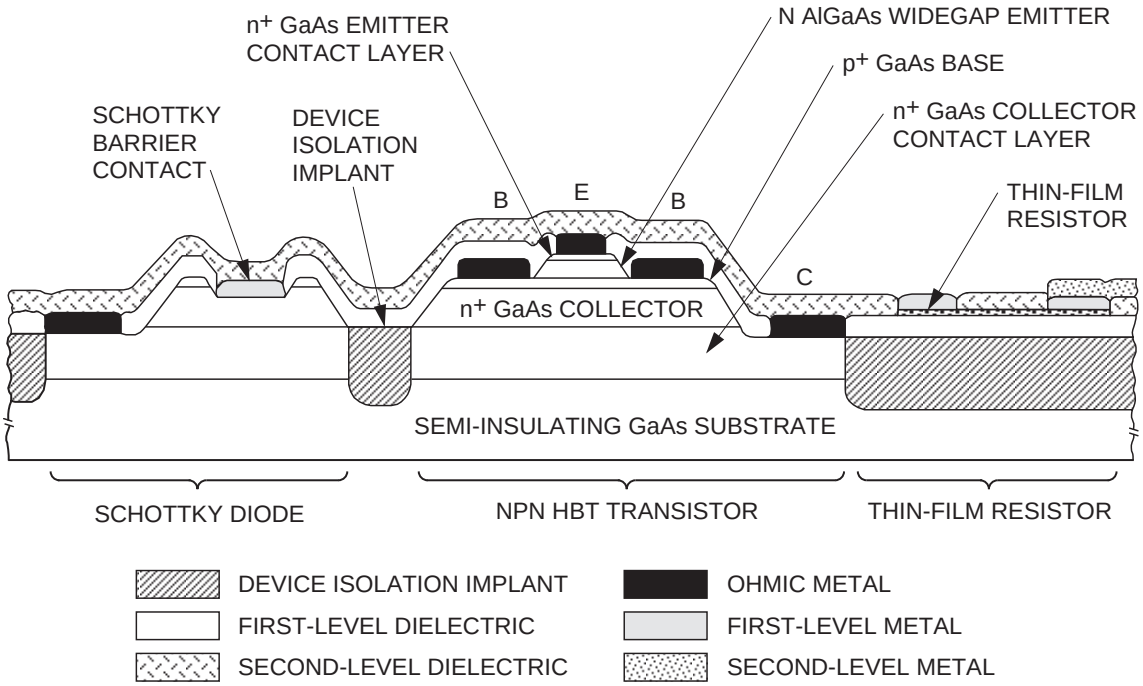

Figure 3-37. Schematic cross section of the self-aligned HBT IC structure with integrated NPN transistor, Schottky diode, thin-film resistor, and metal–insulator (SiN)–metal capacitor (not shown). (Courtesy of Artech House.)

IX. Monolithic Microwave Integrated Circuits

A. N. Downey, G. E. Ponchak, and R. R. Romanofsky

A. General Description

Monolithic Microwave Integrated Circuits (MMICs) are used in satellite systems that require smaller, less expensive circuits or when the parasitic reactance inherent in hybrid integrated circuits degrades the circuit performance, typically in the upper microwave and the millimeter-wave spectrum. Examples of systems that use MMICs are receivers and transmitters for communications, phased-array antennas where small size and uniform circuit performance are required, and sensors and radars that operate at high frequencies. The types of circuits required for each of these systems are illustrated by examining the simple receiver and transmitter systems shown in Figures 3-38 and 3-39, respectively. In both schematics, a phase shifter—which may be placed in either the local oscillator (LO), the RF, or the IF portion of the system—has been added to make the system perform as if each circuit were coupled to a single radiating element of a phased-array antenna. For non-phased-array applications, the schematic is unchanged except for the removal of the phase shifter. A photograph of a completely monolithic 30-GHz receiver is shown in Figure 3-40. Although the high level of circuit integration illustrated in Figure 3-40 decreases the packaging and interconnect costs, this integration is not necessary or common. Instead, each function of the system is typically fabricated on an individual die to permit the optimization of the material system and device type for each application. Regardless of the level of circuit interconnection, the reliability of the system is dependent on the continuous operation of each circuit.

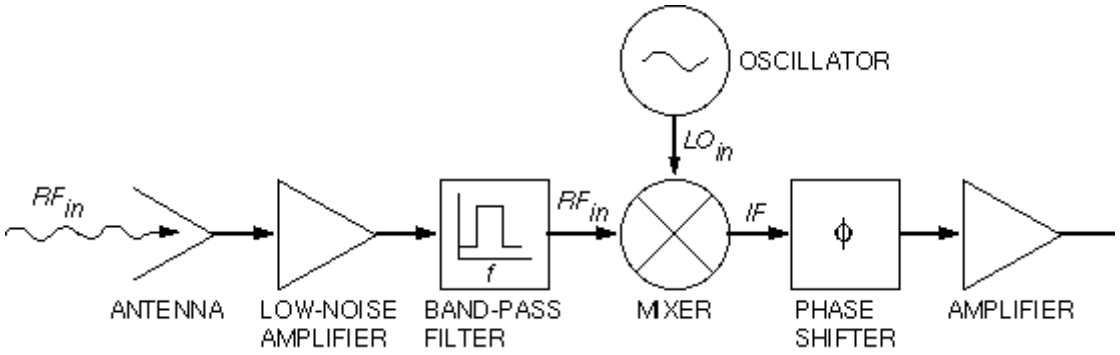

Figure 3-38. Schematic of microwave receiver.

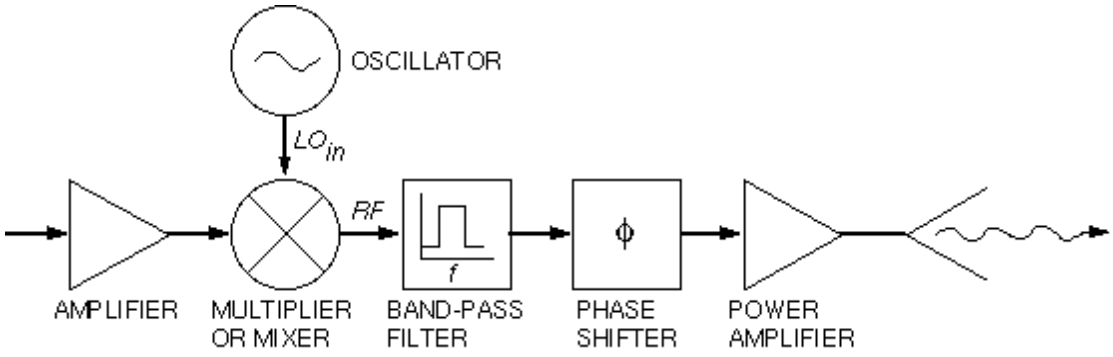

Figure 3-39. Schematic of microwave transmitter.

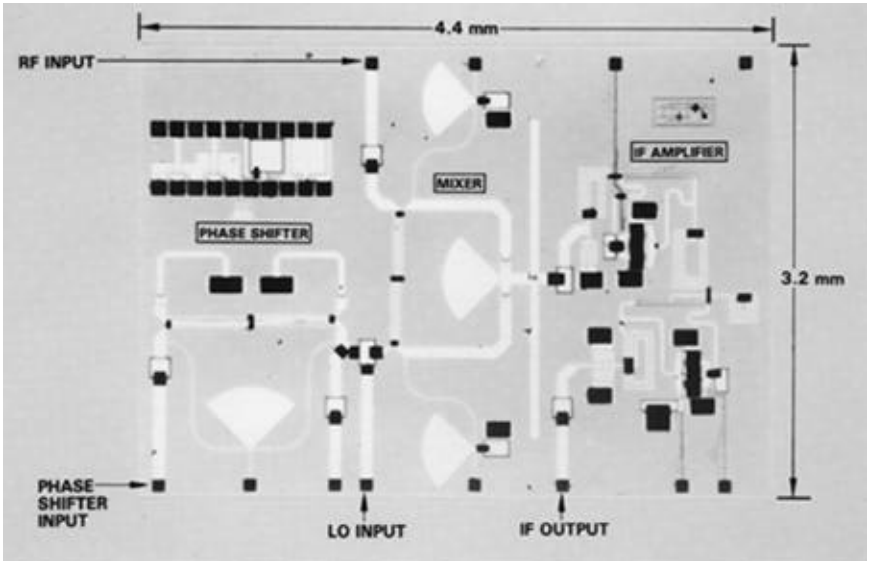

Figure 3-40. 30-GHz MMIC receiver. (Fabricated by Hughes Aircraft Company for NASA Lewis Research Center.)

This is understood by examining the receiver circuit shown in Figure 3-38. The input (RF) signal typically has a very low power level that may be close to the noise floor. The low-noise amplifier (LNA) amplifies the received signal while at the same time introduces very little new noise. If the gain of the LNA is sufficiently large, the noise contributions of the rest of the system will be small since the noise created by later circuits is divided by the gain of the LNA. Thus, the LNA gain and noise figure, the measure of noise added by the LNA, determine the receiver noise characteristics. If the receiver has poor noise characteristics, it will not be able to receive weak signals. The signal may then pass through a narrow-band filter and into the mixer. The LO generates a signal that is also fed into the mixer. The mixer combines the two signals through a nonlinear device, such as a MESFET or diode, and generates a signal at the intermediate frequency (IF) of $f_{RF} - f_{LO}$ or $f_{LO} - f_{RF}$ and harmonics of the IF, RF, and LO frequencies. All but the desired IF components must be filtered out. The conversion efficiency of the mixer is usually dependent on the LO drive power. In addition, a variation in the LO frequency will cause a shift in the IF that may cause the signal to be attenuated in the narrow-band filters that are part of the mixer. If the system is to be associated with a phased-array antenna, the direction and shape of the main beam radiated or received by the antenna is dependent on the relative phase shift and power level of each transmitter (and receiver). The relative phase of each radiating element is set by the phase shifter. Thus, if the phase shift through the circuit varies because of unexpected conditions, the efficiency of the entire antenna will degrade. It is thus seen that a parametric shift by any of the components may cause the entire system to fail.

The phase shifter, local oscillator, and mixer circuits are common to the receiver and transmitter with the exception of a shift in the design frequency. The real difference

between the two systems is in relation to the amplifiers. As described above for the receiver, the LNA must be capable of amplifying a weak signal sufficiently for the mixer to work and for the noise contributions of the rest of the system to be minimized, while at the same time introducing as little new noise as possible. In a transmitter, the critical performance specifications are the amount of power transmitted and the efficiency of the circuit. Thus, the power amplifier must be able to provide gain to a very strong signal.

In early MMICs, all of the circuits were made from GaAs MESFETs, impact ionization avalanche transit time (IMPATT) diodes, and varactor diodes, but as GaAs technology matured, HBTs, HEMTs, and PHEMTs have found increasing use in niche applications. Table 3-4 identifies the devices now most commonly used in each of the circuits. Although most MMIC failures originate at one of the active devices, some reliability concerns relating to each specific circuit will be presented below. A more detailed discussion of reliability problems related to specific devices or components is provided in Chapter 4.

Table 3-4. Matrix of solid-state devices and their applications in MMICs.

Device	Varactor Diode	Schottky Diode	PIN Diode	HBT	MESFET	HEMT	PHEMT
Low-noise amplifier, GHz							
f < 12				√	√	√	√
12 < f < 26						√	√
f > 26						√	√
Power amplifiers, GHz							
f < 12				√	√	√	√
12 < f < 26				√	√	√	√
f > 26						√	√
Mixers, GHz							
f < 12		√		√	√	√	√
12 < f < 26		√		√		√	√
f > 26		√		√		√	√
Oscillators, GHz							
f < 12	√			√	√	√	√
12 < f < 26				√	√	√	√
f > 26				√	√	√	√
Multipliers, GHz							
f < 12		√			√	√	√
12 < f < 26		√				√	√
f > 26		√				√	√
Analog phase shifter	√						
Switched-line phase shifter			√	√	√	√	√

B. Amplifiers

Both low-noise and power amplifiers are used to increase the power of the RF signal. In almost all systems, this is accomplished by using the transconductance of MESFETs and HEMTs or the current gain of HBTs. The amount of signal increase is called "gain" and is usually given in dB, where gain in dB = 10 log (gain). For example, if the output power is twice the input power, the amplifier has 3 dB of gain. Typically, the input power and the output power are also specified in dB, permitting the output power to equal the sum of the input power and the gain. This ideal operation of an amplifier is accurate for low power levels. Unfortunately, as power levels increase, the amplifier becomes nonlinear. In the nonlinear region of operation, the output power is less than the sum of the input power and the amplifier gain in the linear region, or it can be stated that the amplifier gain is lower in the nonlinear region. Figure 3-41 shows a typical amplifier characteristic. The point at which the output power drops by 1 dB from the linearly extrapolated value is called the 1-dB compression point [1]. This value separates small-signal or linear amplifiers from large-signal or power amplifiers. Note that this is also the criteria used to differentiate small-signal and large-signal transistors, since a transistor can be viewed as a simple, unmatched amplifier. This differentiation is important in determining the failure mechanisms that need to be addressed and the type of reliability tests that should be performed.

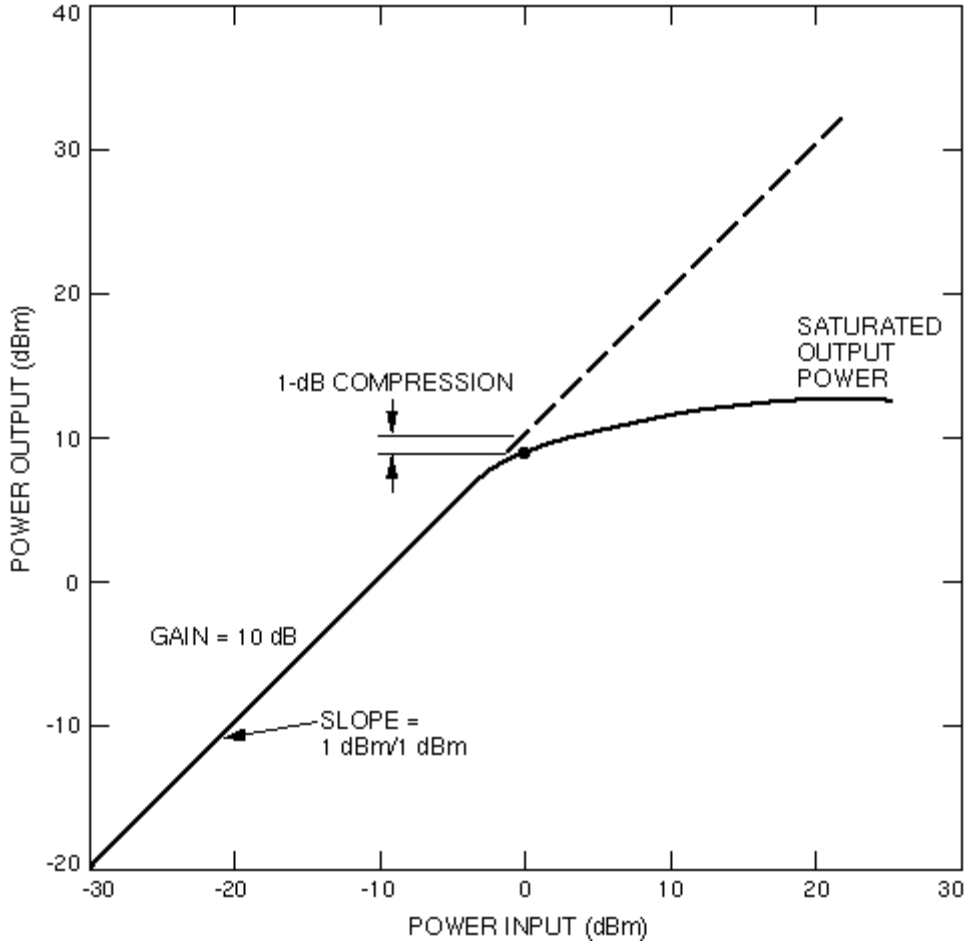

Figure 3-41. Output power as a function of input power for a typical amplifier.

The choice of the bias point is critical in the amplifier operation. If the bias point is chosen so that the output signal from the power amplifier appears as an amplified

version of the input signal over the entire period of the voltage wave, the amplifier is called a Class A amplifier. More typical of power amplifiers, the large voltage swing of the input signal will cause the power amplifier to operate at a bias point such that the current is in cutoff or saturation over part of the input signal voltage swing. Thus, over part of the input voltage swing, the output waveform will be zero. If the output signal is zero over half of the period, the amplifier is called Class B. Other classes of amplifiers are based on the amount of time the output signal is at zero voltage. The choice of the bias point or amplifier class determines the linearity and power-added efficiency defined as

$$\eta = \frac{P_{out}^{RF} - P_{in}^{RF}}{P^{DC}}$$

Class A amplifiers are linear, but since they draw dc power over the entire period, they are not efficient. On the other hand, Class B amplifiers are not as linear, but since the amplifier is not drawing dc current over half of the input voltage swing, they have higher efficiency [2].

1. Power Amplifiers

Power amplifiers, by their very nature, must handle high input and output powers. The maximum voltage swing of the input signal is limited by the breakdown voltage of the transistor, and thus transistors with high breakdown voltages are required. The current through each transistor is limited by the resistance in the gate or emitter of FETs and HBTs, respectively, since ohmic losses are converted to heat, which decreases the device's reliability. To increase the current handling capability of the device, power transistors combine many gates or emitters in parallel. This parallel combination increases the total gate width or emitter area and decreases the resistance, while at the same time increases the difficulty in matching the input impedance of the transistor to the output impedance of the prior stage. In addition, the spacing required between the transistor elements to permit sufficient thermal dissipation creates large devices that are more difficult to maintain with a uniform voltage [3]. To dissipate the heat from the transistors, power amplifiers are fabricated on thin wafers, less than 100 μm thick and typically between 25 and 50 μm, to reduce the thermal path between the transistor's active region and a good heat sink, such as a metal or diamond carrier. Generally, thermal constraints limit the design and performance of power amplifiers more than frequency constraints. Thus, the efficiency of power amplifiers is one of the most critical specifications, especially in space applications where satellite power is limited, where dissipation of the thermal load requires heat sinks that increase the system weight, and where circuit heating can decrease reliability.

Power amplifiers designed with multiple stages (one stage is one transistor or one parallel combination of transistors) are used to accommodate the thermal constraints, peak-voltage and current constraints, and limited gain available from each transistor. Figure 3-42 shows a 30-GHz power amplifier with three stages that consist of increasingly larger transistors. The number of stages required in the amplifier is dependent on the gain specification and frequency, since transistor output power decreases with increasing frequency. Since the power dividing and combining networks on the MMIC will typically introduce 0.5 to 1 dB of loss and the input impedance of transistors decreases with an increasing number of gate fingers, the degree of power dividing and combining that can be used to increase the power level is limited.

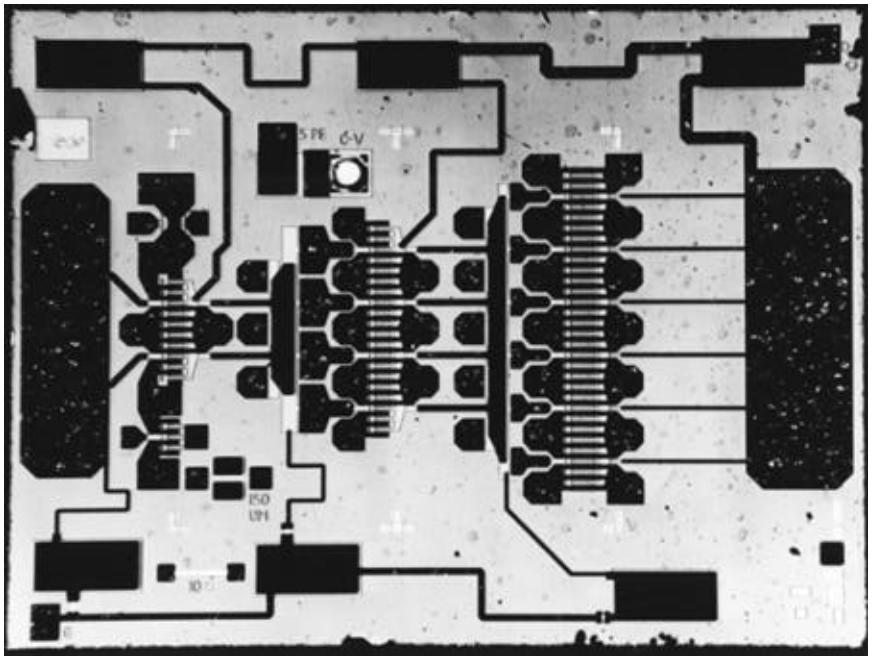

Figure 3-42. 20-GHz high-power amplifier. (Fabricated by Texas Instruments under contract to NASA Lewis Research Center.)

Models for microwave devices, both active and passive, are usually derived from *S* parameters measured on a vector network analyzer. These models are good for low-power circuit designs, but transistors exhibit significant nonlinearity or a power dependence at high-power levels. Therefore, nonlinear models based on load–pull measurements are required for high-power designs. In addition, the nonlinearity of the power transistors creates intermodulation distortion (IMD), which is power at frequencies other than the input frequency: $2f_{RF}$, $3f_{RF}$, etc. Failure to account for these frequency terms in the matching circuit can lead to signal distortion, oscillations, lower efficiency, and package resonances. IMD is specified as the ratio of the power at the IMD frequency to the power at the desired frequency and is usually given in dB. The ability of engineers to design a power amplifier will depend on the availability of good nonlinear models.

Besides thermal-stress-related problems, power amplifiers exhibit some unique failure mechanisms, such as hot-electron trapping, which is covered in Chapter 4. Electromigration and metal diffusion must also be addressed in power amplifier designs due to the large currents and high voltages used during operation.

2. Low-Noise Amplifiers

Since low-noise amplifiers are used on the front end of receivers, they are designed to handle very low power levels. Thus, the thermal problems and high bias currents and voltages that affect power amplifier reliability are generally not a concern for LNA designers. The most important criterion in specifying or measuring an LNA's performance is the noise figure, and since HEMTs and PHEMTs have the lowest noise figure, they are used in almost all LNAs. To minimize the noise figure, small gate lengths and low parasitic gate and source resistances are required [4]. Thus, state-of-the-art LNAs are usually comprised of 0.1 to 0.25 μm gate-length HEMTs or PHEMTs, and the reliability concerns—such as gate metal sinking and ohmic contact diffusion (see Chapter 4)—arising from small gate lengths and corresponding small channel thicknesses are the most important.

To decrease the noise figure of the system, it is important to reduce the circuit losses, especially before the first stage of the LNA. This includes the package feed losses and transmission line losses from the antenna since they introduce noise into the system before the LNA. Besides reducing the circuit losses, noise can be reduced by operating the amplifier at lower temperatures and lower bias currents and voltages. Lastly, the noise figure of the LNA is dependent on the matching circuits, which are designed with an input matching network that minimizes the noise figure and an output matching network to maximize the gain. The optimum input matching network can be found through noise parameter measurements of the HEMT. From these measurements, an equivalent circuit model of the HEMT that includes noise sources can be generated.

C. Mixers

Mixers convert an input signal at one frequency to an output signal at another frequency to permit filtering, phase shifting, or some other data processing operation at a frequency more easily implemented by the circuits. For example, a system may require the data to be received at W-band, 75 to 110 GHz, but W-band filters have a low *Q* or a high loss, which degrades the receiver noise characteristics. Therefore, it may be advantageous to shift the received signal's frequency to a lower value where low-loss filters are possible. Ideally, this operation is accomplished without degrading the input signal's amplitude or introducing additional noise.

Frequency conversion is accomplished by devices with nonlinear I–V characteristics. Early mixers were all made with diodes, but MESFETs, HEMTs, and PHEMTs have been used more as the technologies have matured. Consider first the diode mixers that can be represented in the simple diagram shown in Figure 3-43. The nonlinearity of the diode I–V characteristics is given in Equation (3-6) of Section 3-II, which is plotted in Figure 3-44. If two voltage signals, labeled the LO and the RF signals, are placed across the diode terminals, the output current shown in Figure 3-44 can be represented by [5]

$$i(v) = I_0 + A\sum_{j=1}^{\infty} B_j\left[V_{LO}\sin\left(\omega_{LO}t\right) + V_{RF}\sin\left(\omega_{RF}t\right)\right]^j$$

which, upon performance of some trigonometry, can be shown to yield signals with frequencies of

$$f_0 = mf_{RF} \pm nf_{LO}$$

Usually, the desired output frequency is $f_{RF} - f_{LO}$ and this frequency is called the intermediate frequency or the IF.

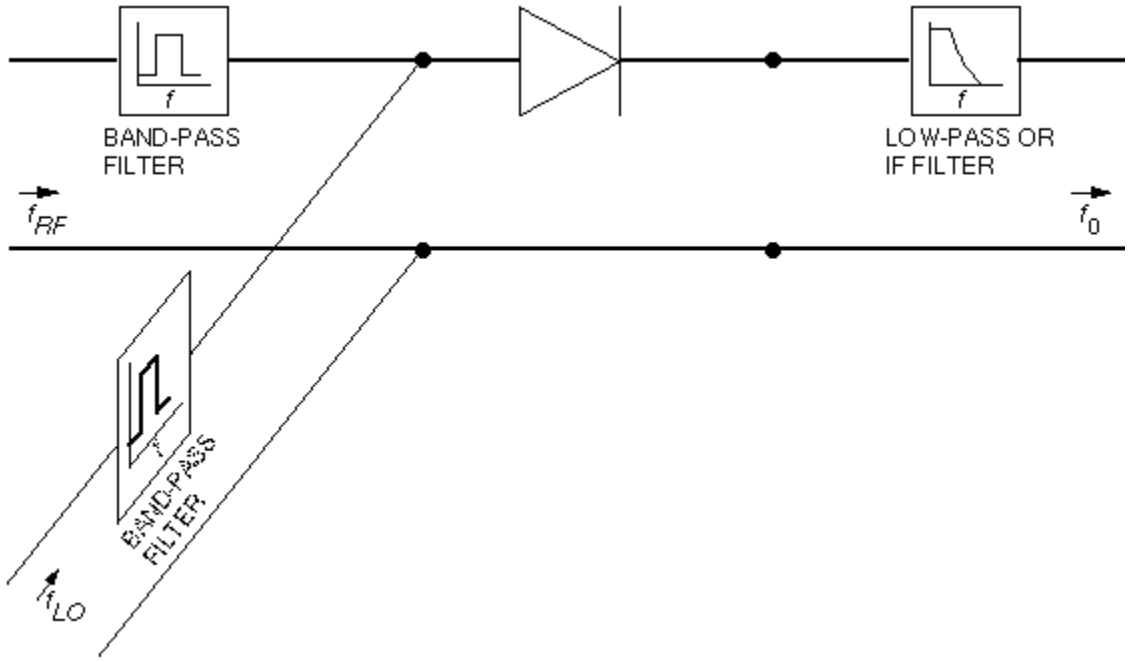

Figure 3-43. Schematic of simple mixer.

A figure of merit for mixers is the ratio of the IF power to the RF power, which is called the conversion loss and is usually specified in dB. There are several contributions to conversion loss. The first is due to poor impedance matching at the RF and IF ports. The second is due to the I–V characteristics of the diode junction, which, if optimized, yields a minimum conversion loss contribution of 3 dB while the remaining half of the power is converted to other frequencies, primarily the image frequency, which is $f_{RF} - 2f_{LO}$ or $2f_{LO} - f_{RF}$. The final contribution to conversion loss is due to the diode parasitics [3] and is given by

$$loss(dB) = 10\log\left(1 + \frac{R_{ohm} + R_{chan}}{R_j} + \left(\omega C_j\right)^2 \left(R_{ohm} + R_{chan}\right) R_j\right)$$

where the parameters are shown in Figure 3-11(c). This loss is minimized when R_j equals $1/\omega C_j$, but since C_j is dependent on the LO power, the conversion loss of a diode mixer is strongly dependent on the LO power and typically decreases to a minimum value with increasing LO power. In addition to optimizing the LO drive power, the cutoff frequency of the diode should be at least 10 times greater than the RF, LO frequency, and IF.

Most FET mixers rely on the nonlinearity of the transconductance by applying the LO and RF signal to the gate of the FET and extracting the IF from the drain. The advantage of FET mixers is that the transistor provides gain that yields mixers with conversion gain instead of conversion loss. The disadvantage of FET mixers is that they also amplify low-frequency noise, $1/f$ noise, which can be converted to a frequency in the desired spectrum.

Note that power at all frequencies—whether from the LO, the RF, noise, or LO instability—applied across the nonlinear device generates power at other frequencies. The elimination of these noise-generated signals and the harmonic frequencies of the RF and LO is critical to the system performance. Many design configurations are possible

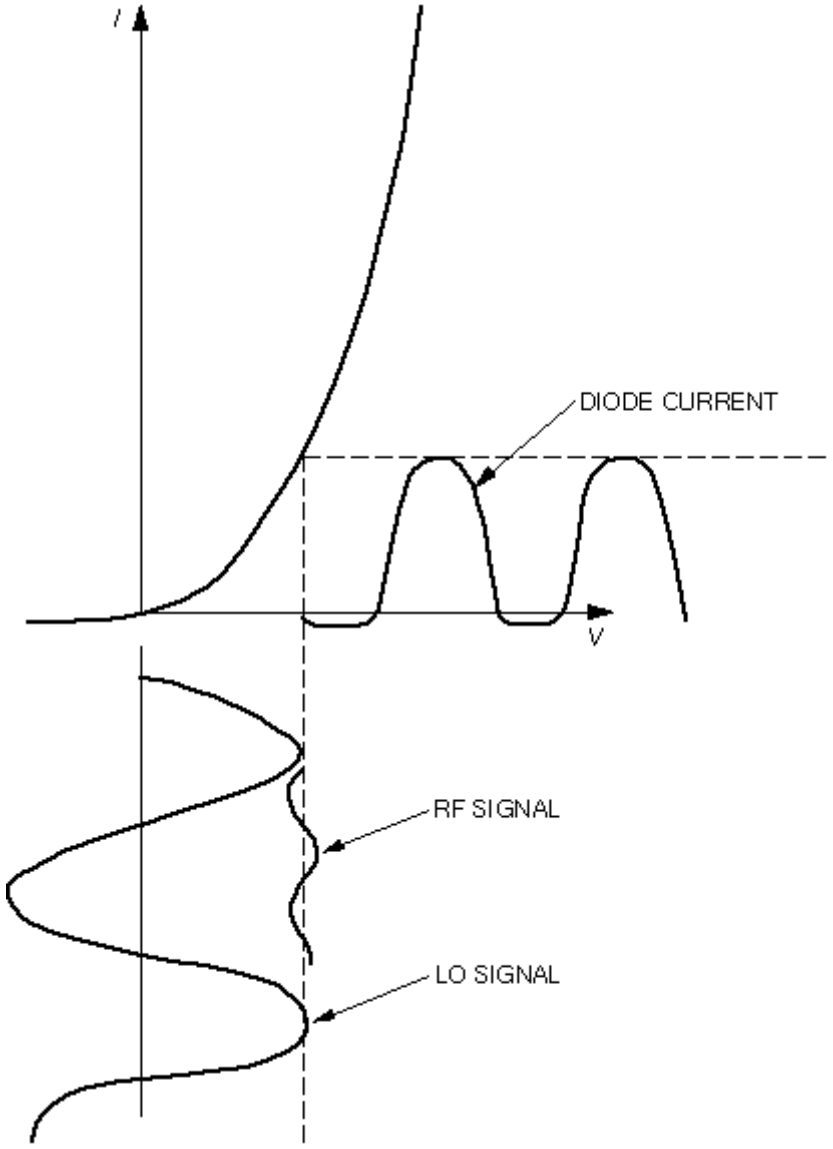

Figure 3-44. Mixer diode I–V characteristics.

from the simple single-ended mixers that require only one diode or FET (Figure 3-43) to those requiring up to eight diodes. The more complex circuits use symmetry to cancel frequency components that are not desired and to help eliminate noise created by amplitude variations in the LO. The disadvantage of the mixers with more diodes or FETs is the need for more LO power, which is difficult to obtain at higher frequencies. Reliability problems associated with mixers relate to the generation of harmonics that can cause oscillations in other circuits or the package, distortion of the signal created by harmonics, $1/f$ noise, and device burnout.

D. Oscillators

Oscillators generate microwave energy for communications, radars, and navigation systems. For example, modulators, superheterodyne receivers, and phased-locked loops depend on a good microwave source to function. In principal, any amplifier could be made into an oscillator by providing positive feedback to the input terminals so that the reflection coefficient of the amplifier is greater than one. More often than not, this is accidentally done by amplifier designers. Therefore, an oscillator is basically an LNA with a feedback loop that introduces delay-of-integer multiples of 2π. The choice of the load and terminating impedance to achieve this condition should also guarantee the proper oscillation frequency and maximize the efficiency or RF power delivered to a load. In general, there are two types of oscillators: fixed-frequency oscillators designed to operate at a single frequency and variable-frequency oscillators or voltage-controlled oscillators (VCOs) with tuning circuits that change the oscillation frequency. The schematic of a simple oscillator is shown in Figure 3-45. It consists of a transistor with feedback between the gate and drain, an output matching circuit, and a resonant structure on the input. Oscillator performance specifications or figures of merit that affect the system reliability include phase noise and thermal stability.

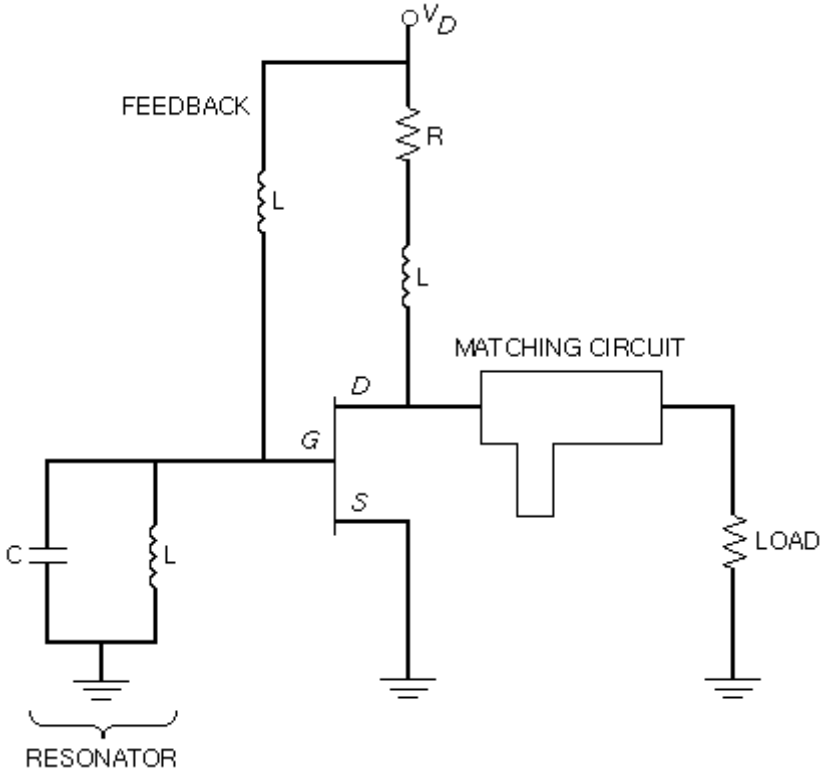

Figure 3-45. Schematic of oscillator.

The phase noise of an oscillator is a measure of the short-term instability of the generated RF signal and is critical in radar applications and digital communication systems where phase noise degrades the system bit error rate (BER). To describe the phase noise, consider a general signal described by

$$V(t) = \left[V_0 + P(t)\right] * \sin\left(\omega_0 t + \phi(t)\right)$$

where $P(t)$ is the amplitude noise term and $f(t)$ is the phase noise term. For $|P(t)| << V_O$ and $|f(t)| << 1$ rad,

$$V(t) \approx V_0 \sin(\omega_0 t) + V_0 \phi(t) \cos(\omega_0 t) + P(t) \sin(\omega_0 t)$$

where the first term represents the desired RF signal and the last two terms represent the amplitude modulated RF signal due to the phase and amplitude noise, respectively. In practice, the phase noise manifests itself as continuous energy sidebands around the carrier in the frequency domain. For the usual case when the phase noise is significantly greater than the amplitude noise, the spectrum around the desired carrier frequency is symmetric.

Noise can be generated by several mechanisms. The first is associated with the kinetic energy of electrons, which is proportional to the temperature of the materials, and thus it is usually called thermal noise. Thermal noise is essentially uniform in magnitude across the entire frequency spectrum, or it is very broad band, which is why it is also referred to as "white" noise, since white light is broad band. The second type of noise is proportional to $1/f$ and is frequently called "flicker" noise because of historical observations of the plate current in vacuum tubes. Flicker noise in active solid-state devices is due to the generation and recombination of carriers at the semiconductor surface [4].

The power spectral density (rad^2/Hz) of phase fluctuations is proportional to the rms phase deviation squared, which results in the spectral slopes of the white and flicker noise becoming twice as steep. It has been shown that the power spectral density decreases at 9 dB/octave where flicker noise dominates, at 6 dB /octave up to the feedback loop half-power bandwidth, and at 0 dB/octave up to the system filter bandwidth, as shown in Figure 3-46 [6]. The figure of merit most often used in oscillators is the ratio of the single sideband noise power per hertz to the carrier signal power at a specific offset frequency.

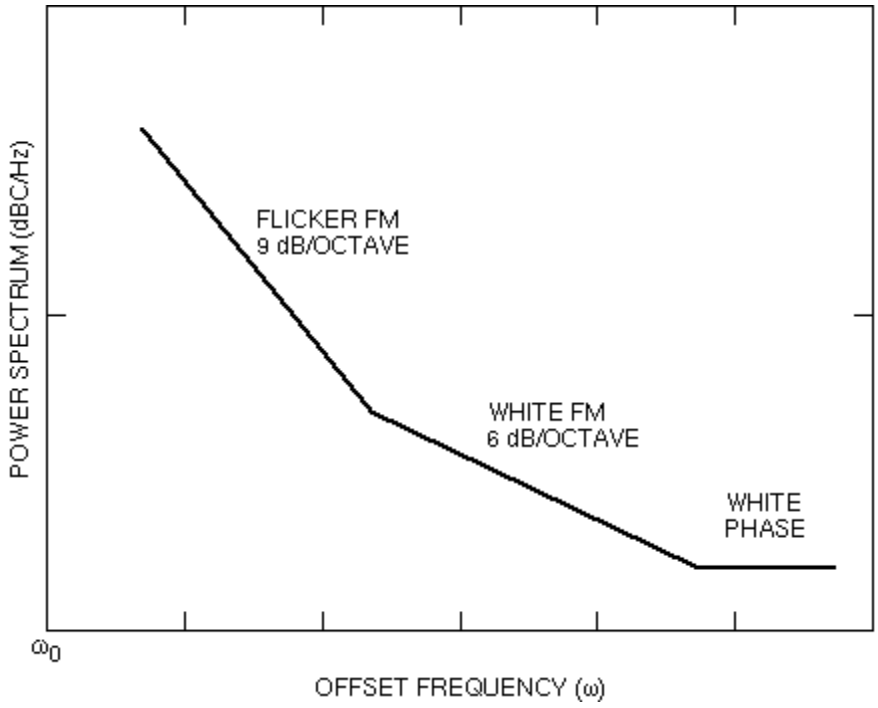

Figure 3-46. Power spectrum for a typical oscillator.

To minimize phase noise, high-*Q* resonators are required to lock in the frequency of the oscillator by providing a reflection coefficient greater than one over a very narrow bandwidth, and transistors with low 1/*f* noise are required. For MMICs, the development of high-*Q* resonators is the more difficult of the two to obtain since thin-film circuit elements on thin GaAs substrates have high conductor loss. HBTs have low 1/*f* noise and are thus frequently used in oscillators. Thermal drift can change the transistor characteristics and cause a shift in the oscillation frequency or cause the circuit to stop oscillating. Temperature compensation can be built in through the use of varactor diodes or other controllable elements with sensors and control circuits.

E. Phase Shifters

Phase shifters are used to impart a repeatable and controllable change of phase to a microwave signal with no effect on the signal's amplitude. Although they are usually associated with phased-array antennas, where they are used to control the beam shape and direction, they are also used in communication systems, radar systems, and microwave instrumentation. Two methods are commonly used to change the phase in MMICs. The first method switches the signal between a short and a long length of transmission line to develop a phase shift of $\beta\ell$ where β is the propagation constant of the transmission line and ℓ is the differential transmission line length. This type of phase shifter is called a switched-line phase shifter and is a true time-delay phase shift. The second method changes the reactance of a transmission line, which changes the propagation constant along the line. The implementation of MMIC phase shifters is broadly characterized as either reflection type or transmission type.

1. Reflection-Type Phase Shifters

Reflection-type phase shifters are one-terminal devices that rely on the reflection of the microwave signal from a termination (e.g., short, open, or other impedance) that has an ideal reflection coefficient with a magnitude of one. An example of a reflection-type phase shifter that employs a switch to add a length of line before the reflective load is shown in Figure 3-47. The resultant transmission-line/termination combination yields a phase shift of $2\beta\ell$ plus a phase that is due to the difference in the termination reactance. Typically, the switches are PIN diodes or MESFETs. Alternatively, an analog phase shifter can be made by removing the switch and replacing the termination load with a varactor diode. Reflection-type phase shifters are primarily used in reflect-array radar applications, or with a coupler to form a transmission-type phase shifter.

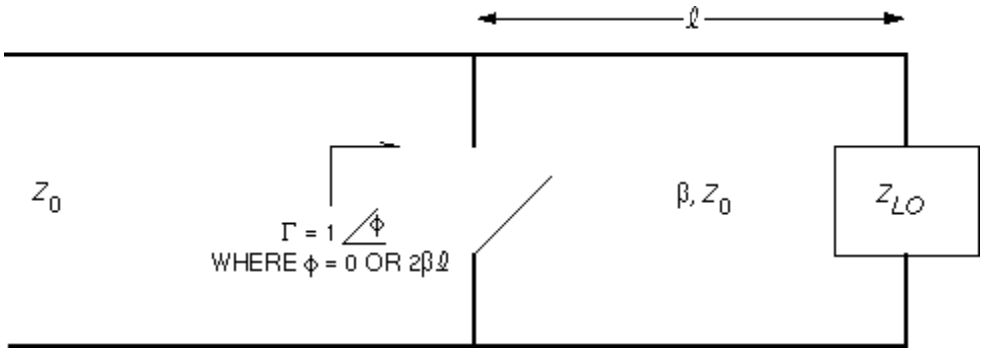

Figure 3-47. Schematic of reflective-type phase shifter.

2. Transmission-Type Phase Shifters

Transmission-type phase shifters are two-terminal devices that change the phase of the input signal as it passes through the circuit. There are three commonly used MMIC implementations of transmission-type phase shifters: hybrid coupled, loaded line, and switched line. The hybrid coupled phase shifters use a reflection-type phase shifter with a coupler to separate the input port from the output port, yielding a two-terminal device. Figure 3-48 is a photograph of an analog hybrid-coupled phase shifter that uses a Lange coupler and varactor diodes.

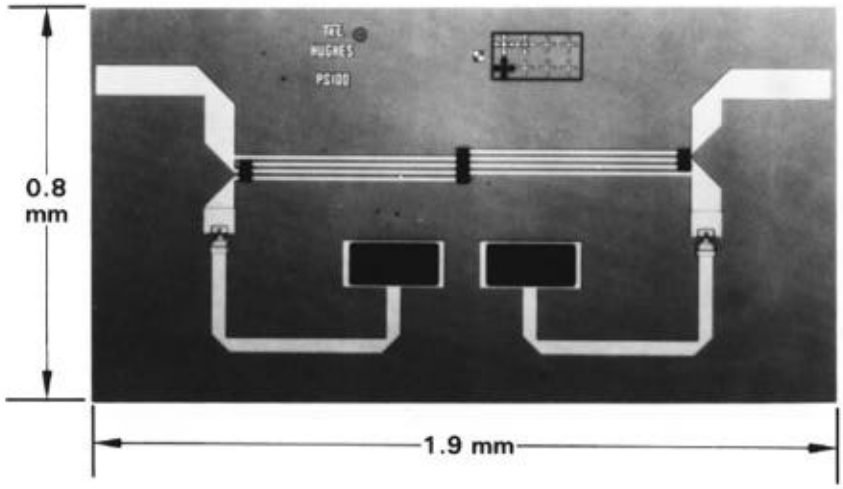

Figure 3-48. Analog phase shifter comprised of a varactor-tuned reflective load and a Lange coupler. (Fabricated by Hughes Aircraft Co. under a contract to NASA Lewis Research Center.)

Figure 3-49 shows a phase shifter comprised of both a switched-line section and a loaded-line section; the schematics of the individual phase-shifter elements are shown in Figure 3-50. The loaded-line phase shifter shown in Figure 3-50(a) is typically comprised of two identical sets of reactive elements separated by about a quarter-wavelength transmission line so that reflections from the reactive elements cancel at the input terminal of the phase shifter. Phase shift is generated by changing the loading on the transmission line and therefore changing its propagation constant, which is approximated by $\beta = \sqrt{LC}$. This type of phase shifter is used for phase shifts less than about 45 deg. The switched-line section is the most straightforward of all. It offers a true time-delay phase shift by switching between two different lengths of transmission line.

Phase shifters are not usually high-power circuits, and therefore the reliability concerns associated with high-power circuits do not need to be considered. Furthermore, the phase shift created by the switch-line type of phase shifters is dependent on transmission line lengths only, and they are therefore very stable over time and temperature. Parametric drift of the active components in the analog type of phase shifters normally translates directly into a phase-state degradation. The active devices

used for switching elements (PIN diodes and MESFETs) may also suffer from parametric drift, but this usually manifests itself as a degradation in the insertion loss of the circuit

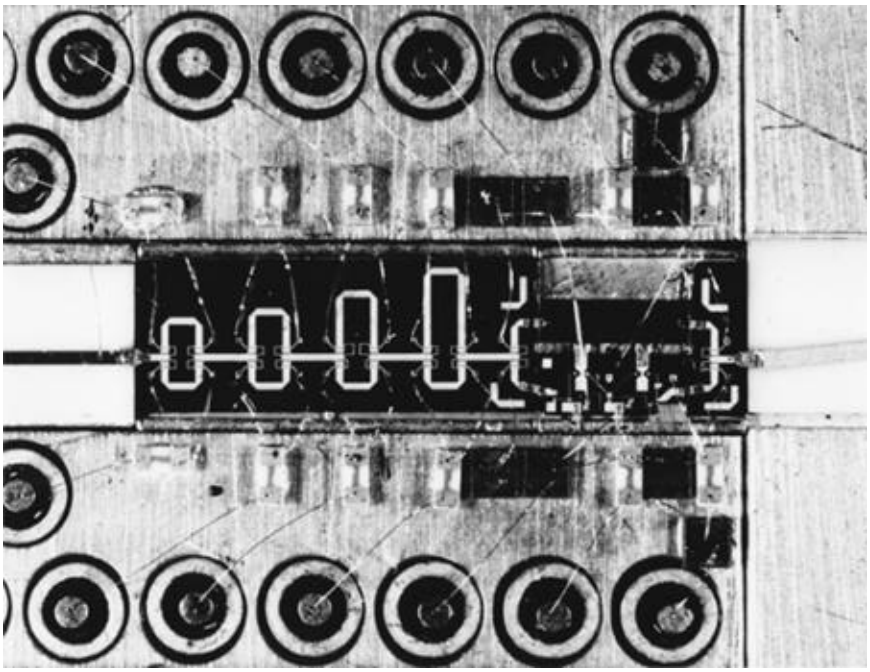

Figure 3-49. Phase shifter comprised of loaded-line and switched-line sections. (Fabricated by Honeywell under a contract to NASA Lewis Research Center.)

and not the phase shift. Finally, burnout of the switches or varactors must be avoided by proper device and system design.

References

[1] I. Bahl and P. Bhartia, *Microwave Solid State Circuit Design*, John Wiley & Sons, New York, 1988.

[2] T. T. Ha, *Solid-State Microwave Amplifier Design*, John Wiley & Sons, New York, 1981.

[3] *Monolithic Microwave Integrated Circuits: Technology & Design*, R. Goyal, Editor, Artech House, Norwood, MA, 1989.

[4] S. M. Sze, *Physics of Semiconductor Devices*, John Wiley & Sons, New York, 1981.

[5] C. M. Howell and S. J. Parisi, *Principles, Applications and Selection of Receiving Diodes*, MA-COM Application Note AG314.

[6] D. B. Leeson, "A Simple Model of Feedback Oscillator Noise Spectrum," *IEEE Proceedings Letters*, Feb. 1966, pp. 329–330.

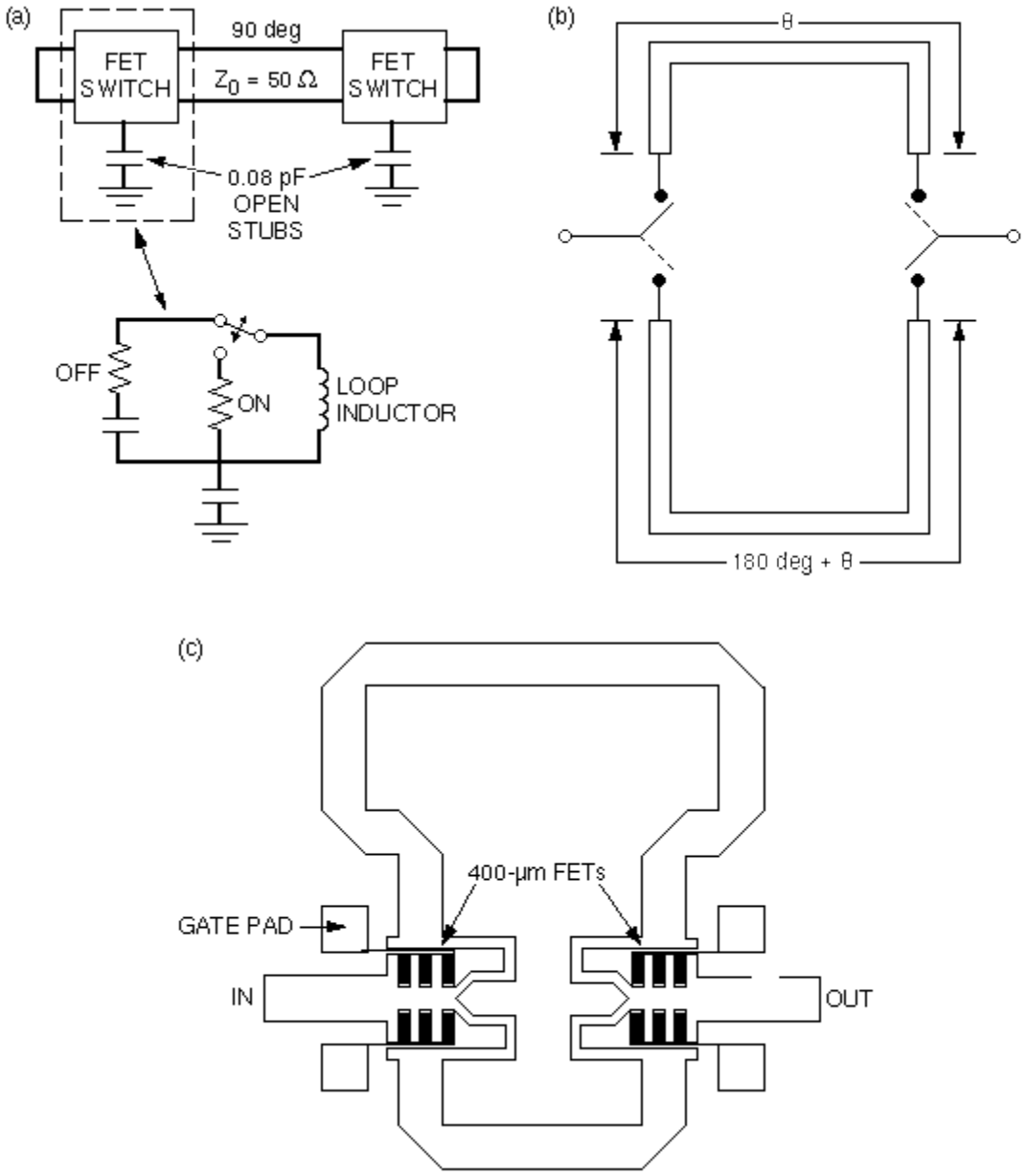

Figure 3-50. 22.5-deg phase-shifter elements: (a) schematic of circuit that is implemented by a loaded-line section; (b) schematic of 180-deg bit with switched lines; (c) layout of 180-deg switched-line bit using four series FET switches.

Chapter 4. Basic Failure Modes and Mechanisms

S. Kayali

Failures of electronic devices, in general, can be catastrophic or noncatastrophic. Catastrophic failures render the device totally nonfunctional, while noncatastrophic failures result in an electrically operating device that shows parametric degradation and limited performance.

This chapter provides a description of some of the more common failure modes and mechanisms affecting GaAs-based MMICs. The current understanding of the topic will be presented along with a discussion of some possible solutions, practiced process improvements, and references.

I. General Failure Modes

GaAs devices exhibit some general failure modes that can be attributed to a defined failure mechanism. The most common failure modes are observed via degradation of the MMIC parameters such as I_{DSS}, gain, P_{OUT}, and others. The degradation observed in MMIC devices is normally a function of the material interactions and the environmental conditions during test or operation. The importance of a particular parameter degradation depends greatly on the design and function of the MMIC and the relationship between the observed degradation and the general health of the device in question. A list of the most common failure modes is provided in Table 4-1. Life tests, with RF or dc excitation and performed under controlled conditions, are the most common means of failure-mode detection. These tests can provide valuable information as to the type of degradation to which the particular device under test may be most susceptible, and the severity of the effect on the performance of the device.

Table 4-1. Common MMIC failure modes.

Failure Mode	Method of Detection	Related Failure Mechanisms	Possible Solutions
Degradation in I_{DSS}	Life test, operation	Gate sinking, surface effects, hydrogen effects	Derating criteria, temperature control, environmental control
Degradation in gate leakage current	Life test, high-temperature storage test, high-temperature reverse bias	Interdiffusion	Temperature control, gate current control, proper passivation
Degradation in V_P	Life test, operation	Gate sinking, hydrogen effects	Temperature control, use of stable barrier materials, environmental control
Increase in R_{DS}	Life test, operation	Gate sinking, ohmic contact degradation	Temperature control, use of stable barrier materials
Decrease in P_{OUT}	Life test, operation	Surface effects, hydrogen effects, gate sinking	Temperature control, use of stable barrier materials, environmental control

While dc testing is much easier and more cost effective to implement, RF testing has the advantage of providing the user with direct information in regard to device degradation under conditions similar to those of the actual application. Correlation between the results of dc tests and actual RF application has been a topic of great interest and debate in the GaAs reliability community, but to date there is little understanding or agreement of the relationship.

A. Degradation in I_{DSS}

This failure mode is one of the most common and easiest to detect. Accelerated life tests have been used to provide an estimation of the lifetime of devices based on the observed level of I_{DSS} degradation. Various failure mechanisms can be attributed to be the cause of this observed degradation. One of the most common is referred to as "gate sinking." In this mechanism, a reduction in the active channel of the device results in a decrease of I_{DSS} among other parameters. Another common mechanism, which can cause similar degradation, is referred to as "hydrogen poisoning." This mechanism is theorized to cause a decrease in the donor density in the channel, which in turn causes a reduction in I_{DSS} [1]. A detailed discussion of failure mechanisms will be presented in Section II.

B. Degradation in Gate Leakage Current

This failure mode is generally observed in devices subjected to an accelerated life test or to high operating temperatures. The degradation is observed as an increase in the gate leakage current over the duration of the test. No experimentally identified failure mechanisms have been linked to this failure mode, but surface-state effects have been suspect.

C. Degradation in Pinch-Off Voltage

Pinch-off voltage (V_p) degradation is another common failure mode for GaAs devices. This degradation results primarily from metal–semiconductor interactions and instability of gate-metal structures. The degradation is normally observed on devices subjected to accelerated life tests or high-temperature operation. Reliability related effects of metal–semiconductor interactions may render the associated barrier layers ineffective due to poor manufacturing practices or material choices. The choice of the appropriate barrier material to limit Au/GaAs interdiffusion is the best method to limit the effects of this degradation.

Hydrogen-related degradation may cause the same observed pinch-off voltage degradation effects [1]. This degradation is theorized to be caused by either a reduction of carrier concentration in the active channel of the device or a change in the surface state built-in potential. Further information on this degradation is found in Section II.D.2.

D. Increase in Drain-to-Source Resistance

The increase in the drain to source resistance (R_{DS}) can be attributed to either gate sinking or to ohmic contact degradation. Both of these failure mechanisms are metal–semiconductor related degradation mechanisms that are accelerated with temperature.

Therefore, devices subjected to accelerated life tests or operation at elevated temperatures generally exhibit this degradation.

E. Degradation in RF Performance

Various RF parameters can exhibit degradation over the lifetime and operation of the devices. Although it is very difficult to extrapolate RF performance from dc test data, some manufacturers use particular dc parameters as predictors of resultant RF performance. I_{DSS}, for example, can be used as a predictor of saturated power performance, while g_m can be used for prediction of gain and noise figure degradation in small-signal and low-noise devices.

The causes of RF parameter degradation vary, depending on the technology and operating conditions of the devices under test. In general, surface-state density and resultant surface effects play a role in overall RF device performance and stability over time. Material interaction effects also play a major role in long-term device performance. Other factors, such as hydrogen-related degradation and other environmental effects, can also contribute to the overall degradation.

II. Failure Mechanisms

Failure mechanisms of electronic semiconductor devices can be divided into the following general categories:

(1) Material-interaction-induced mechanisms.
(2) Stress-induced mechanisms.
(3) Mechanically induced failure mechanisms.
(4) Environmentally induced failure mechanisms.

Material-induced mechanisms can in turn be subdivided into two general categories, the first being semiconductor die material and metal interactions, and the second being a result of die packaging and interconnect. Stress-induced failure mechanisms can be directly attributed to either poor device design or poor and careless device application. Environmentally induced failure mechanisms can cover a wide spectrum of possible environmental conditions, such as humidity and hydrogen effects.

Reported device-failure mechanisms can be a result of one or a combination of these factors. Therefore, care must be exercised in understanding the operating and environmental conditions and process variables associated with the reported failure. Table 4-2 shows the main areas of responsibility for the failure-mechanism categories. In this chapter, a discussion of the general categories of failure mechanisms will be provided, along with reference examples as applicable.

A. Material-Interaction-Induced Failure Mechanisms

GaAs processes involve a number of metal–semiconductor interfaces which, if not designed and applied properly, may cause device degradation and failure. The two main metal–semiconductor interfaces in GaAs-based devices are the Schottky gate contact and the ohmic source and drain contacts. The common metallization structures for GaAs are based on the industry standard Au/Pt/Ti or Au/Pd/Ti on GaAs. The thermal

Table 4-2. General responsibilities for the failure-mechanism categories.

Failure Mechanism Category	Manufacturer Control	User Control
Material-interaction induced	√	
Stress induced	√	√
Mechanically induced		√
Environmentally induced	√	√

stability and reproducibility of Schottky barriers, the correct choice of metals and their applicable processing parameters, and the GaAs surface conditions all play a role in the reliability of the produced structures and the applicable failure mechanisms. Failures related to Schottky and ohmic contacts occur when the metals diffuse into the semiconductor, and the Ga and/or As diffuse into the contact. A description of the failure mechanisms related to these interfaces will be provided along with relevant examples.

1. Gate-Metal Sinking

The performance of GaAs-based devices relies heavily on the quality of the active channel area of the device. The Schottky gate metal-to-semiconductor interface directly influences the device electrical parameters, such as the drain saturation current and reverse breakdown. The gate structures are based on the industry standard mutilayer Au/Pt/Ti or Au/Pd/Ti on GaAs. Interdiffusion of gate metal with GaAs results in a reduction of the active channel depth and a change in the effective channel doping. This effect is termed as “gate sinking.” This process is affected by the surface conditions of the GaAs material at the time of deposition, the deposition parameters, and the choice of deposited materials.

This failure mechanism is generally observed after exposure to an accelerated life test or operation at elevated temperatures, the driving factor for this mechanism being the thermally accelerated diffusion of Au into GaAs. The common gate metallization structure consists of three layers. The first layer contacting GaAs is a thin Ti layer used primarily for adhesion. The second layer is either Pd or Pt. This layer is used as a barrier to Au diffusion into GaAs. The last layer is thick Au used for conduction. The rate of Au gate-metal diffusion into the GaAs is a function of the gate-metal material diffusivity, the temperature , and the material-concentration gradient. For perfect lattice structures, the diffusion rate at normal operating temperatures is too slow to have an effect on device performance. However, when large grain boundaries or large numbers of surface defects exist, the diffusion rate can be fast [2].

Au has a high diffusion factor into GaAs, therefore a Pt or Pd layer is employed to act as a barrier to Au diffusion into GaAs. Grain boundaries in the barrier layer may allow a diffusion path for Au, which in turn may cause device degradation. The inclusion of some oxygen or nitrogen in the barrier films helps reduce the grain boundary diffusion of Au through the films. Other poor manufacturing processes or material quality may render the barrier layer useless. Several examples of Au interaction with GaAs through different barrier layers have been reported in the literature [3,4,5].

2. Ohmic Contact Degradation

The most common system for ohmic contacts is AuGe/Ni, which is alloyed into the GaAs at temperatures in excess of 400°C to provide the necessary low contact resistance (0.1 to 0.5 Ω/mm). A thick Au layer is then deposited on top of the alloyed contacts to provide conduction. This structure, employed at the drain and source contacts, has been shown to degrade at elevated temperatures. The degradation is the result of Ga outdiffusion into the top Au layer and the diffusion of Au into the GaAs causing an increase in the contact resistance [6]. The Ni layer used in the ohmic contact is intended as a Au- and Ga-diffusion barrier. Some other materials such as Cr, Ag, Pt, Ta, and Ti have been used as barrier materials with varying degrees of success. The activation energy associated with ohmic contact degradation varies between 0.5 eV and 1.8 eV [7,8]. This activation energy may provide reasonable contact life at low operating temperatures (< 100°C) but it also indicates rapid deterioration at elevated temperatures (>150°C) [9].

The general understanding of ohmic contacts attributes the degradation to the following :

(1) Ga outdiffusion into the Au layer, which creates a nonstochiometric defect-rich region of high resistivity under the contact. This effect is reduced by employing a barrier layer sandwiched between the AuGe and the Au conduction layer [10].

(2) Indiffusion of Au and Ni into the GaAs, which can cause a reduction in the doping concentration in the active channel of the device [6,7].

(3) The formation of various intermetallic phases such as AuGa and Ni_2AsGe as a result of the alloying process.

Sputter cleaning of the surface prior to deposition along with deposition of Ni as a first layer can provide for much improved ohmic contact stability and homogeneity [11]. Continuous or noncontinuous contamination at the deposition surface by oxides or other contaminants can result in regions of high resistance. The NiAs(Ge) phase is essential for low contact-resistance values, because it satisfies the condition that the Ge atoms diffuse into the Ga vacancy sites forming a heavily doped n^+ layer at the metal/GaAs interface.

Results of recent accelerated life tests confirm that the stability of AuGeNi alloyed ohmic contacts does not appear to be a major reliability concern under normal operating conditions. However, as the gate and gate-drain/source dimensions of high speed devices shrink, vertical spiking and lateral spreading during the alloying process will not allow good dimensional control of alloyed contacts such as AuGeNi [12]. To overcome this limitation, new ohmic contacts have been developed utilizing low temperature anneal of an epitaxially grown thin layer of Ge on GaAs. Another approach is to grow an epitaxial n^+ Ge layer on n-GaAs followed by deposition of a refractory metal layer. Deposition of small amounts of In along with W contact metal and annealing using rapid thermal anneal techniques have also been used.

3. Channel Degradation

Degradation observed in device parameters can sometimes be attributed to changes in the quality and purity of the active channel area and a reduction in the carrier concentration beneath the gate Schottky contact area. These changes have been

postulated to be a result of diffusion of dopants out of the channel or diffusion of impurities or defects from the substrate to the channel [13]. Deep level traps have also been speculated to cause similar degradation in MESFETs.

HEMT devices, being strongly dependent on the properties of the interface of the AlGaAs/GaAs heterostructure, can suffer a related failure mechanism. A decrease in electron concentration in the channel, caused by a deconfinement of the 2DEG, was postulated to be the cause of the observed failure mechanism [14].

HEMT devices can also suffer from metal-diffusion-related mechanisms, which are manifested as channel-related degradation. Lateral diffusion of Al into the gate recess region changes the conduction band discontinuity and consequently the confinement of the channel electrons. Gold diffusion from the ohmic contact into the active channel region under the gate can also cause similar degradation. Lastly, vertical diffusion of Al from the AlGaAs donor layer and Si from the n^+ AlGaAs layer into the channel layer causes an increase in the impurity scattering in the undoped GaAs, thus deteriorating the high electron mobility of the 2DEG [15].

4. Surface-State Effects

The performance of GaAs-based devices depends highly on the quality of the interface between metal and GaAs or the passivation layer (Si_3N_4 or SiO_2) and GaAs. The quality of the interface can depend on the surface cleaning materials and procedures, the deposition method and conditions, and the composition of the passivation layer. The main effect of an increase in surface state density, as illustrated in Figure 4-1, is the

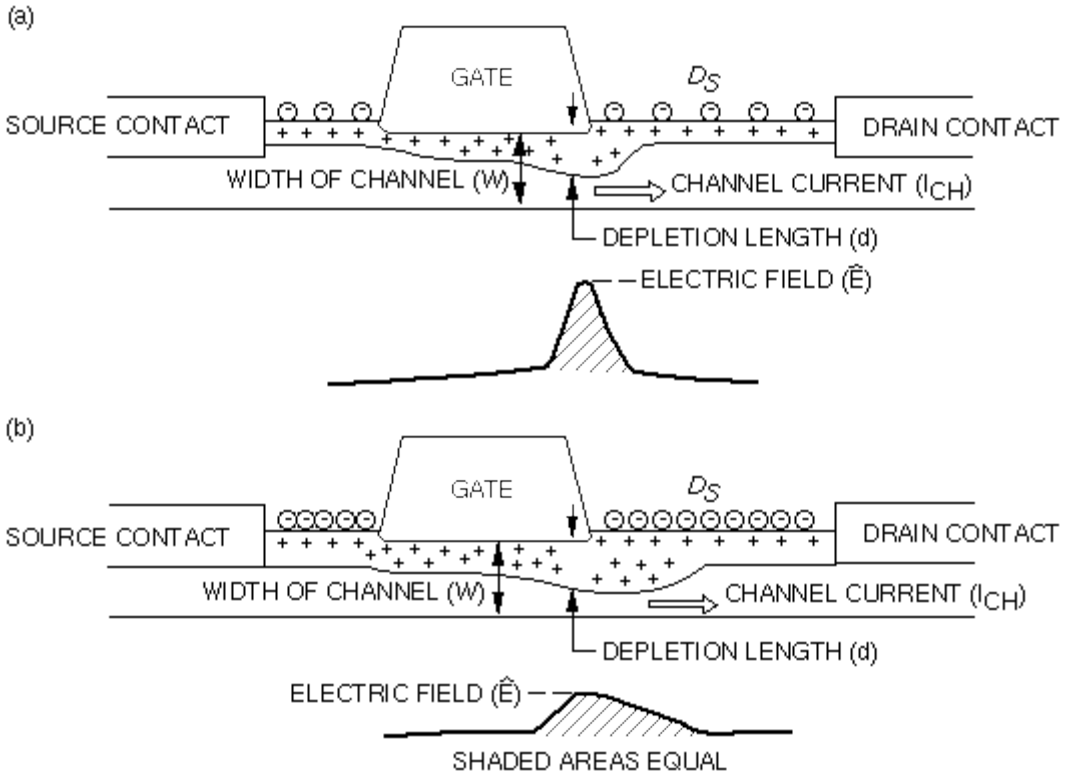

Figure 4-1. Schematic cross section of a MESFET with different surface charges. The gate-drain bias is the same for the two cases: (a) with low density of surface states D_s and (b) with high density of D_s. (From [12]; reprinted by permission of John Wiley & Sons, Ltd.)

lowering of the effective electric field at the drain/gate region, which results in an increase in the depletion region and a change in the breakdown voltage [16].

Unpassivated devices can be susceptible to surface oxidation and loss of arsenic, which may result in an increase in gate leakage current and a reduction of the breakdown voltage. Devices passivated using SiO_2 may experience surface erosion due to the interaction of SiO_2 with GaAs. The use of Si_3N_4 provides a much improved passivation layer with no GaAs surface erosion and a reduced level of arsenic loss. Plasma deposited Si_3N_4 also provides lower tensile stress compared to CVD SiO_2 passivation layers and therefore a reduced effect on surface states.

Surface-state density has a direct effect on the performance of GaAs-based devices. The reduction in the surface-state density at the Si_3N_4/GaAs interface caused by thermal treatments may result in degradation in breakdown voltage, which in turn may give rise to device burnout.

B. Stress-Induced Failure Mechanisms

1. Electromigration

Electromigration is the movement of metal atoms along a metallic strip due to momentum exchange with electrons. Since the mechanism is dependent on momentum transfer from electrons, electromigration is dependent on the temperature and the number of electrons. Therefore, this failure mechanism is generally seen in narrow gates and in power devices where the current density is greater than 2×10^5 A/cm^2, which is normally used as the threshold current density for electromigration to occur. This effect is observed both perpendicular to and along the source and drain contact edges and also at the interconnect of multilevel metallizations.

The metal atoms that migrate along the line tend to accumulate at grain boundaries. The accumulation of metal at the end of the gate or drain contact can create fingers of metal that can short the device. Figure 4-2 shows an example of material accumulation at one end and depletion (voids) at the other end of a drain contact. At the void location, the current density increases due to current crowding, which further increases the temperature due to resistive heating. These effects increase the rate of electromigration, which further increases the void size. Therefore, void creation is a self-accelerated, runaway process. If the void formation occurs in the gate of the device, electromigration may result in catastrophic failure due to the creation of gate open circuits. If electromigration occurs in the drain/source of a device, the voids may result in increased drain/source contact resistance and associated device degradation.

Material accumulation and void formation perpendicular to the source and drain contacts can cause hillock formation over the gate structure. This may result in shorting the gate to source or drain which may result in a catastrophic failure. Figure 4-3 shows an example of void and hillock formation perpendicular to the source and drain contacts. Electromigration problems at the interconnect have also been reported to occur at the AuGeNi interface [17].

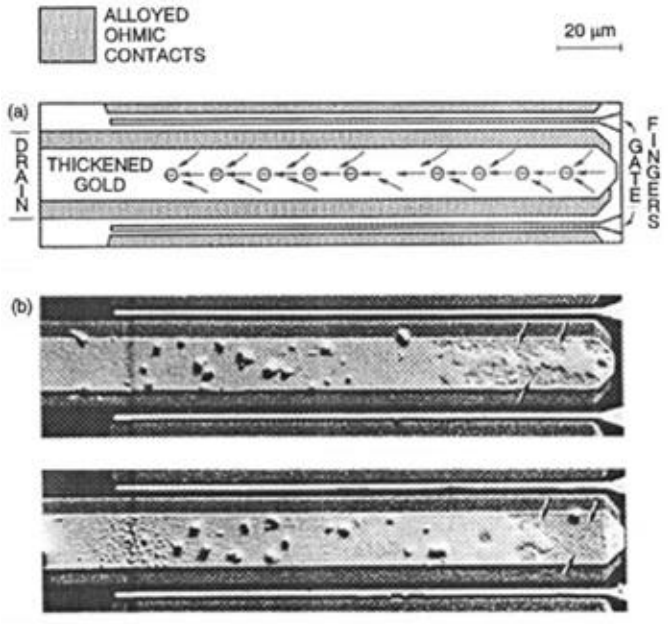

Figure 4-2. Metal-atom migration and accumulation: (a) electron wind along drain fingers; SEM images of (b) accumulation and depletion of a drain contact on a device that endured 5000 h of life testing at T_{ch} = 200°C and j = 5.3 × 105 A/cm^2. (From [12]; reprinted by permission of John Wiley & Sons, Ltd.)

It should be emphasized that electromigration failures can be avoided by limiting the current density and providing a controlled temperature of the devices during operation. Process control assuring a clean and defect-free interface structure is also essential.

2. Burnout

The partial or complete melting of a large device area resulting in catastrophic failure is referred to as burnout. This failure mechanism is considered to be the final result of a combination of other failure mechanisms causing an increase in localized power dissipation.

Burnout can be divided into two forms, “instantaneous” and “long-term.” Instantaneous burnout is caused by sudden events such as electrostatic discharge (ESD), electrical overstress (EOS), and RF spikes. This failure mechanism is related more to device design and robustness than material interaction in the conventional sense of reliability.

Gate-drain burnout can be attributed to avalanche breakdown and therefore depends to a large extent on surface characteristics and device layout and technology.

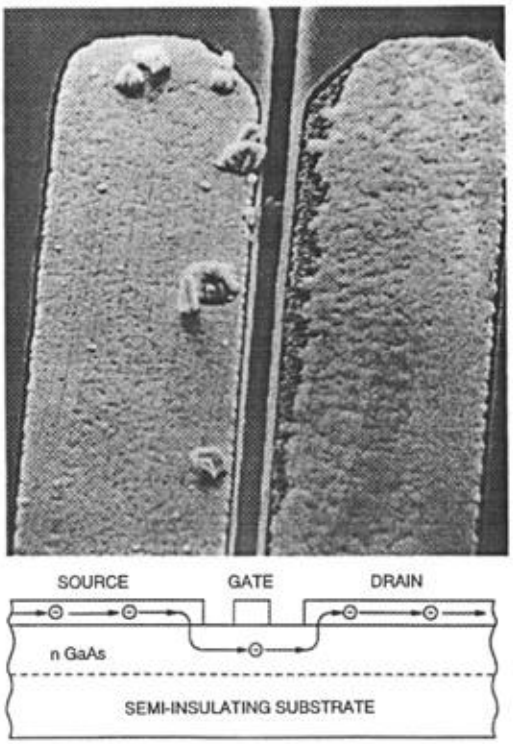

Figure 4-3. Depletion and accumulation of material in AuGeIn source and drain ohmic contacts induced by electromigration in a low-noise MESFET after life test. (From [12]; reprinted by permission of John Wiley & Sons, Ltd.)

To improve the breakdown voltage and burnout characteristics of MESFET type devices, a recessed gate design is usually implemented. Consideration of layout topology, use of an offset gate in relation to the source and drain, and careful characterization of the resultant device under various operating conditions can greatly reduce the occurrence of gate-drain burnout.

Source-drain burnout has been found to be thermally activated and has been shown to initiate at the drain contact where nonuniformities and current crowding cause local hot spots. These hot spots in turn cause a thermal runaway condition associated with the temperature coefficient of the buffer or substrate material [17]. This action has been shown to occur when the buffer and substrate materials reach local temperatures higher than 550°C, leading to a sudden increase in buffer and substrate conductivity and, consequently, in drain current.

Thermally induced metal–GaAs interdiffusion can cause a very similar failure scenario. Metal migration through the grain boundaries and crystalline defects of GaAs can reach the substrate/active channel interface causing a short between the gate or source and drain. Localized heating generated in these locations along with the positive temperature coefficient of the substrate can initiate a positive feedback mechanism in an area of high current density. Thermal runaway is consequently initiated, leading to burnout and catastrophic failure.

Long-term burnout, on the other hand, is believed to be the final result of a parametric degradation occurring during long-term aging and leading to an increased and localized power density dissipation. One of the factors that may contribute to this condition is surface effects such as oxidation reduction of GaAs and the annealing of surface states, which may cause an increase in the leakage current and reduce the breakdown voltage [17]. For example, test results show that a significant improvement is possible when silicon nitride is used instead of silicon dioxide as a passivation layer. This has been attributed to the lower tensile stress and a reduction of the effect on surface states of the plasma-deposited Si_3N_4 compared to those of CVD SiO_2.

Other factors contributing to the long-term burnout include interelectrode bridges and lateral surface metal migration causing an electrical short and, in turn, a burnout condition. Metal–semiconductor interactions resulting in vertical spikes can cause localized heating and thermal runaway conditions and, in turn, burnout.

3. Hot Electron Trapping (modified from [20,21,22])

When RF power transistors are driven into heavy gain compression in order to achieve the maximum power or efficiency, they often suffer the so-called “power-slump” problem, which shows up typically as an approximate 1-dB drop in output power over 1,000 h of RF operation. Initially, the power-slump problem was thought to be unique to GaAs devices—a problem related to metallurgical diffusion and electromigration. However, within the last few years, a hot-electron-induced gradual degradation mechanism in MESFETs was uncovered [18,19,20]. Such a degradation mechanism has been known to take place in Si MOSFET devices: Hot electrons can be trapped in the gate oxide, causing the MOSFET threshold voltage to shift. Few investigators suspected that in a GaAs MESFET, hot electrons can also be trapped in the Si_3N_4 passivation between the gate and drain, thereby decreasing the MESFET’s transconductance without affecting its threshold voltage [21].

Under RF overdrive, hot electrons are generated near the drain end of the channel where the electrical field is the highest. A few electrons can accumulate sufficient energy to tunnel into the Si_3N_4 passivation to form permanent traps. These traps can result in lower open-channel drain current and transconductance, and higher knee voltage, leakage current, and breakdown voltage. Since the traps are located above the channel (see Figure 4-4), there is usually little change in the dc or small signal parameters near the quiescent point. Further, since the traps are located beside the channel, Schottky-barrier height and the ideality factor often remain constant. This selective change in device characteristics helps distinguish hot-electron effects from thermal or environmental effects. In fact, the most distinct feature of hot-electron effects is a weak or negative temperature dependence. This is because, when the channel is hotter, electrons undergo more scattering and, therefore, are less energetic.

Based on the current understanding of the degradation mechanism, work is now being concentrated on improving the device design to reduce the degradation tendency.

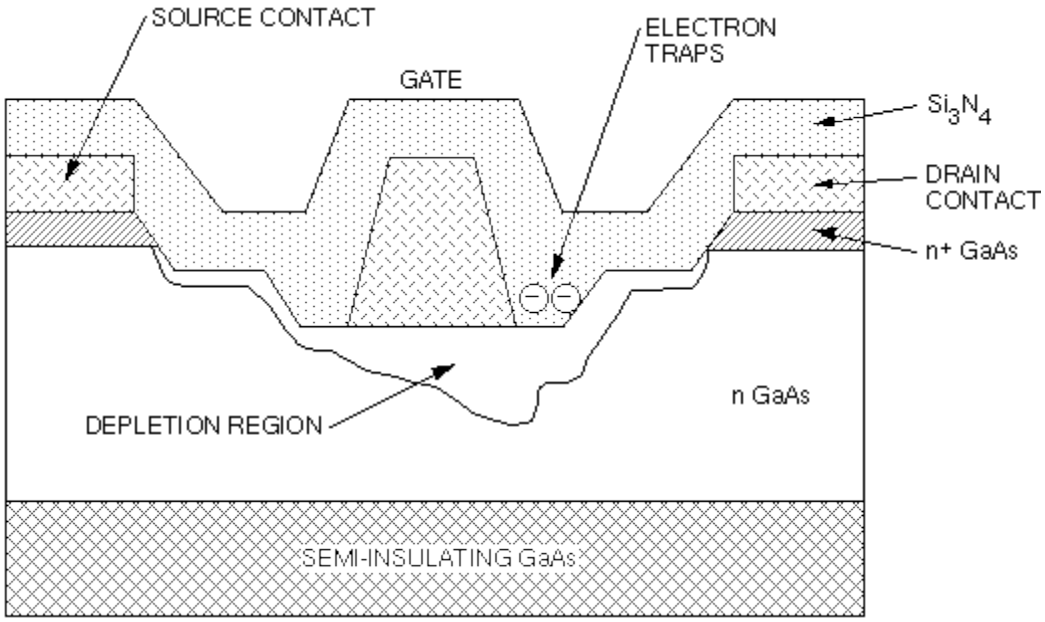

Figure 4-4. Schematic cross section of a degraded MESFET. Hot-electron-induced traps are formed in the SiN passivation layer between the gate and the drain.

This is typically done by trial and error and may take several iterations of wafer processing and device characterization. However, with improved device modeling capabilities and the use of novel measurement techniques, such as high-voltage electron-beam-induced current [22], it is now possible to optimize the shape of the electrical rather than the physical channel without many iterations. Improvement of the Si_3N_4 as a surface passivation is another obvious approach to limiting the described effects. However, perfect passivation of the GaAs surface is yet to be found. Other approaches, such as limiting the operating voltage and including a low-doped drain region as is common in a MOSFET, are either impractical or may actually contribute to further degradation in performance.

4. Electrical Stress

Electrical stress of devices during operation or handling can result in device degradation or catastrophic failure. Electrical overstress (EOS) can result from the improper application or use of the device and may result in parametric degradation or eventual catastrophic failure. Electrostatic discharge (ESD), on the other hand, can result from improper handling and lack of adequate ESD protection during transfer or test of exposed devices. The very small geometry of GaAs devices along with the semi-insulating nature of the material further enhances the sensitivity of the devices to electrostatic discharge effects.

The discharge of large electrical pulses can cause damage to both the gate and ohmic metallization structures, resulting in local melting and pursuant parameter degradation or catastrophic failure. Studies of devices exposed to noncatastrophic ESD levels have observed that MESFETs exhibited an increase in low leakage current and further catastrophic failure at RF power levels below those with no prior exposure to noncatastrophic ESD levels [23]. Other studies have also concluded that damage from

repeated exposure to an ESD level is not cumulative and that noncatastrophic damage does not degrade device lifetime [24].

Other studies have concluded that devices using the AuGeNi ohmic contact structure can exhibit ESD-related interdiffusion of the Au-based ohmic contact with GaAs. The high current densities caused by ESD can result in localized heating at the metal–semiconductor interface leading to Ga diffusion into the metallization and Au diffusion into GaAs [25]. Schottky contacts have also been found to exhibit rapid degradation under ESD stress [26]. The effect is accelerated by the small geometry and the low cross-sectional area of the gate metallization, resulting in very high current densities in response to an ESD stress event. An explosion of the gate metallization can result in response to the heat generated by the high current density; this can result in the gate metallization being physically blown out of the gate recess as shown in Figure 4-5.

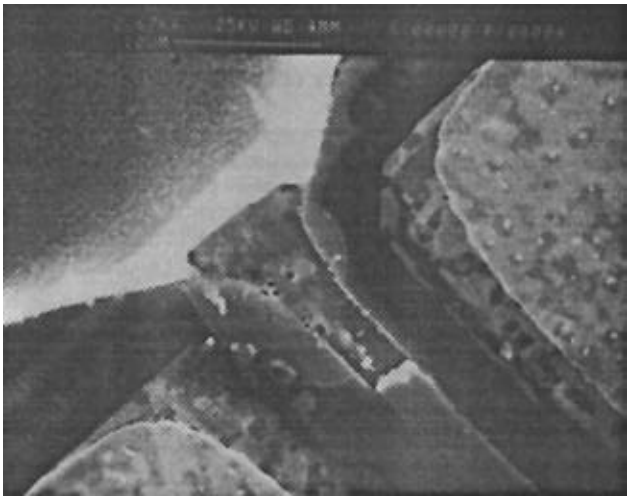

Figure 4-5. Blown-out gate recess. (From [12]; reprinted by permission of John Wiley & Sons, Ltd.)

Passive MMIC elements—such as capacitors, resistors, and interconnect metallization—can also exhibit the detrimental effects of ESD. Gold-based interconnect metallization 2 μm thick has shown limited susceptibility to ESD pulses [24]. Thin-film nickel–chromium resistors, on the other hand, have shown a strong susceptibility to ESD effects as shown in Figure 4-6. The amount of ESD pulse voltage required to cause resistor damage was observed to depend on the width and thickness of the structure [24].

Metal–insulator–metal (MIM) capacitors show a strong susceptibility to ESD damage. Failures in MIM capacitors tend to occur at either the edges of the structure (Figure 4-7) where the electric field is the highest, or at the interior of the capacitor (Figure 4-8). Failures occurring at the interior of the capacitor can be attributed to dielectric defects of surface-related anomalies, while failures occurring at the edges indicate that the ESD performance is limited by the strength of the dielectric material (Si_3N_4).

Increased awareness of the effects of ESD on device reliability and the implementation of ESD precautions and controls—at all facets of device fabrication and

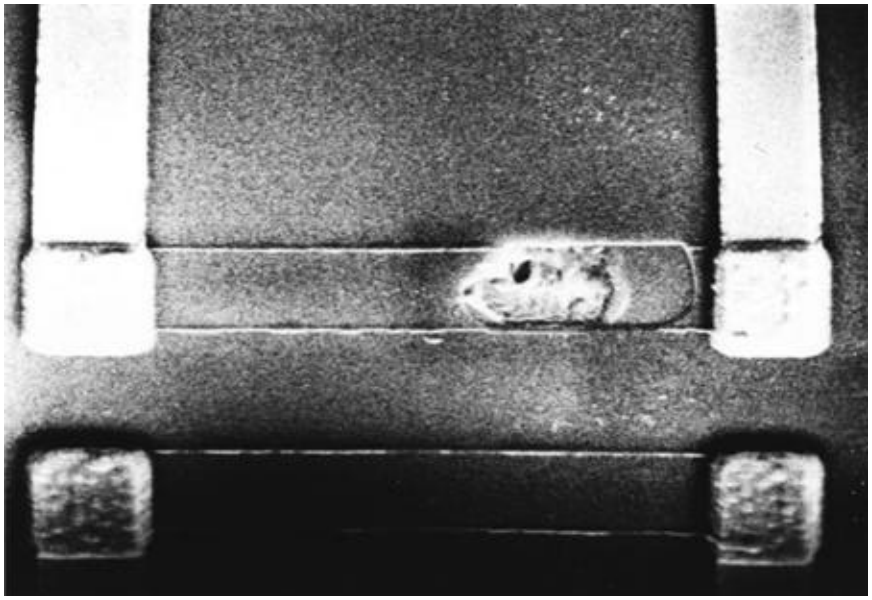

Figure 4-6. SEM photograph of a failed nickel–chromium resistor. (Courtesy of Tri Quint Semiconductor.)

test—can help eliminate this as a device-failure mechanism. If practical, ESD protection circuitry can also be implemented.

C. Mechanically Induced Failure Mechanisms

1. Die Fracture

The difference in the coefficient of thermal expansion (CTE) of GaAs, the carrier or substrate, and the package material can cause mechanical stresses in the die that may result in device failure. Tensile stresses can develop in the central region of the die, while shear stresses can develop at the edges of the die [27]. Thermal cycling either during test or operation may cause surface cracks, which are present at the center or the edges of the die, to reach their critical size and propagate across the surface, resulting in die fracture. Surface cracks can also result from an improper dicing operation, or from an improper die mounting technique.

Die surface cracks and fractures at or close to an active region of the device may result in threshold voltage shifts and general device performance degradation. An increase in leakage current at that location may result in a thermal runaway condition and ultimately catastrophic device failure.

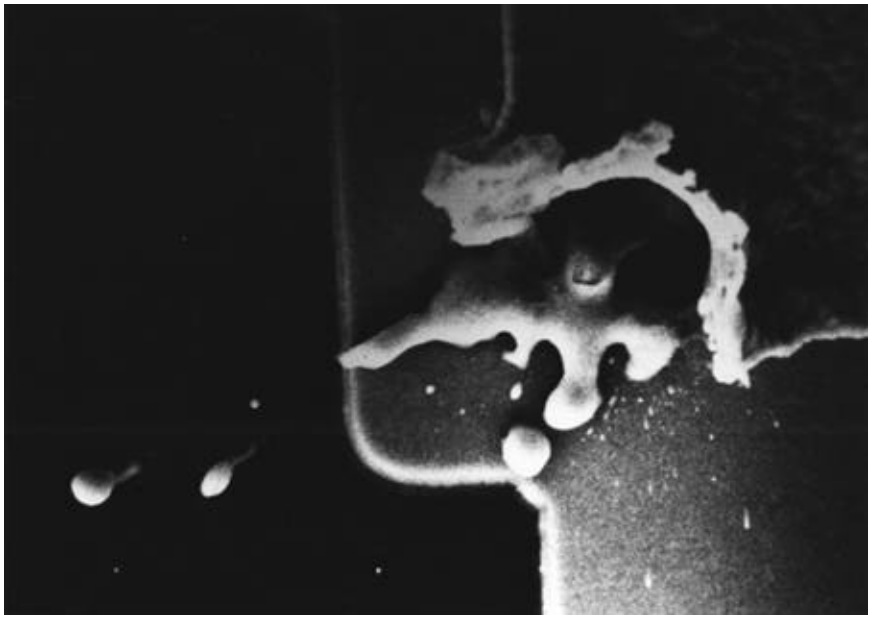

Figure 4-7. Edge-located ESD failure of a MIM capacitor. (Courtesy of Tri Quint Semiconductor.)

2. Die-Attach Voids

Due to the relatively low thermal conductivity of GaAs, die-attach quality and uniformity across the attach surface are essential for proper device operation and reliability. Voids in the die-attach material are one of the most common causes of semiconductor-device thermal runaway and failure. The presence of voids at the edges of the die can induce high longitudinal stresses during power and environmental temperature cycling. Propagation of these voids may result in die delamination and interruption of the thermal path. Physical die detachment from the package or substrate is seldom observed as a result of void propagation.

Although voids can form from a number of sources, process control can limit the effects to an acceptable level. The package or substrate construction, the die-attach material physical properties, the cleaning and application methods, and the overall void concentration and location determine the effect of voids on device reliability [27].

D. Environmentally Induced Failure Mechanisms

1. Humidity Effects

GaAs devices packaged in nonhermetically sealed packages or plastic encapsulated packages suffer from a number of humidity-related or accelerated failure

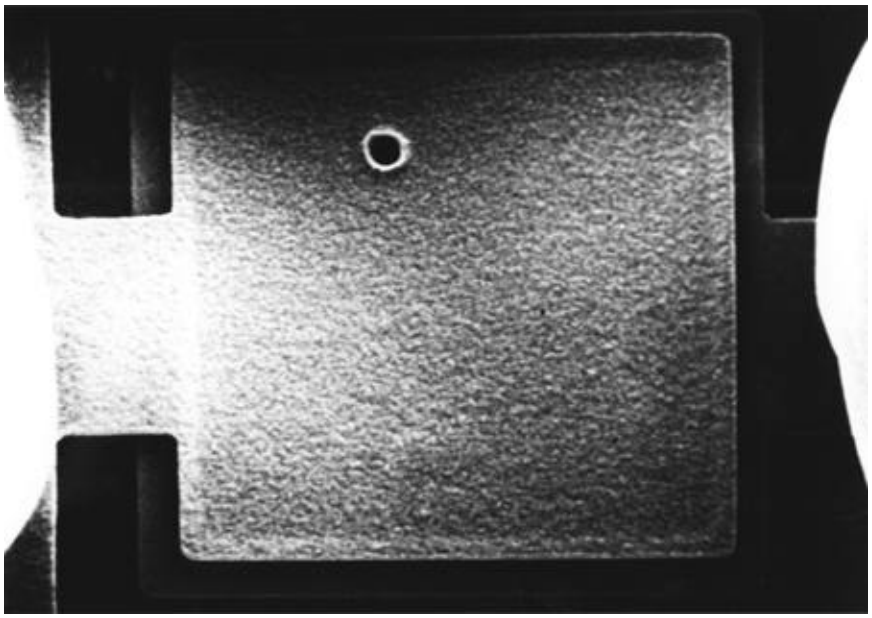

Figure 4-8. Interior-located ESD failure of a MIM capacitor. (Courtesy of Tri Quint Semiconductor.)

mechanisms. Anodic gold corrosion is the main culprit of GaAs device failures in high humidity environments where gold hydroxide ($Au(OH)_3$) has been detected in tests of GaAs ohmic contacts under high humidity conditions [28]. Ni filamentary growth, shown in Figure 4-9, has also been observed along the electric field direction of ohmic contacts adjacent to gate fingers [29].

Arsenic dissolution has also been reported as a humidity accelerated failure mechanism [29]. This effect is theorized to lead to reduction of channel thickness and degradation of device parameters such as I_{DSS} and the channel parasitic resistance.

2. Hydrogen Effects

The effect of hydrogen on the performance and reliability of GaAs devices has been reported over the last few years [30,31,32]. Degradation in I_{DSS}, V_p, g_m, and output power was observed on devices tested in hermetically sealed packages or under hydrogen atmosphere. The source of the degradation has been attributed to hydrogen gas desorbed from the package metals (Kovar, plating, etc.). The exact mechanism by which hydrogen degrades the device performance and the path by which hydrogen reaches the active area of a device are not known and have been under investigation.

Earlier research on GaAs transistors identified the diffusion of atomic hydrogen directly into the channel area of the device where it neutralizes the silicon donors as the

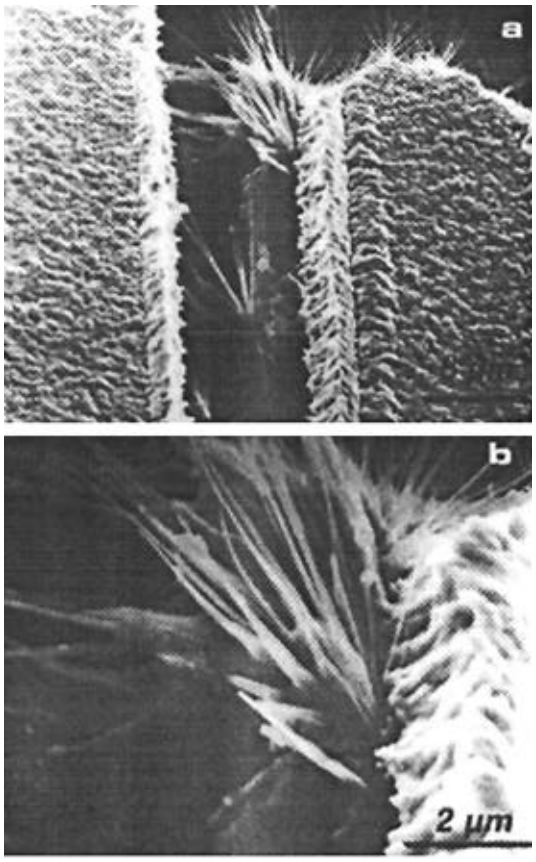

Figure 4-9. Filamentary growth: (a) nickel extrusion from the AuGeNi ohmic contact of an Au/Pd/Ti low-power MESFET passivated by Si_3N_4, submitted to an 85% RH/125°C HAST test and (b) enlarged view, evidencing dimensions of the whiskers. (From [12]; reprinted by permission of John Wiley & Sons, Ltd.)

possible mechanism [31]. It is believed that atomic hydrogen diffuses into the GaAs channel and forms Si-H, thereby neutralizing the donors. Experiments have shown that exposure of Si-doped GaAs to RF hydrogen plasma results in neutralization of the Si donors. Infrared spectroscopy data have also given evidence of $(SiAs_3)As$-H complexes [31,33].

The neutralization of donors can decrease the carrier concentration in the channel, which, in turn, can decrease the drain current, transconductance, and gain of the device. Hydrogen effects in FETs with either Pt or Pd gate metals have been observed. Recent research has concluded that the diffusion of hydrogen may occur at the Pt sidewalls and not at the Au surface of the Au/Pt/Ti gate metal [34].

Other research on GaAs PHEMT and InP HEMT in a hydrogen atmosphere has shown that the drain current may increase in some cases (Figure 4-10). This observation has led to the conclusion that the hydrogen diffuses into the semiconductor surface where it is thought to change the metal–semiconductor built-in potential [35].

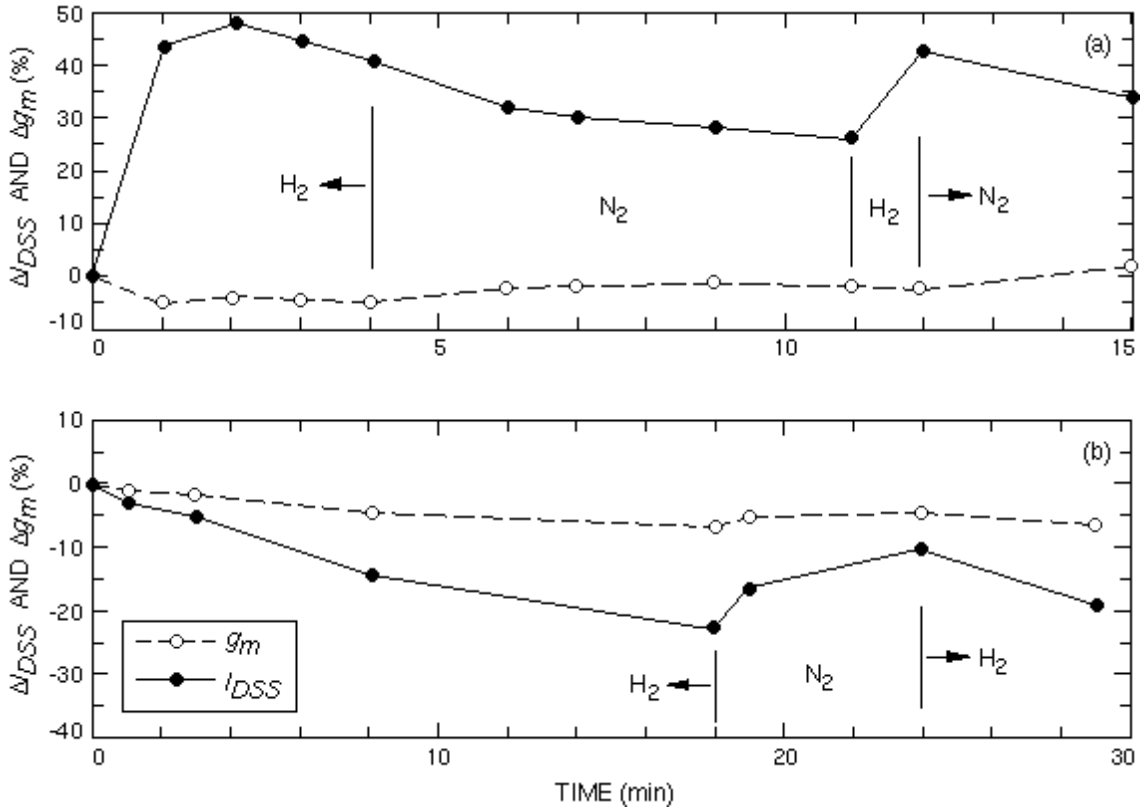

Figure 4-10. Changes in peak transconductance, g_m, and drain current at zero gate bias, I_{dss}, of (a) InP HEMT and (b) GaAs PHEMT under nitrogen and 4% hydrogen treatments at 270°C. The devices were unbiased during the treatments. The measurements were performed at room temperature. (From [35]; ©1994 IEEE.)

Manufacturers and users of GaAs devices used in hermetically sealed packages are currently pursuing an acceptable solution to this problem. Some of the possible solutions include thermal treatment of the packaging materials to reduce the amount of desorbed hydrogen after the seal, the use of hydrogen getter materials in hermetically sealed packages, and the use of barrier materials that do not contain the Pt/Ti or Pd/Ti structures. These solutions have limitations and possible instability problems that must be fully understood prior to implementation in high reliability environments [1].

3. Ionic Contamination

Ionic contamination in semiconductor devices is one of the important failure mechanisms. As a result of mobile ion contamination, GaAs-based devices can suffer changes in the carrier concentration resulting in threshold voltage shifts, an increase in leakage current, and gain reduction. Mobile alkali ions, such as N^+, Cl^-, and K^+, are the most common contaminants and have been identified by spectroscopic analysis to be the

principal causes of failure. The ionic contaminant must be in the form of a solution in order to be mobile and cause the referenced detrimental effects. The ion mobility is thermally and electric-field accelerated.

The existence of surface states and nonuniformities at the material interfaces of GaAs devices promotes the existence of conducting channels. This results in an increase in leakage current and a reduction of breakdown voltage. The surface ions can also contribute to surface leakage currents by creating a conductive path between adjacent metal lines. This may take the form of an electrolytic process involving the corrosion of the metallization, which will result in the formation of voids in the metal, and hence device failure.

Ionic contamination can arise during processing, packaging and interconnect, test, and operation in an unprotected environment. Surface preparation and cleanliness, characterization and control of processing materials and environments, and protection (passivation) of the active area of the devices can reduce or eliminate any ionic-contamination-related failures. High-temperature storage bake and exposure to high temperature during burn-in have been found to be effective methods of detecting ionic contamination problems.

References

[1] S. Kayali, "Hydrogen Effects on GaAs Device Reliability," *Proc. 1996 International Conference on GaAs Manufacturing Technology*.

[2] S. D. Mukherjee, "Interactions of Metal Films on Semiconductors," *Reliability and Degradation*, M. J. Howes and D. V. Morgan, Editors, John Wiley & Sons, New York, 1981.

[3] C. Canali, F. Castaldo, F. Fantini, D. Ogliari, L. Umena, and E. Zanoni, "Gate Metallization Sinking Into the Active Channel in Ti/W/Au Metallized Power MESFETs," *IEEE E. Dev. Lett.*, Vol. EDL-7, pp. 185–187, 1986.

[4] J. Wurfl and H. L. Hartnagel, "Thermal Stability and Degradation Mechanisms of Ti-Pt-Au, Ti-W-Au and WSi_2-Ti-Pt-Au Schottky Contacts on GaAs," *Proc. 4th Int. Conf. Quality in Electronic Components,* Bordeaux, pp. 133–138 1989.

[5] T. Hashinaga, M. Nishiguchi, and H. Nishizawa "Active Layer Thinning due to Metal-GaAs Reactions in GaAs Ics with Ti/Pt/Au Gate Metallization," *Proc. 1991 GaAs Rel. Workshop*, 1991.

[6] K. Mizuishi, H. Kurano, H. Sato, and H. Kodera, "Degradation Mechanisms of GaAs MESFETs," *IEEE Trans. on Electron Devices,* Vol. ED-26, No. 7, pp. 1008–1014, July 1979.

[7] J. F. Bresse, "Reliability of the Structure Au/Cr/Au-Ge/Ni/GaAs in Low Noise Dual Gate GaAs FET," *Microelec. Reliab*., Vol. 25, pp. 411–424, 1985.

[8] T. Irie, I. Nagasako, H. Kohzu, and K. Sekido, "Reliability Study of Power GaAs MESFETs," *IEEE Trans. on Microw. Theory and Tech.*, Vol. MTT-24, pp. 321–328, 1976.

[9] J. C. Irvin, "The Reliability of GaAs FETs," Ch. 6 in *GaAs FET Principles and Technology*, J. DiLorenzo and D. Khandelwal, Editors, Artech House, Mass., pp. 348–400, 1985.

[10] T. S. Kuan, P. E. Batson, T. N. Jackson, H. Rupprecht, and E. L. Wilkie, "Electron Microscope Studies of an Alloyed Au/Ni/Au-Ge Ohmic Contact to GaAs," *J. Appl. Phys*., Vol. 45, pp. 6952–6957, 1983.

[11] Y. C. Shih, M. Murakami, E. L. Wilkie, and A. C. Callegari, "Effect of Interfacial Microstructure on Uniformity and Thermal Stability of AuNiGe Ohmic Contact to n-Type GaAs," *J. Appl. Phys.*, Vol. 62, pp. 582-590, 1987.

[12] F. Magistrali, C. Tedesco, E. Zanoni, and C. Canali, "Reliability Issues of Discrete FETs and HEMTs," *Reliability of Gallium Arsenide MMICs*, A. Christou, ed., John Wiley & Sons, New York, 1992.

[13] M. Omori, J. Wholey, and J. Gibbons, "Accelerated Active Life Test of GaAs FET and New Failure Mode," *Proc. International Reliability Physics Symposium*, pp. 134–9, 1980.

[14] W. T. Anderson, A. Christou, F. A. Buot, J. Archer, G. Bechtel, H. Cooke, Y. C. Pao, M. Simons, and E. W. Chase, "Reliability of Discrete MODFETs: Life Testing, Radiation Effects and ESD," *Proc. IRPS*, 54–59, 1988.

[15] A. Christou, "Reliability of High Speed HEMT Integrated Circuits and Multi-2DEG Structure," *Semiconductor Device Reliability*, A. Christou, and B. A. Unger eds., Kluwer Academic, pp. 454–555, 1990.

[16] P. H. Ladbrooke and S. R. Blight, "Low-Field Low Mobility Dispersion of Transconductance in GaAs MESFETs with Implication for Other Rate-Dependent Anomalies," *IEEE Trans. El. Dev*., Vol. ED-35, pp. 257–267, 1988.

[17] K. Morizane, M. Dosen, and Y. Mori, "A Mechanism of Source-Drain Burnout in GaAs MESFETs," *Proc. GaAs and Related Compounds*, p. 287, 1987.

[18] A. Watanabe, K. Fujimoto, M. Oda, T. Nakatsuka, and A. Tamura, "Rapid Degradation of Wsi Self-Aligned Gate MESFET by Hot Carrier Effect," in *Proc. Int'l Reliability Physics Symp*., pp. 127–130, April 1992.

[19] H. Hasewaga, K. Katsukawa, T. Itoh, T. Noguchi, and Y. Kneko, "High Reliability Power GaAs MESFET Under RF Overdrive Condition," in *IEEE MTT-S Int'l Microwave Symp*., pp. 289–292, 1993.

[20] Y. A. Tkachenko, Y. Lan, D. Whitefield, C. J. Wei, J. C. M. Hwang, L. Aucoin and S. Shanfield, "Gradual Degradation Under RF Overdrive of Power GaAs Field-Effect Transistors," *Proc. GaAs Reliability Workshop*, 1993.

[21] Y. A. Tkachenko, Y. Lan, D. Whitefield, C. J. Wei, J. C. M. Hwang, L. Aucoin, and S. Shanfield, "Hot-Electron-Induced Degradation of Metal-Semiconductor Field-Effect Transistors," *Proc. of IEEE GaAs IC Symp*., pp. 259–262, 1994.

[22] D. M. Hwang, L. Dechiaro, M. C. Wang, P. S. D. Lin, C. E. Zah, S. Ovadia, T. P. Lee, D. Darby, Y. A. Tkachenko, and J. C. M. Hwang, "High-Voltage Electron-Beam-Induced Current Imaging of Microdefects in Laser Diodes and MESFETs," *Proc. Int'l Reliability Phys. Symp*., pp. 470–477, 1994.

[23] D. Ragle, K. Decker, and M. Loy, "ESD Defects on GaAs MESFET Lifetime," *Proc. Int. Reliability Physics Symp*., p. 352, 1993.

[24] A. Rubalcava, D. Stunkard, and W. Roesch, "Electrostatic Discharge Effects on Gallium Arsenide Integrated Circuits," *EOS/ESD Symposium Proceedings*, 1988.

[25] W. T. Anderson, F. A. Buot, A. Christou, Y. Anand, *Proc. 24th Annual Rel.Phys*., p.144, 1986

[26] K. Bock, K. Fricke, V. Krozer, and H. Hartnagel, *EOS/ESD Symposium Proceedings*, 1990.

[27] M. Pecht and P. Lall, “General Reliability Considerations as Applied to MMICs,” *Reliability of Gallium Arsenide MMICs,* A. Christou, Editor, John Wiley & Sons, pp. 40–60, 1992.

[28] E. Brambilla, P. Brambilla, C. Canali, F. Fantini, and M. Vanzi, “Anodic Gold Corrosion in Plastic Encapsulated Devices,” *Microelect. Reliab.,* Vol. 23, pp. 577–585, 1985.

[29] F. Migistrali, D. Ogliari, M. Sangalli, and M. Vanzi, “Degradation Mechanisms of GaAs MESFET Devices in High Humidity Conditions,” *Proc. ISTFA 1989*, pp. 141–151.

[30] W. Camp, R. Lasater, V. Genova, and R. Hume, “Hydrogen Effects on the Reliability of GaAs MMICs,” *Proc. GaAs IC Symposium, 1989*.

[31] J. Chevallier, W. Dautremont-Smith, C. Tu, and S. Pearton, “Donor Neutralization in GaAs(Si) by Atomic Hydrogen,” *Appl. Phys. Lett.* 47, p. 108, July 1985.

[32] S. Kayali, “Hydrogen Effects on GaAs, Status and Progress,” *Proc*. *1995 GaAs Reliability Workshop*.

[33] A. Jalil, J. Chevallier, J. Pesant, and R. Mostefaoui, “Infrared Spectroscopic Evidence of Silicon Related Hydrogen Complexes in Hydrogenated n-Type GaAs Doped with Silicon,” *Appl. Phys. Lett.*, Vol. 50, No. 8, p. 439, February 1987.

[34] D. Eng, R. Culbertson, and K. MacWillams, “The Effects of Hydrogen and Deuterium Incorporation on the Electrical Performance of GaAs MESFETs,” *Proc. 1995 GaAs IC Symposium.*

[35] P. C. Chao, M. Kao, K. Nordheden, and A. Swanson, “HEMT Degradation in Hydrogen Gas,” *IEEE Elec. Dev. Lett.*, Vol. 15, No. 5, p. 151, May 1994.

Chapter 5. Device Modeling

W. Y. Jiang

It is extremely important to complete MMIC device modeling and simulation prior to the fabrication because the technology and design iteration are expensive and the technology often does not allow postfabrication tuning. Therefore, model accuracy is an essential part of first-pass design success. Device modeling is useful not only in design, but also in production control and yield analysis.

This chapter will describe the general subjects related to MMIC device modeling, including the types of models, equivalent circuits, modeling approach, and commercially available modeling software. The issue of model sensitivity will also be discussed. Although the content emphasizes MESFETs, the methodology used can be applied to other MMIC devices, such as HEMTs, HBTs, and diodes.

I. Types of Models

A device model can be composed of a set of equivalent circuit elements in a particular circuit topology or a set of equations that, when evaluated, predict device performance. A modeling process generally includes three steps: characterization, parameter extraction and modeling. The flow chart of a typical modeling process is illustrated in Figure 5-1. Three processes are closely related in a number of important

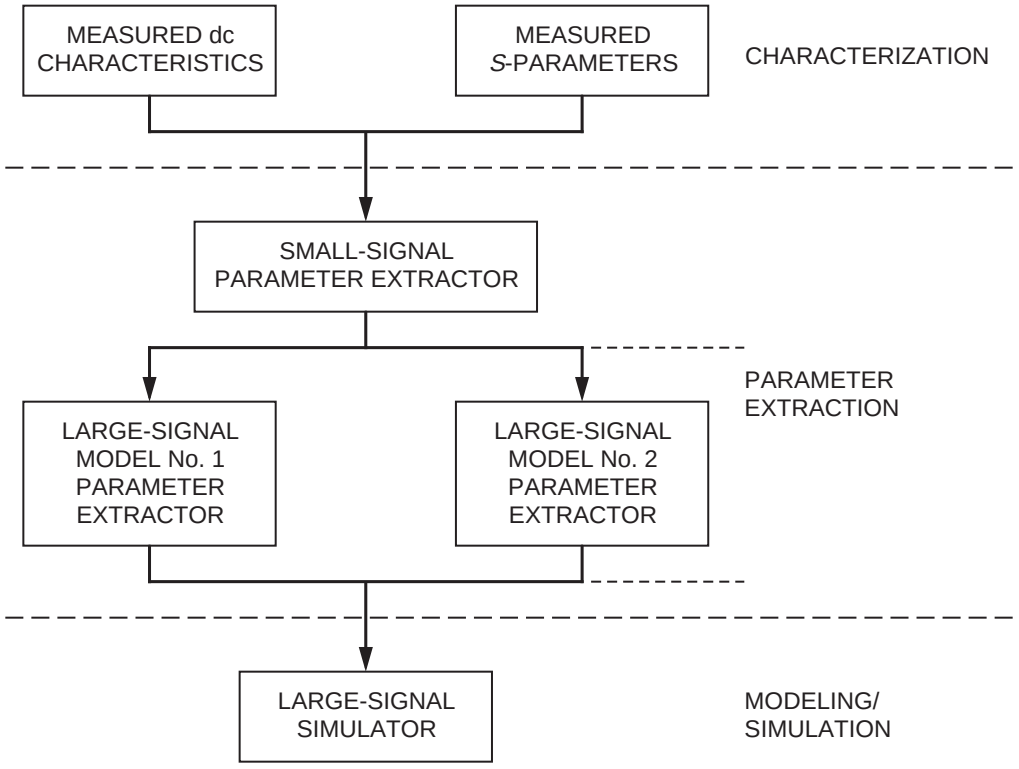

Figure 5-1. Flow chart of the relationship between characterization, parameter extraction, and modeling.

ways. The accuracy of any device model ultimately is limited by how accurately the model parameters are determined. Parameter extraction is dependent on the type and accuracy of available device characterization data. The merits of the device model are partially determined by the amount and type of characterization required. Generally speaking, MMIC device modeling can be classified into three categories: Empirical Device Models (EDMs), Physically Based Models (PBMs), and data-based models.

EDMs use equivalent circuits to simulate the external behavior of devices. Such a model consists of a number of linear and nonlinear elements connected in a predefined topology. Various EDMs, including small signal and large signal, have been widely used in MMIC computer-aided engineering. The advantages of EDMs are simple characterization, implementation, and circuit simulation.

To obtain their performance predictions, PBMs rely on physical parameters that describe the device geometry, materials, and processing parameters. Such parameters typically include gate length, gate width, channel thickness, and doping density. PBMs have an advantage over EDMs: PBMs allow studies of the effects of process variation on the device performance; such effects are critical for process control and yield prediction. However, it is difficult, in some cases even impossible, to obtain the precise physical parameters required to describe the device.

Recently, data-based models (also known as measurement-based) have become popular with device designers. Data-based models are generated directly from measured data without prior knowledge of process parameters. A data-based model can predict behavior exhibited in a new process that may be difficult to represent by empirical functions. However, its lack of physical insight into the actual studied device is a drawback.

II. Equivalent Circuit

The equivalent circuit of an MMIC device is an abstraction and simplification that yields a representation of the device. It must represent adequately all the important physical characteristics of the device. Exploiting the relationship between the equivalent-circuit elements and device physics will be helpful to device modeling.

A. MESFET Equivalent Circuit

The device physics of MESFETs and HEMTs have been discussed in Sections 3-III and 3-IV, respectively. The material and structure features that determine the microwave behavior of a FET are identified on Figure 5-2; some of relevant parameters are [1]

N = doping density in the n - channel layer

W = thickness of the n - channel layer under the gate

Z_G = gate width

L_G = metallurgical gate length

L_{SG} = source – gate separation

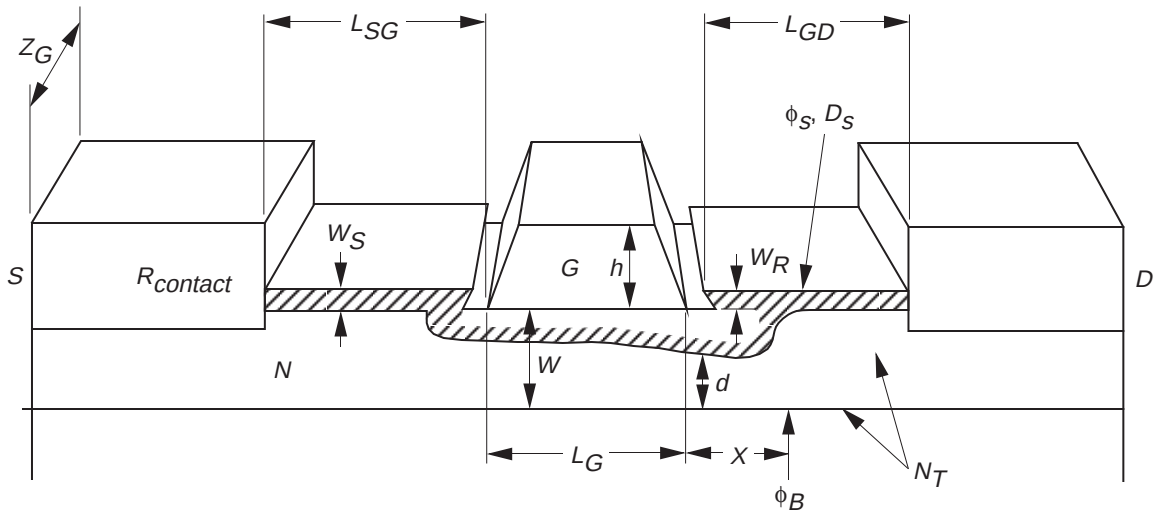

Figure 5-2. Schematic of a MESFET's material and structure.

L_{GD} = drain – gate separation
W_R = depth of gate recess
W_S = surface depletion depth
d = depletion depth
h = gate height
X = extension of the space - charge layer into the gate – drain space

The small-signal equivalent circuit for such a MESFET is presented in Figure 5-3.

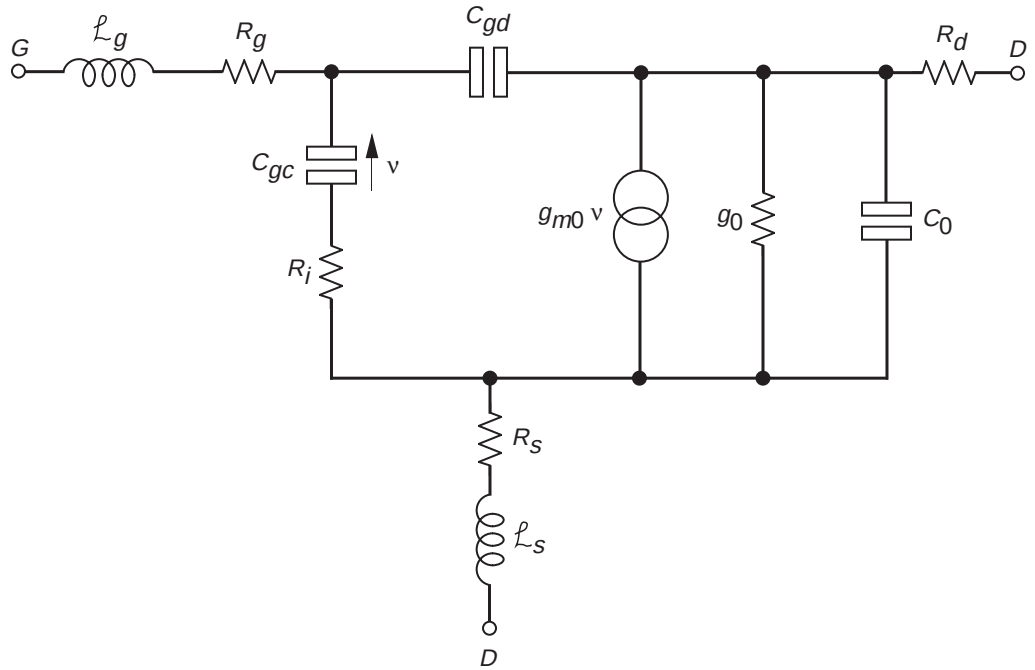

Figure 5-3. Basic GaAs MESFET's equivalent circuit.

In the following, the relationship between some of the equivalent-circuit elements and devices physics will be briefly explained:

(1) Channel resistance, R_i , is the resistance distributed along the channel under the gate, which is the ratio of the potential drop, $E_S L_G$, and the channel current, I_{CH}. The electric field and channel current under the gate are

$$E_S = \frac{v_{sat}}{\mu_0'}; \quad I_{CH} = qNv_{sat}(W-d)Z_G$$

where v_{sat} is the saturated value of electron drift velocity, μ_0' is low-field drift mobility, and q is electron charge. Therefore the channel resistance is

$$R_i \approx \frac{L_G}{\mu_0' qN(W-d)Z_G}$$

(2) Transconductance, g_{m0}, is the ratio of change of drain current and gate voltage. As a first order approximation, it is reasonable to use the channel current to replace the drain current while omitting the substrate current. Using the expressions for I_{CH} and

$$V_{S'G} + V_{BO} \approx \frac{qNd^2}{2\varepsilon}$$

where $V_{S'G}$ is the dc voltage between the gate and virtual source—taking an account of R_S, V_{BO} is the equilibrium contact potential between gate metal and the N-GaAs layer, and ε is permittivity. The transconductance is

$$g_{m0} \equiv \frac{\partial I_D}{\partial V_{S'G}} \approx \frac{\partial I_{CH}}{\partial V_{S'G}} = \frac{\varepsilon v_{sat} Z_G}{d}$$

(3) Gate-channel space capacitance, C_{gc}, is the capacitance of the gate. As a first order approximation, it can be treated as a parallel plate stripline with dimensions of gate length, L_G, and the gate width, Z_G. That capacitance is

$$C_{gc}' = \frac{\varepsilon L_G Z_G}{d}$$

where C_{gc}' is an approximation. Taking account of the capacitance in the velocity-saturated region located near the tail of the gate, the gate-channel capacitance is

$$C_{gc} = \frac{\varepsilon L_G Z_G}{d}\left(1 + \frac{X}{2L_G} - \frac{2d}{L_G + 2X}\right)$$

(4) Gate-drain space capacitance, C_{gd}, is associated with the electron inflow at the right edge of the space-charge layer. The depletion extension X increases slightly as drain–source voltage increases, resulting in charge storage. Assuming W_R equals to W_S, the capacitance is

$$C_{gd} = \frac{2\varepsilon L_G Z_G}{L_G + 2X}$$

(5) Gate series inductance, L_g,, is the inductance determined by the strip's dimension of the gate length, L_G, and the gate width, Z_G. The value of L_g can be assessed by regarding the gate as a section of a parallel plate stripline:

$$L_g = \frac{\mu_0 d Z_G}{L_G}$$

where μ_0 is the permeability of free space.

(6) Gate resistance, R_g, is bulk resistance determined by the strip's dimension in the direction of current flow, which is the cross-sectional area, $L_G \times h$, and the gate width, Z_G. Since h is normally smaller than the skin depth, the whole height of the gate contributes to its conductance. Taking into account the voltage drop across the strip width due to distributed capacitance, the RF resistance, R_g, of the strip is only one third of the dc resistance:

$$R_g = \frac{\rho Z_G}{3hL_G}$$

where ρ is metal resistivity.

(7) Drain resistance, R_d, is the ratio of the voltage change across the length of the bulk region ($L_{GD} - X$) and the channel current, I_{CH}. The channel current in this region becomes $I_{CH} = qNv_{sat}WZ_G$. Thus the drain resistance can be obtained in a way similar to that for R_i:

$$R_d \approx \frac{L_G - X}{\mu_0' qNWZ_G}$$

B. HEMT Equivalent Circuit

The equivalent circuit for HEMT is the same as that for MESFET except for the gate-leaking current in some HEMTs, which may require resistances in parallel with C_{gc} and C_{gd}, respectively. Applying structure and operation conditions, the analytic expressions for some equivalent-circuit elements of HEMTs, such as g_{m0}, C_{gc}, and C_{gd}, can be derived [1].

III. Characterization and Parameter Extraction

The various types of data that might be required for use in the device modeling process include dc I–V characteristics, microwave *S*-parameters, large-signal *S*-parameters or load-pull characteristics, noise parameters, and physical characteristics of the device. The parameter extraction obtains appropriate equivalent-circuit element values from measured data using optimizers to minimize the error between simulated and measured data [2].

A. DC Characterization and Parameter Extraction

The primary advantage of dc data for model parameter extraction is ease of performance. A typical dc-modeling process for MESFET devices is shown in Figure 5-4. Although dc data fail to describe the RF characteristics of the device—such as strong frequency-dependent output conductance, g_0—they are quite useful as first-order estimates of device performance characteristics.

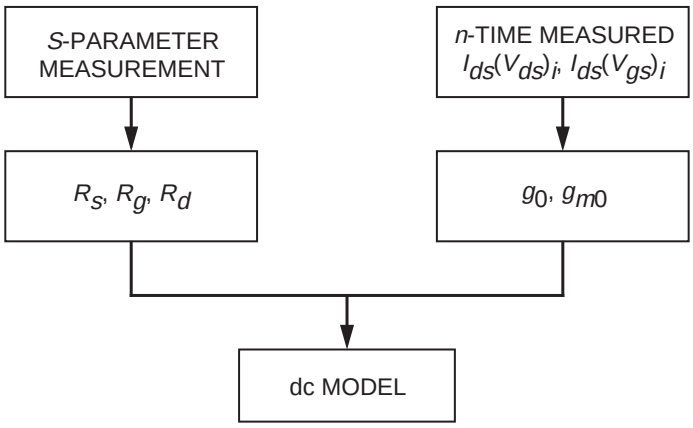

Figure 5-4. Flow chart of dc modeling.

B. RF Characterization and Parameter Extraction

RF or small-signal characterization of devices commonly uses microwave *S*-parameter measurement. Typically, *S*-parameters taken at 5 to 20 frequencies between dc and the upper frequency of interest are sufficient to determine element values. Automated measurement equipment that performs these measurements is readily available. The flow chart for RF characterization and parameter extraction is presented in Figure 5-5. With selected equivalent-circuit topology, the initial value of the circuit elements, which is estimated using a combination of the dc-parameter extraction technique and RF measurement, can be determined. An optimization routine is then run to refine the estimates of the element values until an appropriate agreement between the measured and modeled values is reached. However, the model is valid only under linear operating conditions.

C. Large-Signal Characterization and Parameter Extraction

Large-signal, or nonlinear characterization is important to any MMIC device whose performance objects include gain compression, saturated power, efficiency,

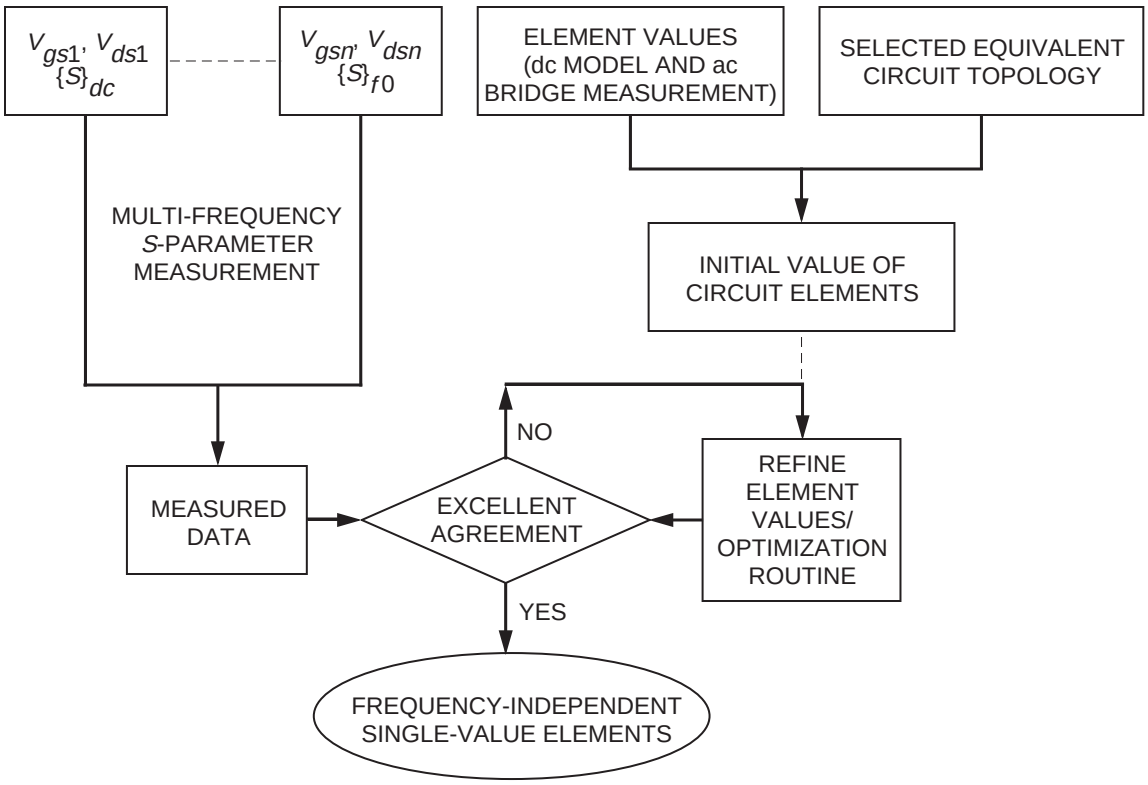

Figure 5-5. Small-signal direct model extraction process for *S*-parameter measurement at multiple frequencies.

harmonic distortion, and multitone intermodulation distortion products. There are a number of models for large-signal GaAs MESFETs; among the popular industry standards are the Curtice quadratic, Curtice cubic, and Statz(Raytheon) models. The equivalent circuit and I–V expression are different from model to model. Taking the Curtice cubic model for example, nonlinear I–V is expressed using a cubic approximation [3]:

$$I_{ds} = \left(A_0 + A_1 V_{in} + A_2 V_{in}^2 + A_3 V_{in}^3\right) \bullet \tanh\left(\gamma \bullet V_{out}(t)\right)$$

The coefficients A_i and γ are arbitrary empirical parameters whose values can be determined through parameter extraction, such that the evaluation of the equation is consistent with measured characteristics.

Large-signal characterization can be accomplished by two commonly used measurement techniques: load-pull and large-signal *S*-parameter characterization. Both techniques require measurements carried out on multiple signal levels to obtain complete characterization. The flow chart for device characterization and parameter extraction of large signal is shown in Figure 5-6. Large-signal models are applicable for both linear and nonlinear applications.

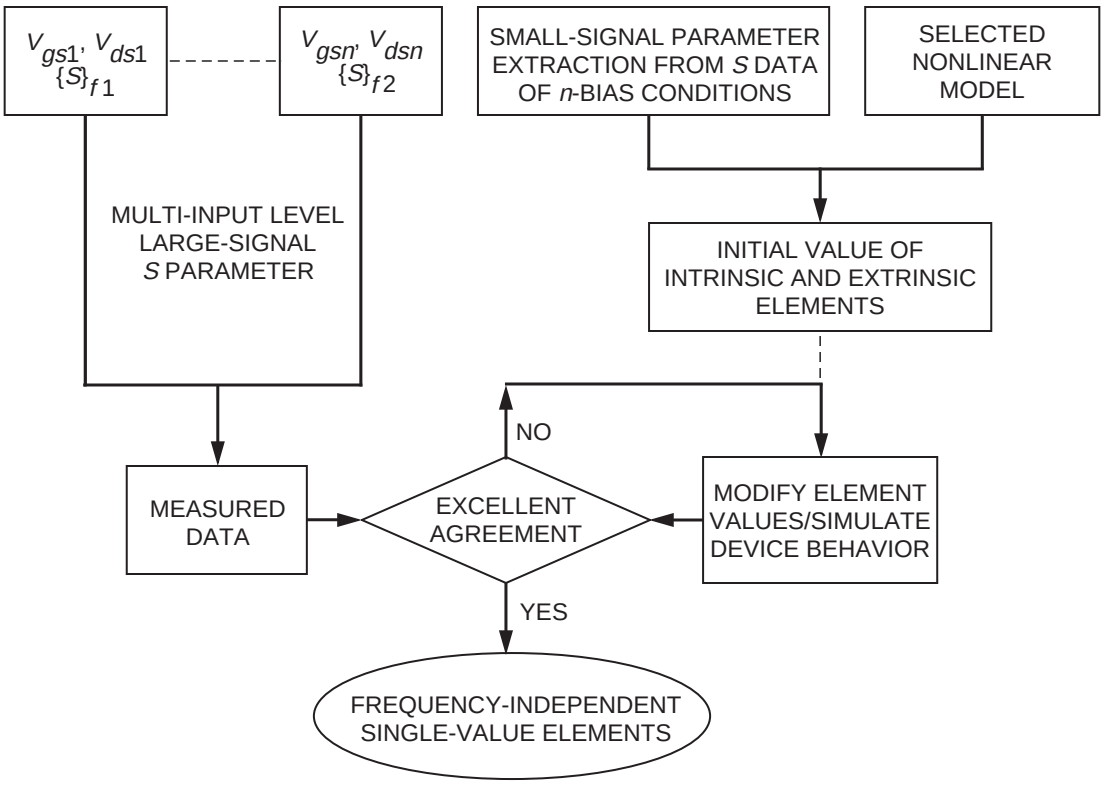

Figure 5-6. Typical flow chart of characterization and parameter extraction for large-signal model.

D. Noise Figure Characterization

A prime application of GaAs MESFETs has been in low-noise amplification. It is important to derive a simple analytic expression for calculating the minimum noise figure of a FET. Since the noise figure of a FET is affected by both bias point and generator impedance, the minimum noise figure, NF_{min}, defined here is an absolute minimum noise figure obtained by adjusting both bias and generator impedance.

Using the four equivalent element values—g_{mo}, C_{gc}, R_s, and R_g, determined by *S*-parameter measurement and small-signal parameter extraction—Fukui empirically derived a simple expression for NF_{min} [4]:

$$NF_{\min} \approx 1 + K_F \omega C_{gc} \left(\frac{R_s + R_g}{g_{mo}} \right)^{1/2}$$

where the factor $K_F \approx 2.5$ to 3.0 for FETs and $K_F \approx 1.5$ to 2.0 for HEMTs. The factor K_F is a gross simplification of the drain-current noise contribution to the overall noise. A noise model with an equally simple expression but retaining more comprehensive physics was derived by Delagebeaudeuf et. al. [5]:

$$NF_{\min} \approx 1 + 2\left(\frac{\mu_0' I_{CH}}{g_{mo} v_{sat} L_G} \right)^{1/2} \omega C_{gc} \left(\frac{R_s + R_g}{R_i} \right)^{1/2}$$

The first bracket is the expression for K_F in Fukui's model, which is related to channel current, I_{CH}, transconductance, gate length, and the saturated value of electron drift velocity. Using the relationship of the equivalent circuit elements and physical parameters presented in Section II, the expression can be further simplified:

$$NF_{\min} \approx 1 + 2\omega \frac{C_{gc}}{g_{mo}} \left(\frac{R_s + R_g}{R_i} \right)^{1/2}$$

Delagebeaudeuf et. al. have concluded that the equation also applies to HEMTs.

In a practical case, the generator impedance, $Z_g = R_g + jX_g$, connected at the input port, is a key factor influencing the noise figure of a circuit. The effect of this on the noise figure is given by [4]

$$NF = NF_{\min} + \frac{R_n}{R_g} \left[\frac{\left(R_g - R_{op}\right)^2 + \left(X_g - X_{op}\right)^2}{R_{op}^2 + X_{op}^2} \right]$$

where R_n is the equivalent noise resistance, and R_{op} and X_{op} are the optimum generator's resistance and reactance, respectively. Therefore, NF_{min}, R_n, R_{op}, and X_{op} are commonly defined as the characteristic noise parameters of the device.

IV. Modeling Software

MMIC modeling software includes device modeling and process modeling. Since there are a number of device-modeling softwares available, it is necessary to examine the compatibility of the software used by customers and the foundry, and between that used for modeling and simulation.

A. Device Modeling Software

HP/EEsof's Integrated Circuit Characterization and Analysis Programs (IC-CAP™) modeling suite Release 4.4 (HP 85190A) is a UNIX-based device-modeling toolset. The software allows users to develop their own model equation and extraction techniques, but also provides turn-key modules for a wide range of popular device models, including MESFET, HEMT, and HBT EDMs, as well as PBMs and HP Root data-based models. The software modules include measurement set-up, mathematical transforms, automation macros, and optimization routines to facilitate modeling. The model parameters are extracted by applying mathematical transforms to measured data. A Parameter Extraction Language (PEL) is built-in to facilitate creation of the transforms. The results of simulation based on the extracted model parameters can be plotted together with the measured data. IC-CAP contains three SPICE simulators and provides direct links to external simulators. The software has several optimization algorithms and user-controlled optimization settings. The sensitivity analysis mode provides information on important parameters for a particular optimization. The distinct feature of this software is the combined capabilities of instrument control, data acquisition, graphic analysis and optimization for device modeling. The software is compatible with HP/EEsof's Series IV™ and MDS™ simulation tools.

Optimization Systems Associates' (OSA) HarPE 2.0™ is a workstation-based nonlinear device-modeling software that includes parameter extraction and advanced statistical modeling. The built-in intrinsic nonlinear models include most popular models in the industry, such as several nonlinear FET models, a Gummel-Poon model for BJTs, models for HEMTs, and models for HBTs. It allows users to modify the built-in models or create user-defined models. It is capable of parameter extraction from harmonic data obtained under RF large-signal excitation or small-signal *S*-parameters taken over a number of bias conditions. It offers the option of extracting the extrinsic parameters from cold (unbiased and pinched-off) measurements. The software statistic modeling capability provides realistic yield analysis and optimization. It accepts *S*-parameter files in the touchstone format or the MDIF format, as well as on-wafer measurement data produced by Cascade Microtech's MicroCAT Test Executive system. The gradient-based minimax, L1, least-square, and Huber optimizers provide flexibility and accuracy in the modeling process. HarPE™ runs under X-windows on Hewlett-Packard, Sun, and DEC workstations.

Compact Software's Compact Scout™ is a PC- and workstation-based active device parameter-extraction and large-signal modeling software. It is based on the Modified Materka-Kacprzak model. The software uses measured data provided by the users to extract and fit the nonlinear model. It has interactive modeling features, which allow the users to modify parameter values and quickly observe the effect on both dc and *S*-parameter. It can optimize the model coefficients to fit both dc and *S*-parameter. An extensive parasitic model has been built around the intrinsic model for chip and package parasitic modeling. The software is compatible with Compact's Super-Compact and Microwave Harmonica.

Optotek's Small and Large Signal Analysis (SALSA™) is a PC-based software dedicated to MESFET and HEMT modeling. For large-signal modeling, it includes most of the popular nonlinear I_{ds} MESFET models, nonlinear C_{gs} and C_{gd} MESFET models, nonlinear I_{ds} HEMT models, and one nonlinear PBM. The software also offers small-signal parameter extraction programs for both intrinsic and extrinsic elements. Two types of Newton optimizers and two types of random optimizers are provided for solving and fitting the measured and modeled data. The automated data acquisition is realized using an automated network analyzer and two programmable power supplies. It is compatible with Optotek's simulation software MMICAD™.

B. Processing Simulation Software

Stanford University Process Engineering Model (SUPREM™) is one of the widely used process simulation programs developed by Integrated Circuit Laboratory of Stanford University. SUPREM-IV.GS™ is an advanced 2D process simulator, which models GaAs and its dopants in addition to modeling silicon fabrication. The software provides physical models for ion implantation, and diffusion and annealing on a cross-section of arbitrary device structures. It also includes basic models for simulating etching and the deposition of thin films on the semiconductor surface, or it interfaces with other programs to accurately simulate these processes. SUPREM-IV.GS™ can model stress gradients produced by overlaying film. The implantation, diffusion, and annealing models are point-defect-based simulations.

SUPREM-IV.GS™ incorporates most features of the previously developed SUPREM 3.5's 1D GaAs models and parameters. SUPREM 3.5™ is primarily designed for modeling processes used to make simple ion-implanted MESFET and JFET structures in semi-insulating GaAs, with or without buried p-layers. The main processes modeled

are ion implantation and active annealing. The dopants modeled are Si, Se, Ge, and Sn (n-type), and Be, Mg, Zn and C (p-type). Also modeled is the diffusion for non-implanted dopants, such as those incorporated during MBE or MOCVD GaAs growth.

Based on SUPREM™ 3.5 and its added 2D capability, SUPREM-IV.GS™ includes some new features. The Pearson-IV implantation parameters have been included. The electron- or hole-dependent diffusion coefficients have been added for the eight dopants. Segregation coefficients, intrinsic carrier concentrations, and defect energy levels have also been included. Furthermore, compensation mechanisms for dopant activation and different diffusivities for implanted versus grown-in dopants have been added.

V. Model Sensitivity

The design of a device always begins with initial fixed parameters. The next problem is to determine the sensitivity of the device to variations in all those parameters, which include material and process parameters along with other factors, such as temperature and bias. A robust design may be achieved with a thorough sensitivity analysis.

A. Sensitivity Analysis

The fabrication of MMIC devices involves a large number of interrelated material and process parameters that influence the performance of MMIC devices. In practice, the material and process parameters inevitably will be different from the designed values, because of the control in the fabrication process. Therefore, it is important to determine the sensitivity of the circuit to the variation of each parameter [1].

Sensitivity analysis includes two levels. The first level analyzes the correlation between variation in a single material or process parameter and variations in equivalent circuit elements. For example, a single variation in the layer of doping, N, leads to correlated changes in C_{gc}, C_{gd}, R_i, g_{mo}, and the gate transit time, τ_{gm}. Other parameters, such as W, Z_G, and L_G also cause correlated changes in various equivalent-circuit elements. The second level analyzes the correlation between the material and process parameters themselves. For instance, the size or position error of the gate strip, L_G, in Figure 5-2 brings correlated changes in L_{SG} and L_{GD}. A typical flow chart for MMIC sensitivity analysis is presented in Figure 5-7.

B. Temperature Effect

The effects of temperature on MESFET performance include variations in transconductance, input capacitance, and device resistance. Transconductance variations are caused by an increase in electron mobility in the active channel as the temperature decreases. Variation of input capacitance is induced by an increase in the built-in Schottky voltage as temperature decreases. Resistance variation is caused by changes in the metallurgical nature of ohmic contacts in the source and drain areas at low temperature. Temperature changes have an impact on device equivalent circuit models and *S*-parameters, as well as noise-figure models. Using extensive *S*-parameter measurements at different temperatures, a modified model including the effects of temperature can be created. The modified GaAs FET model should accurately predict the

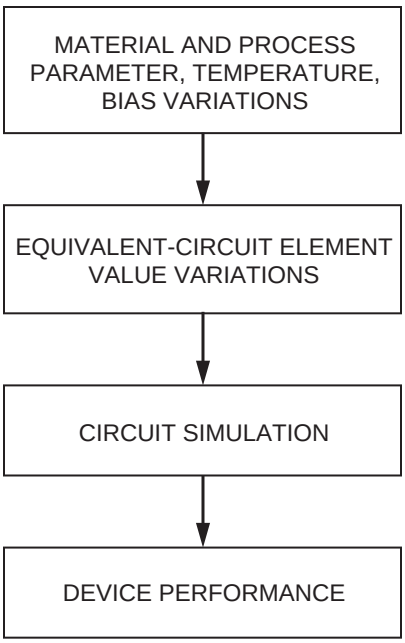

Figure 5-7. Flow chart for MMIC sensitivity analysis.

gain and noise figure at any temperature. Compact Software's Version 4.0 of Supercompact™ PC microwave simulation software has the capability of temperature-sensitivity simulation.

C. DC Bias Effect

As the discussion in Section 5-II indicates, the values of equivalent-circuit elements are directly or indirectly affected by dc biases. Therefore, the analysis of model sensitivity has to take dc bias effects into consideration.

Some elements, such as C_{gc}, C_{gd}, R_d, and τ_{gm}, are affected by the extension X of the depletion layer into the gate-drain space. X increases as $V_{D'G}$ increases and decreases as $V_{S'G}$ increases. Table 5-1 lists some important elements and their variation with biases.

Table 5-1. Relationship between variations of bias and element values.

Bias	Equivalent circuit element
$V_{GS'}$ ↑	g_{mo} ↑
$V_{GS'}$ ↑	R_i ↓
$V_{GS'}$ ↑	L_g ↓
$V_{D'S'}$ ↑, $V_{S'G}$ ↑	g_o ↓
V_{DS} ↑, V_{GS} ↑	C_{gc} ↑
V_{DS} ↑, V_{GS} ↑	C_{gd} ↓
V_{DS} ↑, V_{GS} ↑	τ_{gm} ↑

D. Statistical Analysis

In reality, the material and process parameters vary together, along with temperature and bias, unlike the previously described scenario in sensitivity analysis, where the effect of only one parameter variation is analyzed at a time. To assess the sensitivity of an MMIC to a given process parameter, it is better to approach the problem with an analysis that progressively restricts the parameter in question while continuously varying all others within known statistical limits; such a process is called statistical analysis [1]. The prerequisite for this technique is to establish the statistical distribution of all material and process parameters.

The Monte Carlo method is a popular and powerful technique for statistical analysis. Figure 5-8 is a flow chart of the Monte Carlo method applied to MMIC-yield forecasting. To have a reliable yield forecast using the technique, a large enough trial must be performed. The random values of material and processing parameters should duplicate exactly the empirically determined distribution. Where two or more tightly correlated distributions are involved, only one computer-generated random number is used to generate the correlated values. The Monte Carlo method can be applied not only to yield forecasting but also to assessment of design robustness and process control. It offers an alternative to the vastly expensive and time-consuming approach of practical trial and iteration of MMIC design and fabrication.

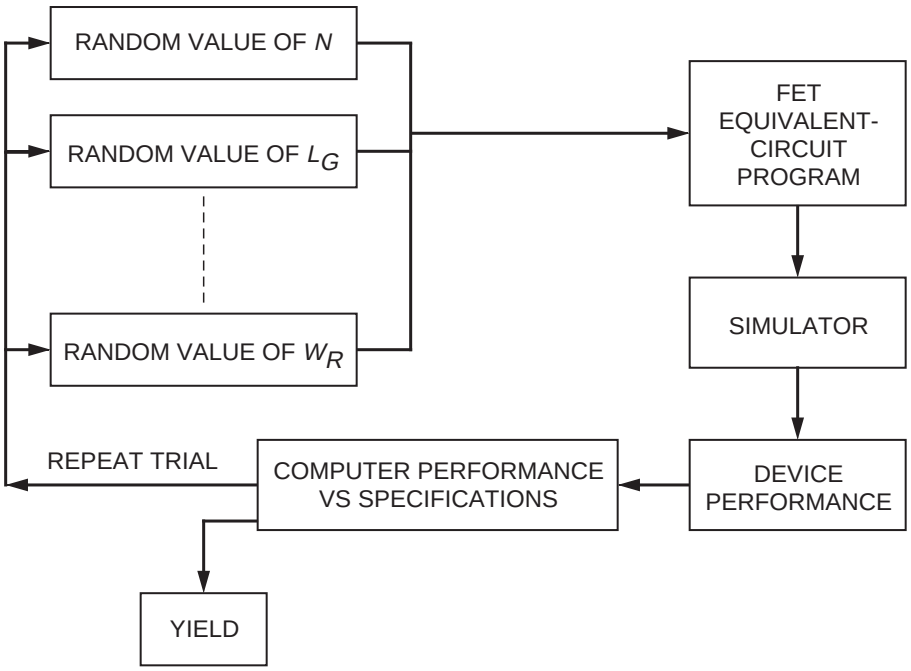

Figure 5-8. Application of the Monte Carlo method to MMIC yield forecasting.

References

[1] P. H. Ladbrooke, *MMIC Design: GaAs FETs and HEMTs,* Artech House, MA, 1989.

[2] *Monolithic Microwave Integrated Circuits: Technology and Design,* R. Goyal, Editor, Artech House, MA, 1989.

[3] W. R. Curtice and M. Ettenberg, "A Nonlinear GaAs FET Model for Use in the Design of Output Circuits for Power Amplifiers," *IEEE Trans. Microwave Theory and Tech.*, Vol. MTT-33, No. 12, pp. 1383–1393, 1985.

[4] H. Fukui, "Design of Microwave GaAs MESFETs for Broad-Band Low-Noise Amplifiers," *IEEE Trans. Microwave Theory and Tech.*, Vol. MTT-29, No. 10, pp. 176–183, 1979.

[5] D. Delagebeaudeuf, J. Chevrier, M. Laviron, and P. Delescuse, "A New Relationship Between the Fukui Coefficient and Optimal Current Value for Low-Noise Operation of Field-Effect Transistors," *IEEE Electron Device Lett.*, Vol-EDL-6, No. 9, pp. 444–445, 1985.

Chapter 6. MMIC Design Methodologies and Verification

C. Chen, S. Kayali, E. Rezek, and T. Trinh

The implementation of a MMIC design involves a number of circuit simulation, layout, fabrication, and testing steps. The large number of variables involved in these steps makes it imperative that all facets of the implementation be documented to assure repeatability of similar designs and improve the yield of the final product.

This chapter will describe the general aspects of MMIC design and the necessary tools available to both the user and manufacturer. This chapter will also provide a typical design methodology and flow used by MMIC foundries.

I. Foundry Documentation

A well-documented MMIC design methodology ensures a much smoother and faster turnaround of circuits. Circuit designs along with information relating to layout, processing, and testing can be preserved and used for future applications, thereby eliminating duplication of effort and providing substantial time and cost savings. This will not only assure a well-controlled and repeatable environment, which is essential for high-volume and high-yield applications, but also increases the probability of first-time success for new designs. As an added benefit, detailed documentation also guarantees a shorter period for the assimilation of new employees.

Documentation of a foundry's capabilities and design rules is necessary for internal use of the foundry's personnel and for external customers using the foundry's services to fabricate their own MMICs. Literature related to design rules, processing, and testing also provides the customer with an overall understanding of the end-to-end MMIC implementation.

In general, the available documentation should provide the interested customer with a description of the CAD tools, semiconductor processing steps, and test methods used at the foundry. From a users perspective, the level of documentation available at a MMIC foundry is a reflection of the maturity of the facility and the extent of process control being applied and practiced. Some typical documentation may include:

(1) Semiconductor processing capabilities.
(2) Design and layout rules used at the foundry.
(3) Design and layout tools used at the foundry.
(4) Available library designs.
(5) Available simulation tools.
(6) Available device and circuit element models.
(7) Design and processing flow.
(8) Design verification and review.
(9) Design and processing schedule information.
(10) Test methods.

II. MMIC Simulation

Circuit simulation is an essential step in the design and fabrication of MMICs for production purposes. Simulations can provide a first-order approximation of circuit functionality and performance under various input and output conditions prior to committing the design to fabrication. Since most simulators also include an optimization capability, circuits can be fine tuned, or in some cases synthesized, to meet the required performance specifications. This greatly reduces design turnaround time and increases the chances of first-time success. With the increased affordability of computing power and the recent advances in software development, many new software techniques and systems have become available for interactive MMIC design. The development of commercial software that integrates the various stages of MMIC design, such as schematic capture, simulation, and layout, has been the result of recent technology advancements and initiatives on MMIC CAD motivated by the identified need in the marketplace for these tools.

One of the biggest obstacles in transferring a customer-developed MMIC design to a MMIC foundry is the issue of design-tool compatibility. The customer must verify that the tools used by both entities for the design and simulation are compatible. The most common MMIC design and simulation tools are offered by HP/EEsof and Compact Software. Both companies provide a wide variety of tools for both linear and nonlinear simulations and links to layout.

Compact Software's Microwave Harmonica® is one of the primary design and simulation tools used for GaAs MMICs. This tool is used for both linear and nonlinear microwave circuit simulations. Microwave circuit structures are simulated using distributed element models, where nonlinear circuits are simulated by using harmonic balance techniques at the interface between the linear and nonlinear portions of the circuit. All the necessary components can be entered in either schematic or netlist form. This simulator also includes optimization, statistical analysis, yield optimization, and voltage synthesis, as well as oscillator and phase-noise analysis and optimization. Various portions of this software package can be used to address specific aspects of MMIC design.

The Compact Software Microwave Explorer® is a 3-D electromagnetic analysis tool used the for simulation of planar passive structures in both open and packaged environments. Circuits can be entered into the program through the use of either the Integrated Polygon Editor or the GDSII import utility. This package includes a graphics interface for viewing results on Smith charts, rectangular plots, or current distribution plots.

The Compact Microwave Success® is a block-level simulator used for examining such data as *S*-parameters and noise parameters for communication systems consisting of RF and microwave components. Success® systems can be built from a large set of models such as mixers, filters, antennae, and amplifiers. The package will also generate data in several other standard formats; it can provide analysis as a function of temperature, frequency, power, and other user-defined variables. The results can be displayed as time-domain waveforms, spectrum plots, sweep results, intermodulation results, and budget analysis.

HP EEsof's Libra® is another primary design and simulation tool used for both linear and nonlinear GaAs MMIC microwave-circuit simulations. Libra® performs frequency-domain simulations by using distributed element models for microwave circuit

structures. Nonlinear circuits are simulated by using harmonic balance techniques. Different portions of this software package can be used for the simulation and optimization of specific aspects of MMIC design. The HP/EEsof Libra Design Suite® is a simulation and layout toolset developed for RF and microwave design engineers. The Series IV Project Design Environment® is a graphical design environment for the design, simulation, layout, and documentation of high-frequency circuits and systems. This package includes capabilities for schematic capture, high-frequency circuit simulation, electromagnetic simulation, system simulation, and circuit layout, along with an extensive design library and various tools and links to third-party software.

The HP/EEsof Microwave Design System® is a UNIX-based computer-aided engineering (CAE) toolset tailored for high-frequency circuit and system design. This package has linear, nonlinear, transient-simulation, and sensitivity analysis capabilities and provides electromagnetic simulation and yield analysis. Design capture and circuit layout are also provided.

Other available simulation tools include Mathematica®, which is an interactive software package used primarily for the solution of complex mathematical problems and the development of mathematical models for microwave components and systems. Microwave Spice® is a time-domain circuit-simulation tool similar to the Berkeley Spice® package. This package includes many microwave effects and components that make it useful for microwave MMIC designers. This package is particularly useful for the development of microwave oscillators.

The Cadence Analog Artist Microwave Simulator Interface® is a linear frequency domain simulation tool that can be used within the Cadence IC environment of Analog Artist® with microwave extensions. The package can provide simulations using HP/EEsof or Compact software tools. Simulations can also be provided using Microwave Interconnect® layout format. Additionally, Cadence Spectre® and SpectreHDL® support frequency-domain blocks in time-domain analysis and behavioral modeling in the time domain; it is written in the SpectreHDL® language. The package can provide transient analysis via *S*-parameter blocks for microstrip lines, behavioral modeling for devices and circuit blocks, and mixed-level time-domain simulation.

EM simulation tools used either in conjunction with time- and frequency-domain simulators or as stand-alone EM simulators include Ansoft Maxwell Eminence®, which is a 3-D EM simulation tool consisting of a solid modeler for model entry, a simulation engine, and several data analysis features. The simulation engine creates its own mesh by way of an adaptive meshing algorithm. The user can specify both the convergence criteria and the maximum number of iterations in order to balance accuracy and time. Sonnet® is another 3-D EM simulation tool capable of accepting inputs in GDSII, HP/EEsof, Cadence, and AutoCAD formats. Outputs are generated as *S*-parameters, current distributions, or radiation patterns.

III. MMIC Layout

Conversion of the MMIC design into a layout can be accomplished in two ways; the first uses a commercially available CAD software program to manually perform the layout from a hard copy of the schematic or netlist. The second uses advanced software tools, such as Cadence or HP/EEsof, to transfer the schematic to layout in real time. The output from these programs can then be modified to comply with foundry design rules.

There are several layout tools available for GaAs MMICs. CALMA was one of the early programs. The GDSII format, developed by CALMA, has become an industry standard for data communication and file transfer, regardless of the actual program being used. The universal acceptance of the GDSII format and its ability to handle up to 63 layers makes it the preferred method of data communication and design file transfer.

Mentor Graphics and Cadence packages offering complete MMIC design capability, including simulation, optimization, and layout, are two of the most commonly used software layout tools in the GaAs MMIC industry.

Tools specializing in design layout verification compare the actual chip layout to the circuit schematic and physical design rules and then provide specific error reports with defined locations. Some packages can also provide suggested corrective actions. One of the most common of these tools is ECAD Dracula® Integrated Circuit Layout Verification System, which provides layout-vs-schematic comparison and design rule checks.

Design rule checks are performed by comparing the geometric spacings of the MMIC layout against predefined physical design rules. These rules depend on the technology and processing capabilities of the foundry and are therefore generated by individual foundries; they are applicable to the foundry's processes only. Advanced trapezoidal approaches along with the introduction of electrical node determination and multilevel conjunctive rule capability can eliminate false errors.

In general, design rules take into account orientation effects, device spacing limitations, and probe and pad placement, among other parameters and physical restrictions. The relative placement of elements such as FETs and diodes at the die and wafer levels can have an impact on the performance of these devices, while most passive components are not sensitive to this effect. For power transistors, the spacing between the gate fingers can have a direct impact on the channel temperature and the overall performance, reliability, and stability of the device. These effects are considered in the design-rule-check stage and should be taken into account during the initial design and layout of MMIC.

IV. Typical Design Methodology

In the competitive marketplace, cost reduction at all stages of design, fabrication, and test is of prime importance. The use of CAD simulation and layout tools plays a pivotal role in first-time success and yield of a MMIC design.

Designing a MMIC involves two critical stages: performance specification, and circuit design and simulation. Other functions such as fabrication and test must also be considered during the design stages to arrive at a manufacturable product with high yield and the desired performance.

The definition of performance specifications must take into account customer needs, the technology, and the processing capabilities. Detailed understanding of the system requirements is necessary to arrive at design parameters applicable to the technology desired by the customer. Attempts to force-fit a design into an unsuitable process or to require performance parameters that cannot be supported by the process can

create substantial cost and schedule delays, low yield, and a product of suspect reliability. In typical applications, the following must be addressed:

(1) Translation of customer requirements to design instructions.
(2) Suitability of technology and process to design requirements.
(3) Availability of existing designs.
(4) Cost, schedule, and performance trade-off.

The translation of the customer's requirements into the manufacturer's design instructions is another critical point in the process. The MMIC design approach must take into account the available CAD software tools and identify the relevant cell libraries and applicable models. The choice of software tools depends on the need for linear or nonlinear simulation, time or frequency simulation, EM simulation, and the desired use of schematic capture for layout.

In the early stages of the MMIC design, the customer must examine the available cell library devices and components to determine their applicability in meeting the performance requirements. Usually, multiple iteration of a design is a common practice and helpful in identifying, through simulation and optimization, the best possible performance characteristics. Conducting a yield and sensitivity analyses as part of this iterative process will enhance the probability of first-time success and ensure an acceptable yield.

MMIC layout is also an important process that makes use of design-rule checks against circuit schematics to arrive at the final layout. A layout review prior to pattern generation is a normal practice.

Design reviews should be an integral part of the overall MMIC design process and should occur at various steps during that process. An initial design review normally examines the performance requirements, the choice of technology, and the identified cell libraries. A preliminary design review examines the suitability of the initial design in meeting the requirements and includes the circuit simulation results. A critical design review is normally used to finalize the chip design and identify the layout approach. A final design review addresses the complete design, simulation results, chip layout, and manufacturing procedures. These reviews may be conducted internally at the foundry or may include customer participation, which is highly desirable at least in the initial and final design review stages.

V. Design for Reliability and Manufacturability

Several factors may have a direct or an indirect impact on circuit yield and reliability. Device parameter variations as a result of process limitations or level of control, raw material variations, and EM proximity effects all play roles in determining overall circuit reliability and yield. The approach for achieving a reliable design should take into account the following:

(1) Definition of realistic performance requirements.
(2) Documentation of design methodology.
(3) Material and processing characterization and variation.

(4) Understanding of potential failure mechanisms.

(5) Use of adequate simulation and test tools.

The definition of realistic and achievable performance requirements is probably the most important initial step in MMIC design. Pushing the design-performance boundaries may result in selection of devices that fall at the edges of the normal Gaussian distribution. This will result in low yield and may have an impact on the reliability of the selected components. Therefore, design-performance requirements should fit very comfortably into the high-yield and assured-performance window of a MMIC foundry's process.

A documented design methodology can provide a clear path for device design, simulation, layout, and fabrication. This approach will also allow a smoother design implementation and identification of unacceptable design limits or points of possible yield loss. Figure 6-1 shows a typical design flow and the various necessary inputs.

The characterization of the processes and materials used throughout the fabrication cycle is also connected to an understanding of the common failure mechanisms and other reliability aspects of device and circuit design. As an example, gate length and placement-variation effects have been shown to be the dominant factors in limiting the yield of semiconductor devices. To meet performance requirements at higher operating frequencies, shorter gate lengths are normally required. However, this results in smaller gate-metal-to-semiconductor contact area, which is more critical from the yield and reliability aspect. Another parameter of importance is the chip length-to-width aspect ratio; it should be kept as close to 1:1 as possible to increase overall yield and reduce chip breakage during the dice and sort operations. A maximum chip length-to-width ratio of 3:1 is the normally recommended ratio for MMICs.

Layout design rules, derived from empirical and process-variation data can be very valuable in increasing the yield and reliability of a MMIC design. Considerations for circuit element placement, sizing, process variations, and physical tolerances play an important role in determining the reliability, yield, and final cost of the product. Yield analysis techniques are normally practiced to determine the overall yield and identify areas of poor yield performance. Additionally, sensitivity analysis techniques are commonly employed to determine a design's sensitivity to variations in bias point, device process parameters, tolerances, and thermal conditions.

In-process and on-wafer testing of MMIC components can provide valuable information on the performance and yield of the final product. Comparison of these data with those of the process can indicate the manufacturability of the design for a particular foundry.

Additional Reading

M. Gillick and I.D. Robertson, "Passive Components," *MMIC Design*, I. D. Robertson, Editor, The Institution of Electrical Engineers, United Kingdom, 1995.

Ladbooke, P. H., *MMIC Design: GaAs FETs and HEMTs*, Artech House, Boston, 1989.

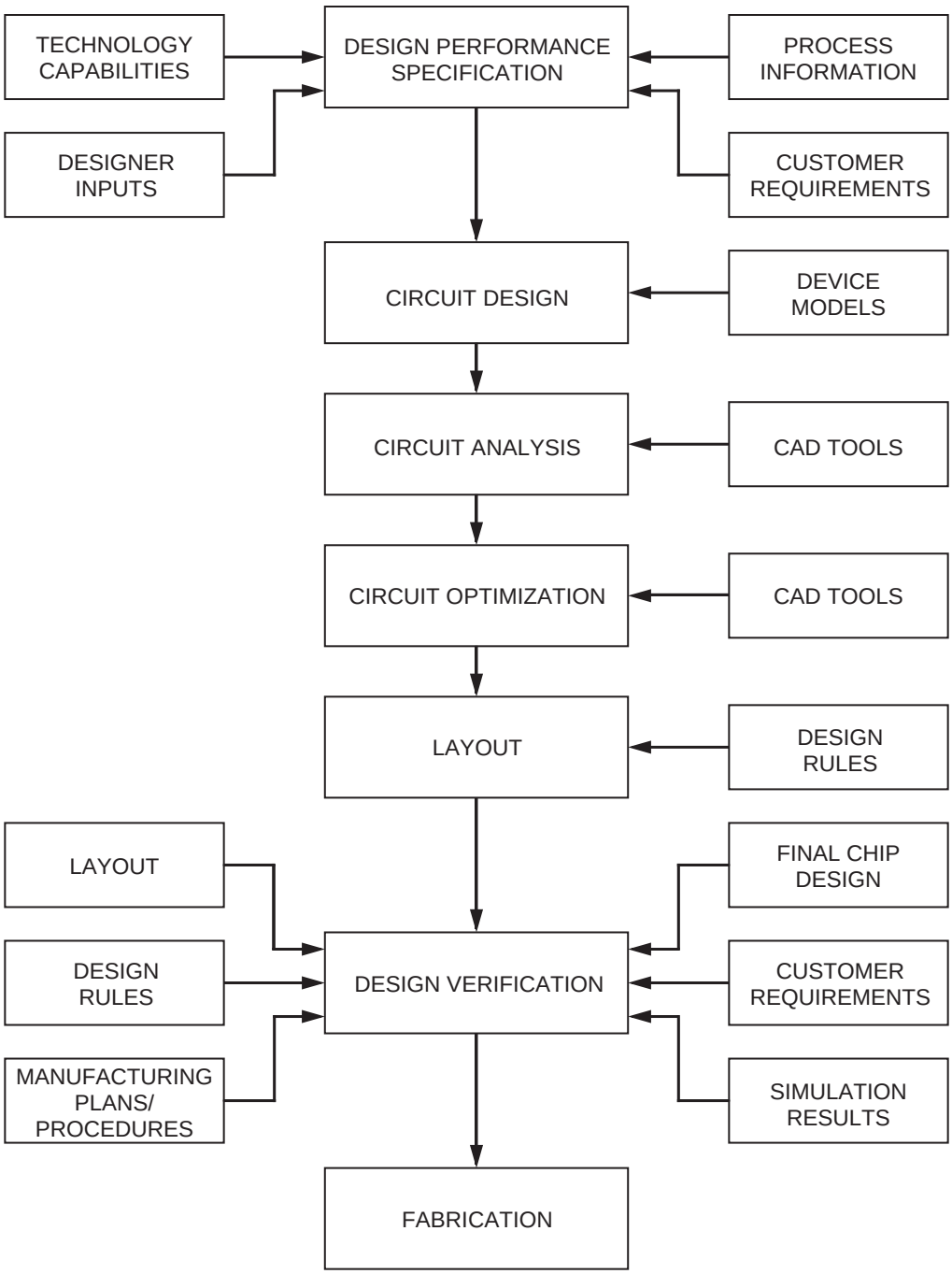

Figure 6-1. Typical design flow.

Chapter 7. Testability and Test Structures

R. Shaw

A major concern in the use of advanced technology in a high-reliability application is the quality and reliability of the product. Generally, the user's confidence that the product will meet an expected level of reliability and quality is based on documented data supplied by the manufacturer. To give the data significance, the manufacturer of the product to be validated for use in a space or other high-reliability system must have the user's expectations defined in a measurable set of values. These expectations and performance values for reliability and quality assurance are often incorporated in a "product specification." A typical product specification identifies electrical and environmental screening tests and visual inspection requirements to be performed on the product at various steps in the manufacture of the unit. These tests and visual inspection requirements are not usually done as part of the product fabrication process, but are rather a verification of conformance after the process has been completed. As such, they require additional handling and may impose stress to the part. Also, it is possible that deficiencies affecting the quality and long-term reliability of the product could be due to the design limits of the technology or the validity of the fabrication process control limits. These deficiencies are usually not found by performing a qualification based on the screening tests and inspections associated with the "product specification."

The qualification methodology proposed in this document includes process and product qualifications to help ensure that the technology, the fabrication, and the MMIC performance meet the expected level of quality and reliability. The MMIC process qualification is an evaluation of the technology's ability to attain certain defined reliability and performance levels using the manufacturer's documented processes. The process qualification usually identifies the design limits and the process deviation limits accepted by the manufacturer. The MMIC product qualification is a validation of the circuit design to perform to a minimally defined electrical performance under stress and environmental conditions.

Elements that allow the measurement of parameters that document the evaluation of the technology and validate the fabrication process are known as test structures. The definition of these structures and when, where, and how to measure these parameters are the testability criteria.

I. Test Structures

Test structures similar to those described in the following subsections are usually employed by the manufacturer. It is important that these test structures be understood by the user for their value in validating the level of expected quality and reliability of the MMIC. The usual documents that describe these structures include design files, simulation results, rationale for the proposed design, and other documents.

A. Technology Characterization Vehicle

The technology characterization vehicle (TCV) is a structure that can be used to characterize a technology's susceptibility to intrinsic reliability failure mechanisms such

as electromigration, interlayer dielectric integrity, and metal diffusion. The TCV can be composed of basic active and passive elements from the cell library including specific elements relevant to the technology. Typical elements contained in the TCV are (1) the basic FET; (2) diodes; (3) usually capacitors of two values, the lowest value in the design and a value between 1 pF and 10 pF; (4) inductors, usually the highest value; (5) resistors, all types in the design (e.g., implanted and metallic); (6) air bridges; (7) via holes; and (8) calibration elements, such as open and short circuits. Parametric Monitor (PM) elements may be included within the TCV or be the actual TCV. The TCV is usually designed close to or at the limit of the design rules and the test structures should verify all relevant material, process, and various fixed-cell dc parameters such as

interface properties	ohmic contacts
sheet resistance	implantation
etching	via hole
side/back gating	air bridge
interconnection layer	leakage currents
discrete capacitors	mask alignment
wire and die bonding	

In general, TCV certification includes subjection of a sufficient number of test structures for each wear-out mechanism to an accelerated life test to produce an estimate of the mean-time-to-failure (MTTF) and a distribution of the failure times. The MTTF and failure distribution are then used to predict a worst-case failure rate or worst-case operating lifetime at the normal operating conditions. The accelerated life tests are usually performed on packaged TCV structures. The TCV is packaged using the same packaging material and assembly procedures as those used for standard circuits in the technology. An example TCV is shown in Figure 7-1. The acceptable certification of the TCV is an integral part of the MMIC process reliability evaluation.

B. Standard Evaluation Circuits

Typically, the manufacturer has a standard evaluation circuit (SEC) used to demonstrate fabrication process reliability for the technology to be validated. The SEC design documentation usually includes the design methodology, the software tools used in the design, simulation of design performance, the design function, and the fabricated size in terms of the utilized transistor or gate count. The documentation for the SEC can be the same as that for a production circuit. The SEC can be designed solely for its role as a quality and reliability monitoring device, or it can be a production or subset of a production circuit. Typical SECs include low-noise-amplifier and high-power-amplifier production level circuits. The SEC can exercise the worst-case design rules, or, if it is a standard product, the normal design-rule limits allowed in the manufacturer's design guideline. The complexity of the SEC should be at least one half of the number of transistors or gate count of the largest MMIC to be fabricated. Usually the SEC is dc life tested if it is a small signal device process and dc and RF life tested if it is a power device process. Generally, temperature-accelerated life testing is used as the aging test. As with the TCV, the accelerated life tests are usually performed on packaged structures. The SEC is packaged using the same packaging material and assembly procedures as those used for standard circuits in the technology.

Unlike the TCV, which includes elements only from the design cell library, the SEC is a circuit device or an actual circuit that can be used as an indicator of the process stability through microwave parameter measurements. The parameters measured on the SEC are usually implemented in a data base to establish comparisons from wafer to wafer

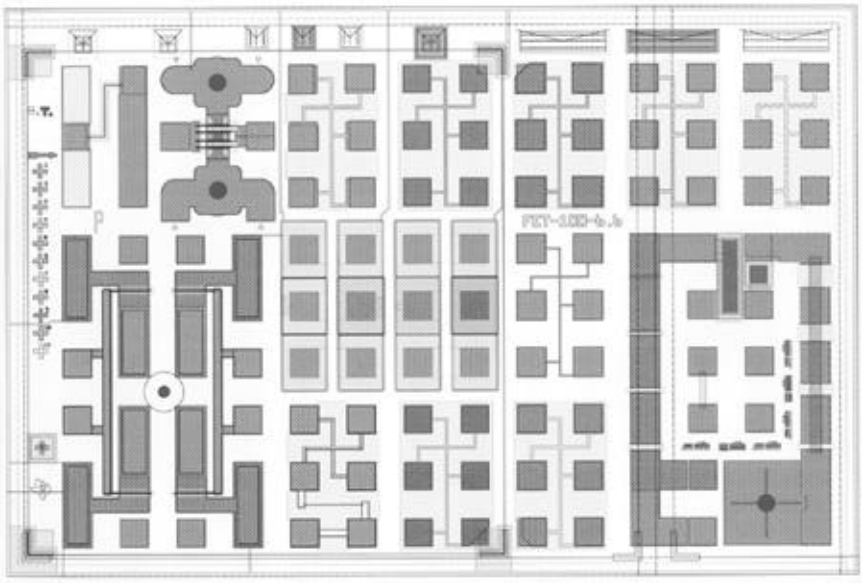

Figure 7-1. TCV example. (Courtesy of Texas Instruments.)

and lot to lot. Acceptable certification of the SEC is an integral part of the MMIC process reliability evaluation.

C. Parametric Monitors

Parametric monitors (PMs) are used as a means of measuring electrical characteristics of each wafer in a specified technology. The PM test structure can be implemented in one of several ways: incorporated into the grid, located within a device chip, designed as a dedicated drop-in die, or any combination of these. Usually several areas on each wafer are reserved for PMs. The location of the PM structures should allow the determination of uniformity across the wafer. An example of the PM locations across a wafer is shown in Figure 7-2. The manufacturer usually documents these PM locations for the user. The manufacturer's documentation could establish reject limits, record retention for wafer-to-wafer variation, and describe measurement procedures, critical parameters used in process control, and how routinely these parameters are scrutinized for statistical process control (SPC) analysis. Table 7-1 gives examples of various PM test structures and the parameter to be monitored by the structure.

Technology characterization vehicles, standard evaluation circuits, and parametric monitors required for design verification, process reliability, and performance information should be described in the Quality Management Plan. Typical information that should be addressed in the Quality Management Plan on test structures is shown in Table 7-2.

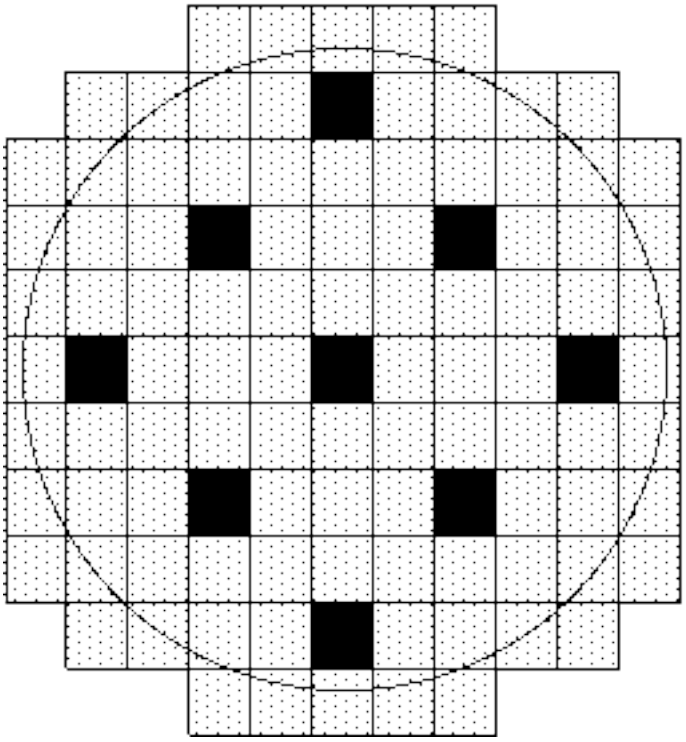

Figure 7-2. Example of parametric-monitor locations across the wafer.

Table 7-1. Common parametric monitors.

Structure	Monitored Parameter
Van der Paaw Cross	Sheet resistance
	Line width
MIM capacitor	Capacitance
	Leakage current
Transmission line structure	Contact resistance
	Sheet resistance
	Transfer length
	Saturation current
Isolation gap	Isolation gap breakdown voltage
Air-bridge chain	Air-bridge defect density
FET	Diode characteristics
RF-probeable FET	Dc characteristics
	S-parameters
	Extracted small-signal model
Via-hole alignment structure	Misalignment resistance

Table 7-2. Typical test structure information.

Test Structure	Typical Information
TCV	Failure mechanism identification. Failure mechanism method of test. Suitability of TCV to identify failure mechanism. Method of TCV implementation. TCV design parameters and applicability. Test conditions and parameters.
SEC	SEC determination methodology. SEC suitability to desired product function and design. SEC test conditions and parameters. Level of available reliability information.
PM	Description of the PM. Function of the PM. Location of the PM and method of determination. Suitability to the technology and design. Test conditions and parameters. Method of test.

II. Testability

Testability is the ability to measure defined parameters associated with the MMIC fabrication process. These measurements can be used to validate the MMIC design, to evaluate the technology in terms of reliability, as a catalog for statistical process control (SPC) to indicate the process stability, and as device/wafer acceptance data records. The manufacturer and user should be in agreement as to the definition of the test structures and the when, where, and how of parameter measurement. The following subparagraphs give information on the levels of testability and typical parameters measured on test structures and MMIC devices.

A. Wafer-Level Testability

Usually dc on-wafer testing is performed on the test structures and MMICs with traceability and wafer mapping included in the measurement documentation. These tests are usually done with an autoprobing instrument. The dc autoprobe is used to test individual FETs, resistors, and capacitors on the circuit for dc characteristics and functionality. The tests are a key screening procedure and are used to ensure that all parts delivered have fully dc functional FETs, resistors and capacitors, and that the parameters for these components are within standard or desired specifications. The FET gate functionality (pinch-off) test is particularly important since it ensures gate functionality across the entire gate width. Usually, dc probing is performed after frontside processing and before backside processing. The lack of "via" connections to a backside ground plane allows for better dc isolation of circuit components. Measurements are taken on several parameters, such as transconductance (g_m), breakdown voltages (V_b), and pinch-off voltage (V_p), to address yield analysis.

Dc wafer probing can also be used at several steps during fabrication: after ohmic contact formation, after gate formation, after interconnect/air-bridge metal deposition, and after final frontside and backside processing are complete. This in-process

monitoring can help catch problems early in the fabrication of the wafer so that further processing of out-of-tolerance wafers can be avoided. This minimizes cost and helps to identify process control problems as early as possible.

Some manufacturers have the capability to perform on-wafer RF measurements for SECs and MMICs. Usually this probing is done by connecting RF measurement equipment such as network analyzers to specialized probes for injecting and detecting the microwave frequency signal through the device. RF measurements are made after both frontside and backside processing is complete but prior to chip separation and normally under the same bias conditions used in the application of the device. The technique generally used in RF probing is to place pads or "footprints" for signal connection at each RF port. The footprints are typically coplanar waveguide topologies and can be designed to be ground-signal (G-S) or ground-signal-ground (G-S-G) configuration depending on the type of probe to be used. On-wafer testing of high power devices can cause a thermal problem because of the relatively high currents that must be provided for bias and the usually poor thermal environment of the wafer probe system. This problem can be minimized by pulsing the dc bias to the device, but it is still a difficult measurement. Also, testing high-frequency devices on-wafer can cause an instability problem at lower frequencies because of the high forward gain of the transistors. Often the probe equipment and data logging equipment cannot detect this oscillation condition, and measurement errors or damaged parts may result. Table 7-3 lists several dc and RF parameters that are typically measured during on-wafer probing.

Table 7-3. Examples of dc and RF autoprobe test parameters.

DC Parameters
Saturated drain current (I_{DSS})
Transconductance (g_m)
Pinch-off voltage (V_p)
Gate-source breakdown voltage (V_{BGS})
Gate-drain breakdown voltage (V_{BGD})
Gate leakage current (I_{lk})
Capacitance (pF)
Resistance (Ω)

RF Parameters
Scattering parameters (S_{nn})
Associated gain (G_a)
Noise figure (NF)
Small-signal gain (SSG)

B. MMIC-Level Tests

Once the wafer has been tested and accepted, it is diced, sawed, or scribed, and the good dies are identified for MMIC chip delivery or evaluation test. At this point in most certification, acceptance, and/or qualification programs, several samples are further tested for individual MMIC reliability or design validation. Usually the test structure or MMIC is packaged using the same packaging material and assembly procedures as standard circuits in the technology. The packaged circuit is then dc and RF tested for validation to the performance level of the design limits. The MMIC validation program should validate, in terms of electrical performance and reliability, all the MMIC functions assessed at the technology evaluation and design review. The electrical performance testing done at this time is usually more detailed than the wafer-level testing described in the previous section. The device can be rigorously tested for design validation since it is packaged and precautions for oscillations and thermal issues can be more easily handled.

The following list gives examples of the RF tests that are typically performed on the MMIC:

Scattering parameters (S_{nn})	Power added efficiency
Associated gain (G_a)	Isolation
Noise figure (NF)	Switching time
Small-signal gain (SSG)	Intermodulation (distortion)
Output power at 1-dB gain compression	Phase linearity

The information collected from test structures can be a very valuable tool in understanding the overall stability and long-term reliability of the product. Also, data collected from test structures at various processing and manufacturing steps can be excellent indicators of the quality of the manufactured lot. The use of test structures and planning for testability are very common practices in the industry and should be used to their fullest potentials in developing the MMIC qualification plan.

Additional Reading

Belfort, M., J. F. Dreyfuss, P. Burgaud, J. C. Le Gouable, J. P. Fortea, P. G. Tizien, J. M. Dumas, and G. Kervarrec, *A Methodology for the Space Qualification of GaAs MMICs*, CNET, Lannion Cedex, France, December 1992.

Guidelines for GaAs MMIC and FET Life Testing, JEP118 JEDEC Publication, Electronic Industries Association, Washington, DC.

Chapter 8. Qualification Methodologies

S. Kayali, G. E. Ponchak, and R. Shaw

I. Introduction

This chapter outlines a recommended procedure for the design, manufacture, and acceptance of space qualified MMICs. First and foremost, the reader must understand that although the methodologies recommended in this chapter may appear rigid and specific, they should not be viewed as such. In fact, it is the authors' intention that the qualification methodology not only permit but rather require the manufacturer and customer to determine many of the details. Instead of presenting specifications for reliability, this chapter presents the questions a MMIC user should ask of the manufacturer to assure a reasonable level of reliability, and at the same time it tries to present to the MMIC manufacturer the methodologies that have been accepted and practiced by some members of the industry in the hope that a standard qualification procedure may develop. This chapter, like the previous chapters, is also an educational guide. Furthermore, it should be used with the other chapters: The details of this qualification methodology depend on the type of circuit being fabricated and the devices incorporated into the circuit, along with the reliability concerns and failure mechanisms (Chapters 3 and 4), the testability of the circuit (Chapter 7), and the effect the package has on the MMIC reliability (Chapter 9).

The rationale for not publishing a strict qualification standard is derived from the fact that the GaAs industry is rapidly evolving, and, therefore, it would not be prudent to set limits on that evolution. In addition, it is not possible to guess the needs of every system being planned or the reliability requirements of every system. For example, MMIC users may request a relaxation of the recommended qualification methodology to lower the part cost, if the mission has a short expected lifetime or if the total satellite cost is small. Alternatively, very expensive satellites with a long projected lifetime will normally be qualified to a higher standard than even that recommended in this guide. The important point is that whenever reliability qualification is relaxed, either through the deletion of some tests, or screens, or a reduction in the number of parts tested, up-front MMIC costs are lowered at the price of increased risk of system failure.

A four-step procedure followed by most satellite manufacturers includes some practices recommended by the Qualified Manufacturers Listing (QML) programs [1] with screening procedures from more traditional qualification methodologies ; that procedure is recommended in this guide. The steps are (1) Company Certification, (2) Process Qualification, (3) Product Qualification, and (4) Product Acceptance, as summarized in Figure 8-1. Company Certification outlines the procedures and management controls the manufacturer should have in place to assure the quality of its MMICs. Process Qualification outlines a procedure the manufacturer should follow to assure the quality, uniformity, and reproducibility of MMICs from a specific fabrication process. Product Qualification encompasses a set of simulations and measurements to establish the electrical, thermal, and reliability characteristics of a particular circuit design. Lastly, Product Acceptance is a series of tests or screens performed on the deliverable that is normally practiced by GaAs MMIC manufacturers and their customers to satisfy high-reliability program requirements and provide specific reliability and qualification information pertinent to that particular product.

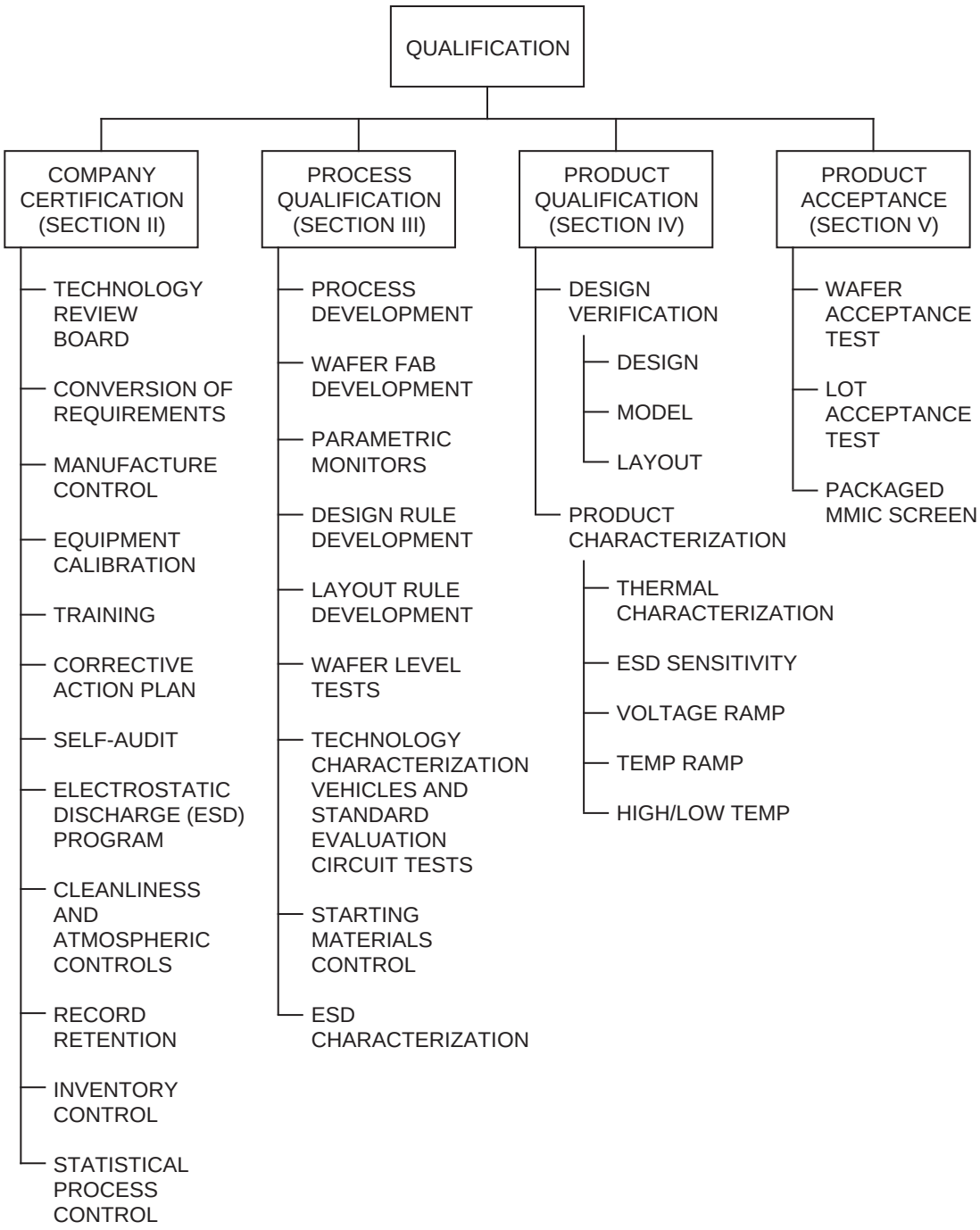

Figure 8-1. Recommended qualification methodology.

Before these four steps are presented in detail, a few important aspects of MMIC qualification must be discussed. First, although the manufacturer is ultimately responsible for delivering a reliable MMIC, the reliability of the total system rests with the MMIC user. Therefore, it is within both parties interests to understand the expected electrical performance requirements and operating environment of not just the MMIC, but the system itself. While this helps the manufacturer select the best technology for the MMIC and deliver a more reliable part, it requires the MMIC user to share information with the manufacturer. Furthermore, although the organization of the qualification

methodology is representative of what MMIC manufacturers and users currently use, the content of the qualification process is the essential ingredient. The MMIC user should not discount a manufacturer's proposal because the manufacturer does not organize its procedures in the same way or use the same terms and phrases offered in this chapter.

II. Company Certification

Procurement of MMICs is often the result of a long-term partnership between the customer and the manufacturer in which both parties add knowledge and experience to the process to assure reliability of the final circuits and satisfaction of the required performance specifications. This close, working relationship evolves after mutual trust is established. If the parties have never worked together, the MMIC user can still gain the necessary confidence in the manufacturer if the manufacturer can show that it has documentation, procedures, and management practices that control the facilities, equipment, design processes, fabrication processes, and personnel. These items are typically part of an overall Quality Management Program and outlined in a Quality Management Plan. This step of the qualification process is often referred to as "company certification" and is usually verified by the MMIC user through either a written or facility audit. It is recommended that the audit and company certification be completed before a contract for the purchase or development of an MMIC is established. The MMIC user may even consider this the first and most important criterion in selecting a company from which to buy parts. A company that cannot demonstrate a formal structure to address the issues of quality and reliability should not be used as a supplier of MMICs for high-reliability or space applications.

Since most of the information sought during company certification is based on established QML programs [1] and standard industry methodologies, the audit should be easy and inexpensive for both the user and manufacturer. In fact, most of the data sought in the audit should be compiled and available for distribution by the manufacturer. Furthermore, if the manufacturer has passed previous audits, either for other MMIC procurements or ISO 9000 certification, this step in the qualification process may be reduced to a simple updating of past audits, or eliminated entirely.

A simplified version of the audit is shown in Figure 8-2. The audit for a specific MMIC must be developed on a case-by-case basis. The major items in the Quality Management Program are presented in the rest of this section, but it must be remembered that this is only a partial list. As stated before, company certification is the first opportunity a MMIC user has to determine the credibility of a manufacturer's reliability program. This credibility should be established before a contract has been signed. Beyond the following list, the inclusion of additional items in the company certification procedure that are specific to the user's needs would be expected.

A. Technology Review Board

To assure the quality and reliability of MMICs, manufacturers will typically have a permanent committee or board in place with knowledge of the entire MMIC fabrication process and the authority to change the process if the quality of the parts is not maintained. This board is commonly called the Technology Review Board (TRB) from the QML program [1]. The TRB is responsible for

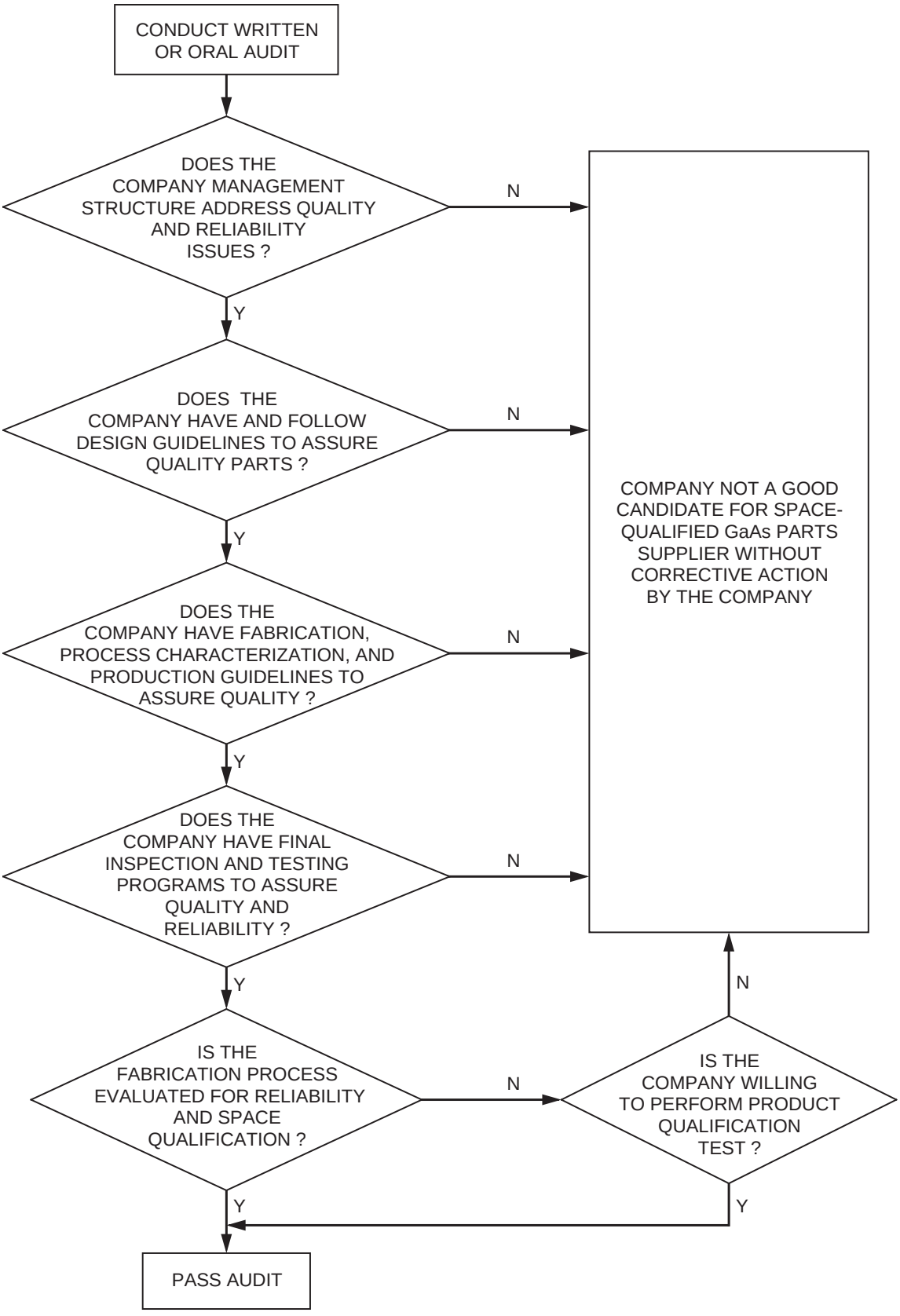

Figure 8-2. Reliability audit.

(1) The development, implementation, and documentation of the manufacturer's Quality Management Program and Quality Management Plan.

(2) The development, implementation, and documentation of the manufacturer's Process Qualification, Product Qualification, and Product Acceptance plans.

(3) Compiling and maintaining all records of the fabrication process, statistical process control (SPC) procedures, SPC data, certification and qualification processes, reliability data analysis, and corrective actions taken to remedy reliability problems.

(4) Examining standard evaluation circuits (SECs) and MMIC reliability data and establishing and implementing corrective actions when the reliability of the circuits decreases.

(5) Notifying customers when the reliability of a wafer lot is questioned and supplying the customers an evaluation of the problem and any corrective actions required.

(6) Supplying reliability data to customers.

Because of these great responsibilities that cover a broad area of knowledge, the members of the TRB should have good hands-on knowledge of device design, technology development, wafer fabrication, assembly, testing, and quality-assurance procedures. The members of the TRB are normally from the manufacturing company, but a customer requesting custom products may request a seat on the board for those products only.

B. Conversion of Customer Requirements

Not all customers express their specifications in the same way, and not all manufacturers publish MMIC performance specifications and operating guidelines in the same way. For example, a user will not normally specify the type of transistor, substrate thickness, or transmission lines they want in the fabrication of a circuit. Instead, they simply ask for an amplifier with 15 dB of gain and a maximum output power of 1 W at 10 GHz. For the MMIC manufacturer, these performance specifications are the starting point in determining the type of transistor, substrate, and transmission lines, among other things, required. Only after conversion from the customer's specifications to the manufacturer's specifications can the manufacturer bid on the contract and the user know what reliability questions to ask. It is recommended that the procedure by which customer requirements—as expressed, for example, in specifications and purchase orders—are converted into working instructions for the manufacturer's personnel be documented. A typical document will describe the procedures a company performs, the order in which they are performed, and the typical schedule. Some of the items commonly found in such a conversion are

(1) Relating customer circuit requirements to manufacturer circuit requirements.

(2) Converting circuit requirements to a circuit design, using controlled design procedures and tools (i.e., established geometric, electrical, and reliability design rules).

(3) Establishing a design review team.

(4) Selection of SECs and Parametric Monitors (PMs).
(5) Mask generation procedure within the controlled design procedure.
(6) Wafer-fabrication-capabilities baseline.
(7) Circuit-fabrication procedures in accordance with approved design, mask, fabrication, assembly, and test flows.
(8) Incoming inspection and supplier procurement document covering design, mask, fabrication, and assembly.
(9) Establishment of screening and traveler documents.
(10) Technology Conformance Inspection (TCI) procedures.
(11) Marking requirements.
(12) Rework procedures.

C. Manufacturing Control Procedures

MMIC manufacture is a very complicated process involving many materials and steps, all of which are critical to MMIC performance and reliability. Only a properly controlled manufacturing line can be expected to routinely produce quality MMICs. Thus, the customer should be assured that the manufacturer is using only certified processes and qualified technologies at every step in the manufacture of the MMIC—from the ordering of materials to the shipping of the MMIC. To obtain that level of assurance, the company certification audit should review the manufacturer's procedures for

(1) Traceability of all materials and products to the wafer lot.
(2) Incoming inspection to assure conformance to the material specification.
(3) Electrostatic discharge (ESD) control in handling the material in all stages of manufacturing.
(4) Conformance with design requirements at
    (a) Device procurement specification.
    (b) Simulation-model verification.
    (c) Layout verification.
    (d) Testability and fault coverage verification.
    (e) Electrical parameter performance extraction.
    (f) Archived data.
(5) Conformance of fabrication requirements at
    (a) Mask fabrication.
    (b) Mask inspection.
    (c) Wafer fabrication.
(6) Assembly and package requirements.
(7) Electrical testing.

Most of this information can be obtained if the MMIC user asks for documentation of the manufacturer's production flow.

D. Equipment Calibration and Maintenance

It would be difficult to maintain the quality of MMICs produced on equipment that is not properly maintained and calibrated. Therefore, all equipment used in the design, fabrication, and testing of the MMIC should be maintained according to the equipment manufacturer's specifications. In addition, the equipment should be calibrated on a regular basis. Documentation showing the maintenance and calibration schedule, deviations from the calibration and maintenance schedules, and any corrective action taken will normally be kept by the manufacturers. This documentation will also highlight any major discrepancies found in the calibration and maintenance of a piece of equipment since it may affect the reliability of the MMICs. The TRB will review this document to determine if any corrective action is required. Further information on equipment calibration and maintenance documentation can be found in [2].

E. Training Programs

Even well maintained and calibrated equipment cannot produce quality MMICs without skilled operators. To assure the skills of the personnel employed in the design, fabrication, and testing of the MMICs, each engineer, scientist, and technician should have formal training relative to their tasks. Furthermore, retesting and retraining should be provided regularly to maintain the worker's proficiency, especially if new equipment or procedures are introduced into the manufacturing process. It is therefore recommended that the work training and testing practices employed to establish, evaluate, and maintain the skills of personnel engaged in reliability-critical work be documented as to form, content, and frequency.

F. Corrective Action Program

One of the best ways to continuously improve the reliability of manufactured parts is to test and analyze failed parts—including returns—from all stages of manufacturing, and, based on the findings, make corrective actions to the manufacturing process or the education of the MMIC users. The plan that describes these corrective actions is normally documented. The corrective action plan should describe the specific steps followed by the manufacturer to correct any process that is out of control or found to be defective and the mechanism and time frame that a manufacturer will follow to notify customers of potential reliability problems.

G. Self-Audit Program

To promote continual quality improvement, manufacturers regularly review their manufacturing procedures through an internal, independent self-audit program under the direction of the TRB. The self-audit program should identify the critical review areas, their frequency of audit, and the corrective action system to be employed when deviations from requirements are found. Typical areas included in a self-audit are

(1) Calibration and preventive maintenance
(2) Fabrication procedures
(3) Training programs
(4) Electrical tests

(5) Failure analysis programs
(6) Test methods
(7) Environmental control
(8) Incoming inspection
(9) Inventory control and traceability
(10) Statistical Process Control (SPC)
(11) Record retention

The self-audit checklist, the date of the previous audits, and all findings from the audits are maintained typically by the TRB, which will use these findings to recommend corrective actions and prepare a self-audit follow-up.

H. Electrostatic Discharge Handling Program

Because of the catastrophic failure that normally follows ESD, all personnel that work with GaAs MMICs should be trained in the proper procedures for handling the devices. Furthermore, these procedures should be documented and available for reference. Typically, the procedures include the methods, equipment, and materials used in the handling, packaging, and testing of the MMICs. Further guidance for device handling is available in the Electronics Industry Association (EIA) JEDEC Publication EIA 625 [3] and MIL-STD-1686 [4].

I. Cleanliness and Atmospheric Controls

The quality of GaAs MMICs and the yield of the fabrication line is directly linked to the manufacturer's control over the cleanliness of the environment in which the parts are fabricated. Therefore, manufacturers often spend a large amount of their resources to assure that the MMICs are fabricated in ultraclean rooms where the atmosphere is tightly controlled. Since the yield of the fabrication process is so strongly dependent on the success of maintaining those conditions, regular measurements are taken to assure the temperature, humidity, and cleanliness of the fabrication areas. In addition, during transit and storage prior to seal, the die/wafer should be protected from human contact, machine overspray, or other sources of contamination. All of these procedures and measurements are recorded and compiled into a single document by the clean-room manager or alternate for future reference.

J. Record Retention

Documentation is the only method to gauge the reliability of MMICs fabricated today vs those produced last week or last year and to correlate changes in the reliability to variations in the processing steps. Although many sections in this guide recommend the documentation of certain data or procedures, it is helpful if a list of documents and the period of retention for each document is made. Furthermore, the list should contain a record of when each document was last changed, who is responsible for maintaining the document, and where the document is stored. The typical documents to be retained are relevant to

(1) Inspection operations (i.e. production processes, screening, qualification).

(2) Failure and defect reports and analyses.

(3) Initial documentation and subsequent changes in design, materials, or processing.

(4) Equipment calibration.

(5) Process, utility, and material controls.

(6) Product lot identification.

(7) Product traceability.

(8) Self-audit report.

(9) Personnel training and testing.

(10) TRB meeting minutes.

K. Inventory Control

The proper inventory of all incoming materials and outgoing parts is not only required for the management of a profitable company but also for the manufacture of reliable MMICs. Many materials and chemicals used in the fabrication of MMICs have shelf lives that must be adhered to if process yield and reliability are to be maintained. The tracking of in-process and completed MMICs is essential for the establishment of MMIC history, which is critical if failure analysis is ever necessary. Therefore, the methods and procedures used to control the inventory of all materials related to the MMIC manufacturing process should be documented. Typically documented inventory control procedures include

(1) Incoming inspection requirements and reports.

(2) Identification and segregation of non-conforming materials.

(3) Identification and control of limited-life materials.

(4) Control of raw materials.

(5) Data retention for required receiving reports, test reports, certification, etc.

(6) Supplier certification plan.

L. Statistical Process Control

The establishment of a statistical baseline for judging the continuous improvement of a manufacturer's processes is important. To establish that baseline, the manufacturer should develop an SPC program using in-process monitoring techniques to control the key processing steps that affect device yield and reliability. As part of the SPC process, every wafer lot typically has built-in control monitors from which data are gathered. The resulting data should be analyzed by appropriate SPC methods to determine the effectiveness of the company's continuous improvement plans. Additional information on SPC analysis can be found in the Electronics Industry Association JEDEC EIA 557A [5] and in MIL-I-38535 [1].

III. Process Qualification

A manufacturer who has standardized production around a single technology will often qualify the entire production line. In doing so, the manufacturer attempts to demonstrate that the entire process of designing and fabricating an MMIC using the stated technology is under its control. In addition, the manufacturer establishes an electrical performance and reliability baseline for all components fabricated using the process. This has advantages for both the manufacturer and the user of the MMIC. For the manufacturer, it saves costs and time on the fabrication of future MMICs, since the reliability and functional performance of the components constituting the MMIC have already been established. For the MMIC user, there is a certain level of comfort in buying parts from a production line with a history of supplying reliable MMICs, in addition to the reduced qualification time and therefore delivery time that should be possible.

The term usually applied to this procedure is "process qualification." Process qualification is a set of procedures a manufacturer follows to demonstrate that they have control of the entire process of designing and fabricating an MMIC using a specific process (e.g., MESFET, HBT, HEMT). It addresses all aspects of the process including the acceptance of starting materials, documentation of procedures, implementation of handling procedures, and the establishment of lifetime and failure data for devices fabricated using the process. Since the goal of process qualification is to provide assurance that a particular process is under control and known to produce reliable parts, it needs to be performed only once, although routine monitoring of the production line is standard. It is critical to remember that only the process and basic circuit components are being qualified. No reliability information is obtained for a particular MMIC design.

Although process qualification is intended to qualify a defined fabrication procedure and device family, it must be recognized that GaAs technology is constantly evolving, and this technology evolution requires the continual change of fabrication procedures. Furthermore, minor changes in the fabrication process to account for environmental variations, incoming material variations, continuous process improvement, or minor design modifications may be required. All of these changes in the process are permitted and frequently occur under the direction of the TRB. Thus, strict application of the commonly used phrase, "freezing the production process," does not apply.

The internal documents and procedures used by most manufacturers for process qualification are summarized in Figure 8-3. In addition, the QML program [1] provides guidelines for process qualification. The first step in the procedure is for the manufacturer to determine the family of devices to be fabricated and the technology that will be used in the fabrication—for example, a 0.5 μm, ion-implanted MESFET technology with Si_3N_4 MIM capacitors and NiCr resistors. Second, the manufacturer will establish a TRB to control the process qualification procedure. After all of the processing steps have been defined and documented, the workmanship, management procedures, material tracking procedures, and design procedures should be documented. The information contained in the documentation describes the process domain that is being qualified.

The qualification process also involves a series of tests designed to characterize the technology being qualified. This includes the electrical as well as the reliability characteristics of components fabricated on the line. Some of these tests are performed at wafer level and include the characterization of PMs, Technology Characterization Vehicles (TCVs), and SECs, which were all discussed in Chapter 7. Other tests require

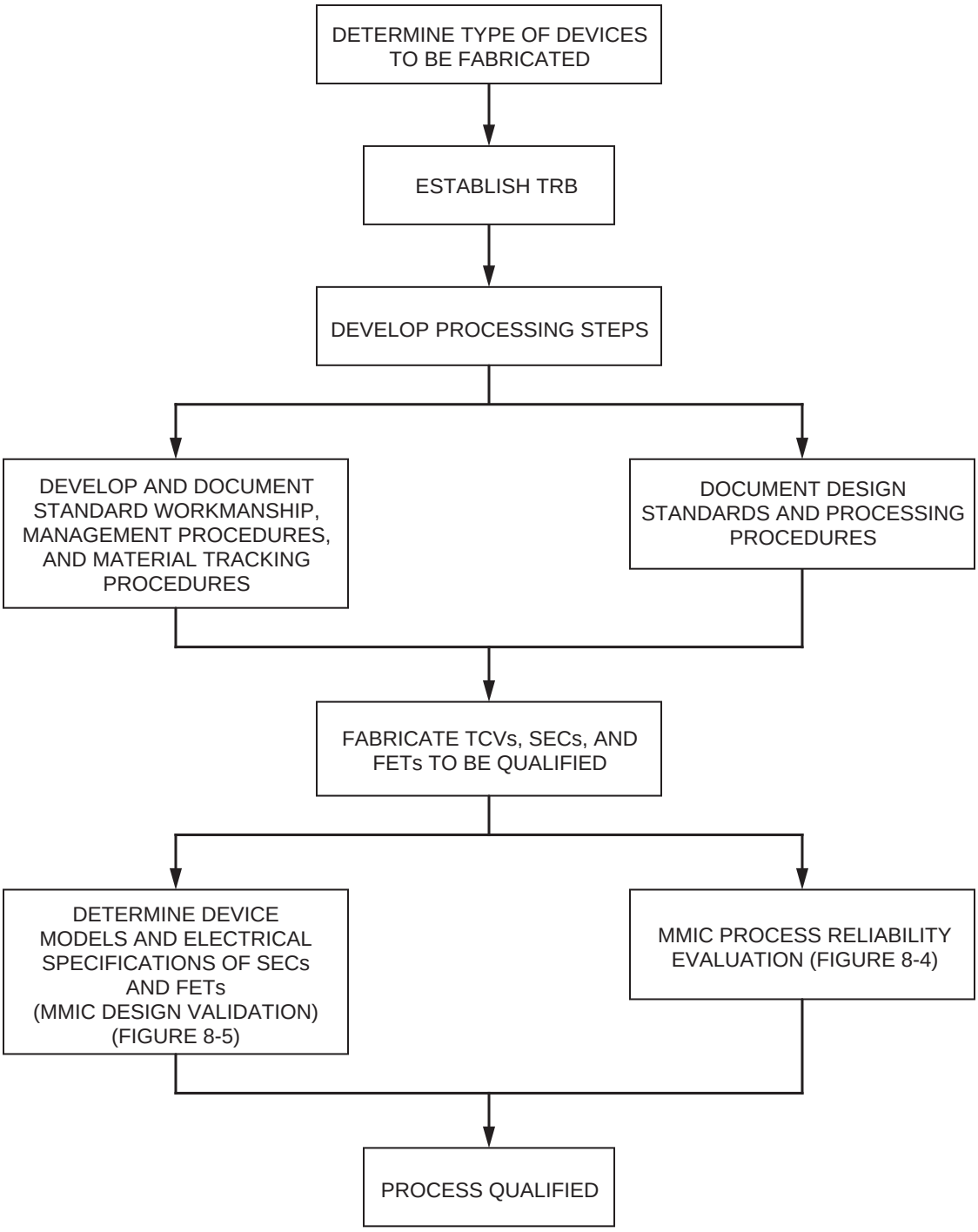

Figure 8-3. MMIC die process qualification.

the mounting of circuits or elements onto carriers. All of these tests and the applicable procedures are an integral part of the qualification program and provide valuable reliability and performance data at various stages of the manufacturing process. Figure 8-4 outlines a recommended series of tests for MMIC process reliability evaluation. The number of circuits or devices subjected to each test will normally be determined by the TRB and the rationale for their decision will become part of the process qualification documentation. In general, a higher level of confidence in the

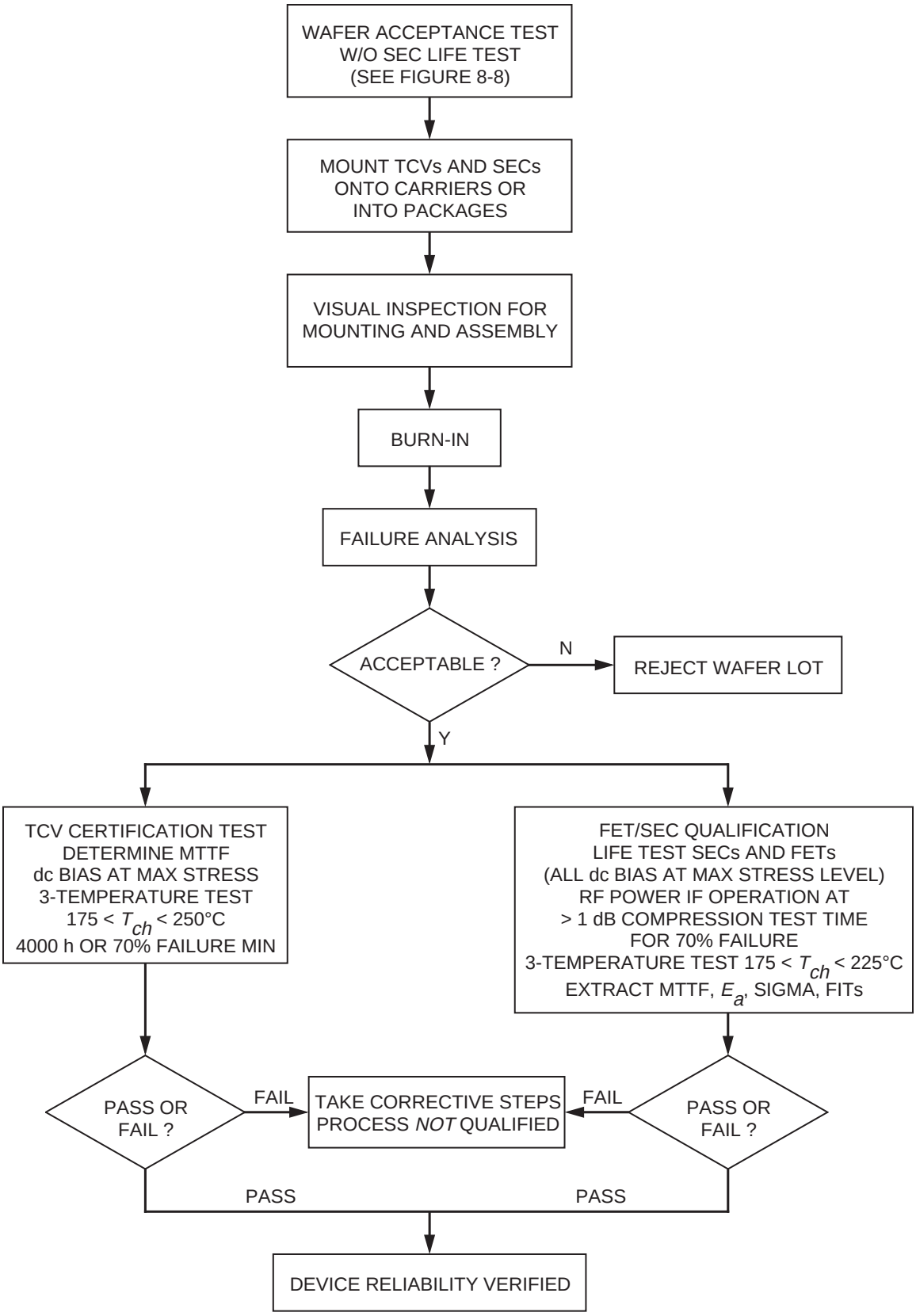

Figure 8-4. MMIC process reliability evaluation.

reliability data exists if more circuits are tested, but this is offset by the fact that after a certain level of testing, the incremental gain in confidence is minor compared to the cost of testing. Since the stability of the process is being determined as part of the process qualification, the manufacturer will typically fabricate and test components from several wafer lots. Figure 8-5 provides a series of tests that is recommended to characterize the electrical and thermal limitations of the devices or circuits. The performance limitations obtained from these tests often become the basis for limits incorporated into the design and layout rules.

Note that the process-qualification procedure is QML-like and therefore addresses topics similar to those of the company certification discussed in Section 8-1. The major difference is that company certification is performed by the customer, whereas process qualification is self-imposed by the manufacturer, often before customers are identified. Other items particular to process qualification are discussed below.

A. Process Step Development

Although all of the items described in Section 8-II are important to the process-qualification procedure, the actual process of turning a bare GaAs wafer into a MMIC by technicians in a clean room is often the only task associated with process qualification. Indeed, it may be the most critical step in the process and probably requires the most time and resources to develop. In addition, it is truly the fabrication procedures and the components fabricated on the line that distinguish one production line from another. Therefore, it follows that the first step in the process-qualification procedure is the development and documentation of the processing steps required to build a MMIC. Although all of the steps in the fabrication process, including wafer surface preparation, photolithography, active layer formation, passivation, and the metallization system and formation (Section 3-VIII), should be included in the documentation, the details are typically considered proprietary by the manufacturer. Therefore, the MMIC customer may expect to see a general list of processing steps or the process flow, but not the level of detail actually required to fabricate the parts.

B. Wafer Fabrication Documentation

Once a process is qualified, reliability concerns may still arise from minor variations in the process flow, environment, or starting materials. Therefore, all wafer fabrication steps and conditions will normally be recorded by the manufacturer in order to maintain repeatability of the product. Documentation of these steps and fabrication conditions should be maintained to trace any future quality or reliability concerns to a specific step. Although process travelers can be used to document the fabrication and manufacturing steps, they usually lack the detail necessary to trace quality or reliability problems to specific fabrication steps. The wafer fabrication steps themselves and the documentation describing them are usually considered proprietary by the manufacturer.

C. Parametric Monitors

Parametric monitors are essential for monitoring a production line's quality or continuous improvement. Parametric monitors were fully described in Chapter 7: they are mentioned in this section only to emphasize the fact that choice of the PMs is dependent on the process and technology being monitored. Therefore, this choice is a critical element in the process-qualification procedure. The complete list of PMs is

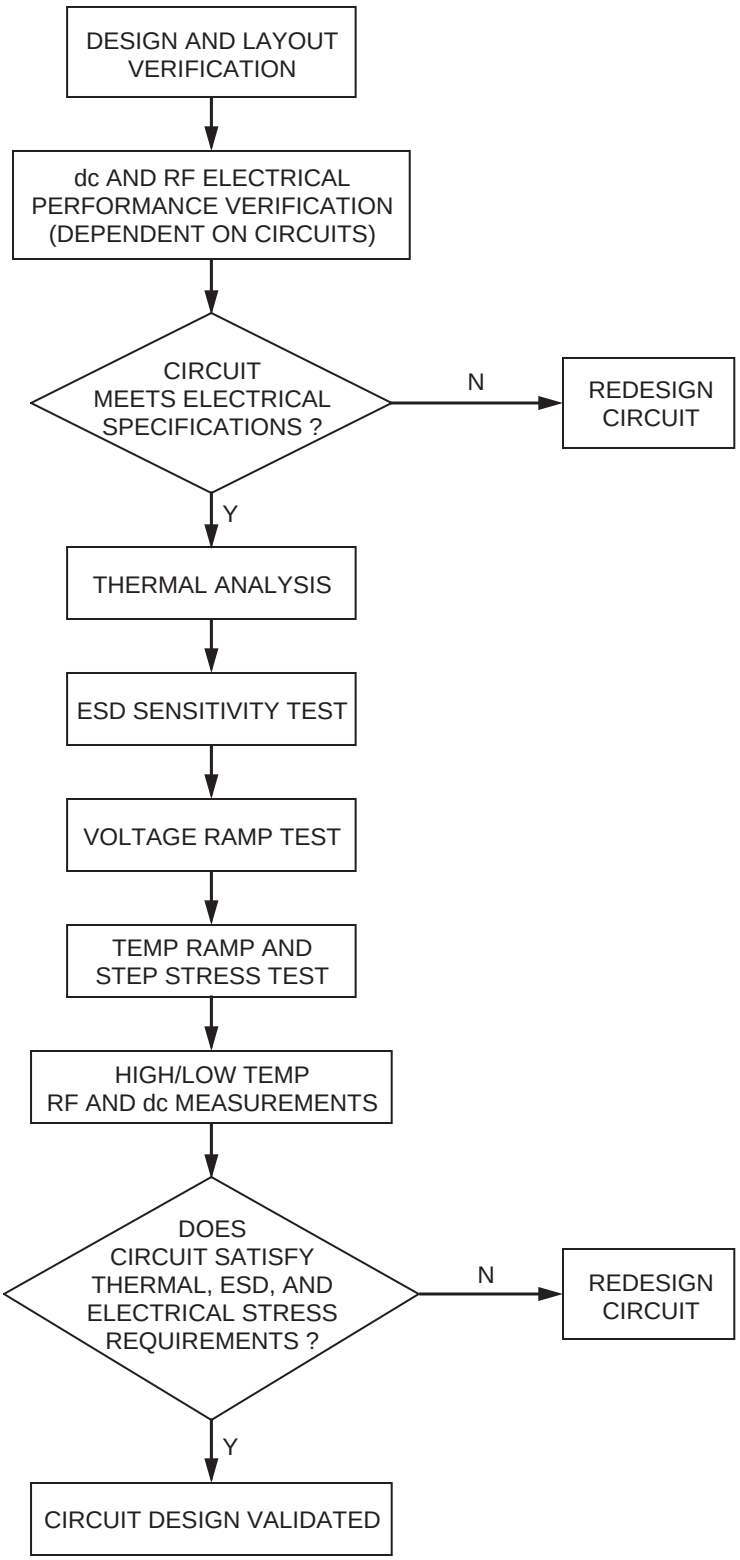

Figure 8-5. MMIC design validation.

usually combined into a single reticle that is included on all wafers. The data obtained from the PMs will normally be stored in a data base that permits the quick comparison of

each wafer fabricated on the line to all of the other wafers. This permits determination of process stability.

D. Design-Rule and Model Development

The reliability of MMICs fabricated using a qualified process will greatly depend on whether they are designed with qualified components and according to prescribed rules. In addition, the standardization of the component types also brings a certain degree of cost reduction. Therefore, part of the process-qualification procedure is to determine and document design rules that are specific to the process. Typical information included is the minimum resistor size, the maximum capacitance, the minimum air-bridge width, the maximum air-bridge length, the minimum separation between via holes, and the active device geometry. In addition to these characteristics, a list of rules relating to such issues of circuit design as the maximum power handling capability, maximum linear gain, and minimum noise figure of the devices should also be included. Finally, manufacturers will often develop standard cells or small circuits that perform specific functions, such as couplers, gain blocks, bias networks.

To fully use the standard components in circuit designs, models must be developed; although models contained in commercial software packages may be adequate, they often need to be adapted to fit the measured characteristics. Commercial software packages are available to extract the RF and dc characteristics from measurements and fit the model to the data. Once each of these components and cells is described and characterized, circuit designers can use them to increase the chance of first-time design success.

E. Layout-Rule Development

Layout rules should be followed in any circuit design to assure manufacturability and reliability. The layout rules may be specific to a particular process, and therefore, must be developed for the process being qualified.

F. Wafer-Level Tests

The GaAs industry strives to reduce production costs by shifting as much testing as possible to the earliest possible point in the production cycle—this to weed out bad wafer lots before more value has been added to them. The best strategy performs wafer-level tests that include dc and RF characterization, PM characterization data, and temperature performance. Limitations may exist in the level of test detail depending on the device design and the manufacturer's test capabilities discussed in Chapter 7. In general, wafer-level tests are performed, but they should be supplemented with other verifications, such as test fixture or in-package tests. Once agreement between the wafer level and the package-level tests has been established, the manufacturer may rely on the wafer-level tests for production monitoring.

G. TCV and SEC Tests

One of the most important steps in the process-qualification procedures is to determine the thermal, electrical, and reliability characteristics of devices fabricated within the domains of the process. This data is obtained through the characterization of

TCVs and SECs, as shown in Figures 8-4 and 8-5. Both of these test structures and the testing procedures for them were presented in Chapter 7. All data obtained from these tests should be gathered and stored by the TRB. In most cases, the success of a manufacturer in qualifying the process will depend on the data from these tests.

H. Starting Materials Control

The manufacturer should have in place a mechanism to assure the quality and characteristics of every starting material from the GaAs wafers and chemicals used in the processing steps to the shipping containers used for die/wafer transportation and storage, since they all have a direct impact on the quality and reliability of the final product. Analyses of the chemicals and gases used in processing GaAs is normally performed by the device manufacturer or the supplier of the chemicals. Traceability and documentation of the characterization results to the individual wafer process lot are essential in resolving any process variation questions or concerns. The facility audit program can be the vehicle used to determine the manufacturer's level of control.

Most GaAs device manufacturers procure the GaAs wafers from outside suppliers. Procurement requirements imposed by the device manufacturer identify the dislocation density, type of starting material, resistivity, and other characteristics that are very important to the device user. This information can help determine the suitability of the starting material to the process and the material's capabilities. The traceability and documentation of the procurement requirements and wafer characterization can be used to resolve any process variation concerns. Wafer preparation steps, such as initial surface cleaning, can also alter device characteristics and are an important aspect of process control.

Integral to the complete process flow is the mask preparation and the method of identification of any changes to the applicable mask set. The repeatability and quality of the masks should be assessed and documented prior to initiation of the fabrication process.

I. Electrostatic Discharge Characterization and Sensitivity

If not handled properly, several elements used in MMICs can be damaged by ESD. Damage may occur at tune-and-test, assembly, inspection, and other places, if proper precautions are not taken. Therefore, every process and design should be characterized to determine ESD sensitivity. Regardless of the test results, all GaAs-based devices should be treated as highly sensitive to ESD damage. An ESD handling and training program is essential to maintain a low level of ESD attributed failures.

Inspection, test, and packaging of MMICs should be carried out in static-free environments to assure that delivered products are free of damage. Devices should be packaged in conductive carriers and delivered in static-free bags. All handling and inspection should be performed in areas meeting "Class 1" handling requirements. Both the manufacturer and the user share the responsibility of assuring that an adequate procedure is in place for protection against ESD.

In general, the following steps can help reduce or eliminate the ESD problems in device manufacturing and test areas:

(1) Ensure that all workstations are static free.

(2) Handle devices only at static-free workstations.

(3) Implement ESD training for all operators.

(4) Control relative humidity to within 40 to 60%.

(5) Transport all devices in static-free containers.

(6) Ground yourself before handling devices.

IV. Product Qualification

A consumer expects the manufacturer to verify that his products are properly designed. A person buying a radio, for example, would expect it to receive RF energy in the AM and FM frequency bands and to convert that energy into a clear, audio wave. The consumer may also expect the manufacturer to specify the operating environment for which the product was designed. Again considering the radio, the consumer will want to know if it will work properly after storage in a shed in upper Michigan during the winter or on a boat during a summer sail of the Caribbean. The manufacturer can give these assurances and information only if he has tested the product after fabrication.

For MMICs, the process of obtaining this data is called product qualification or design validation, and, as implied above, every MMIC design should pass product qualification before it is sold. Because the data desired in product qualification is specific to a particular MMIC design, this applies as well to circuits fabricated on process-qualified fabrication lines. Figure 8-5 shows a product qualification procedure that addresses the issues critical to MMICs. The first step of design verification occurs before mask generation and includes design, simulation, and layout verification of the circuits. The rest of design verification includes full electrical characterization of the circuit to establish its operating performance, thermal analysis, and electrostatic discharge characterization, and verifies the results of the voltage ramp test, temperature ramp test, and temperature cycling tests. Although the sequence of the tests may be altered, it is recommended that design and layout verification be performed first, and this should be followed by electrical performance verification, since any out-of-specification parameters found during these tests will require a redesign of the circuit. This is a recommended approach, and all of the tests may not be required for some circuit designs. All participants in the MMIC design, manufacture, and end-product integration should be involved in deciding which tests are required.

The rationale for and a description of the steps recommended in the design validation follow.

A. MMIC Design, Model, and Layout Verification

One of the best ways of reducing MMIC engineering cost and improving reliability is to verify the design, model or simulation , and layout of the MMIC before fabrication begins. This critical step was addressed separately in Chapter 6. During the MMIC design cycle, these verifications are normally addressed through a series of design reviews that include representatives from all companies involved in the manufacturer and use of the MMIC. Furthermore, the representatives should come from all departments involved in the MMIC integration, including the MMIC designers, the fabrication staff, the RF metrology personnel, the packaging engineers, and the system designers.

Typically, the reviews are held before the circuits are sent to layout, after layout but before mask making, and after final MMIC characterization.

B. Thermal Analysis and Characterization

Thermal analysis determines the hottest part of the device during normal bias conditions and the temperature difference between the hottest point on the surface of the die and the case temperature; this is critical in determining the expected life of the MMIC. The analysis should be performed over the entire temperature range of the MMIC's intended application. Typically, this theoretical analysis is difficult and requires detailed knowledge of the power dissipation, geometry of the gold plating layers around the channel, method of attaching the die to the substrate, and the thermal boundary conditions of the substrate. A preferred method is actual thermal measurements using either liquid crystal or infrared scanning techniques.

C. Electrostatic Discharge Sensitivity Tests

GaAs devices are very sensitive to ESD damage, and therefore the ESD characterization given in [6] should be conducted to determine the sensitivity of the design. GaAs FET structures can be damaged by ESD voltages in the 20- to 2000-V range [4]. Thus, classification and treatment of the devices from the fabrication stage to the actual application as a Class 1 ESD-sensitive device is highly recommended. The device's normal electrical parameters should be used as a reference for degradation of performance due to testing.

Tests have shown that noncatastrophic damage can occur in the 50- to 75-V range for some devices. This damage is characterized by a slight increase in gate leakage current. As an example, typical leakage currents of 8 μA have been observed to increase to 30 to 40 μA after being subjected to 60- to 75-V ESD per MIL-STD-883 [6] test methods. The MMIC used in the test still operated properly and met all RF specifications, but catastrophic damage has been observed in the 50- to 200-V range.

Thin-film capacitors and resistors can be damaged by static charges of less than 2000 V and are therefore also "Class 1" devices. The voltages needed to damage these components are, however, much higher than those needed to damage FETs. Several hundred volts would damage these circuit elements; FETs are more susceptible to damage than capacitors and resistors.

Input and output blocking capacitors will not protect internal FETs from damage in most cases since ESD is usually present in the form of voltage transients and as such will be coupled through most capacitors. Therefore, it is recommended that all operators be careful when connecting these devices to RF test setups. Grounding the test technician prior to connecting the bias or RF leads is good practice.

It is not known what impact noncatastrophic damage will have on device lifetime. Tests on intentionally damaged devices have shown that they continue to operate for over 500 h at 85°C without further degradation. It is anticipated, however, that lifetime will be shortened when compared to undamaged devices.

D. Voltage Ramp

The sensitivity of an MMIC design to voltage overstress and the absolute maximum voltage ratings are determined during the voltage ramp test. Testing is normally done by ramping each device's power supply until a catastrophic failure occurs. Ramp rates and step duration are a function of the design limitations, but the test should allow thermal stabilization of the device at each successive step. After the test, an analysis to determine the exact failure site is recommended. Failure-point definition should be in conservative agreement with the device data sheet and design limits.

E. Temperature Ramp and Step Stress

Temperature ramping can serve more than one purpose. It can indicate which portion of the design is most sensitive to high-temperature operation, indicate the absolute maximum ratings applicable, give an indication of high-temperature operation characteristics, and it can determine the appropriate temperatures applicable for life tests. The test is normally done by ramping the temperature of the devices until catastrophic failure. Ramp rates and step duration should be designed to allow thermal stabilization of the devices at each successive step. Afterwards, failure analysis to determine the exact failure site is recommended. Failure point definition should be in conservative agreement with the device data sheet and design limits.

F. High/Low Temperature Tests

Data sheets will always specify the highest and lowest temperature at which an MMIC will operate, and it will give the percentage change in electrical parameters at the temperature extremes. The high/low temperature test is designed to obtain that data. The test temperature at both extremes may be obtained from step stress tests or from system requirements. Once the data have been measured for a specific MMIC design, the temperature limits and percent change in electrical parameters can be used in product acceptance screens.

V. Product Acceptance

Although an MMIC may be designed by highly qualified engineers, fabricated on a process qualified production line, and verified through measurements to meet the design goals, parts with poor reliability characteristics still exist. This may be due to variations in the fabrication process, or material flaws that were undetected, or, as is more often the case, to the MMIC package and stresses imposed on the MMIC during packaging. Regardless of the cause, these weak MMICs must be found and removed before they are integrated into the system. Therefore, manufacturers of all high reliability systems, including space systems, require the MMICs to pass a series of product acceptance screens, whose sole purpose is to increase the confidence in the reliability of the MMICs. Note that this step in the qualification methodology is the major difference between space qualified MMICs and commercial grade MMICs.

The level of testing performed under product acceptance is a function of the form of the deliverable. For example, the first level of acceptance testing, called "wafer acceptance test," is performed at the wafer level to assure the uniformity and reliability of the fabrication process through a wafer to wafer comparison. "Lot acceptance test for die" is a second level of testing that provides further reliability information, but only on a sample of the MMICs because of the difficulty in performing full characterization on

unpackaged MMICs. It is used if the MMIC user has requested the MMICs to be delivered in die form for integration into a larger module. This sample testing will provide the user with only an estimate of the MMIC's reliability. Furthermore, the user will not have an understanding of the MMIC's performance in the final package and any of the reliability issues that the package may cause. If packaged parts are requested though, a full 100% screening can be performed and the user should have assurance that the delivered parts are reliable. The acceptance testing procedure is summarized in Figure 8-6, where it is seen that the MMICs are not space qualified until they have passed the 100% screening tests, and the user takes responsibility for final space qualification screening if they request unpackaged parts.

The recommended product acceptance test for die deliverables is shown in Figure 8-7. Note there are three levels of testing within the procedure and each starts with the wafer acceptance test shown in Figure 8-8. The lowest level of testing is required for MMICs that have already been product qualified and have been manufactured on a qualified process line, whereas the highest level of testing is required for a new circuit design that is fabricated on an unqualified process line. Whichever level of testing is required, the same level of reliability assurance should be granted to the MMIC upon completion of the lot acceptance test. The cost and time advantage of buying MMICs from manufacturers with qualified processes and validated circuit designs can be large, and it is for this reason that manufacturers incur the cost of qualifying their processes.

A recommended flow chart for product acceptance of packaged parts is shown in Figure 8-9. It is assumed that a product acceptance of die deliverables is performed on the MMICs before they are inserted into the packaging process, or a subgroup of the parts can be removed from the packaged parts and life testing performed on them in a way similar to that recommended for the die deliverables. Thus, this screen adds further reliability information to the data obtained from the wafer and lot acceptance tests. As stated above, 100% of the packaged MMICs are recommended to be screened using Figure 8-9. Some of the steps require the selection of a particular screen, and this must be based on the intended application and device type.

Table 8-1 shows the recommended screening tests that can be used for MMIC packaged devices and the reference for the screen. This information is modified from MIL-PRF-38534 Class K requirements and should be applied after careful consideration of the applicability and the desired requirements.

Throughout the rest of this chapter, a brief description of and the rationale for product acceptance test or screen will be given.

A. Stabilization Bake

Some GaAs circuits have an initial period when their electrical parameters vary vs time. Most of the parametric variations decay to a steady-state value within 20 h, but during the initial life of the circuit, the variations can be large. Measured variations in I_{DS} of 20% over 2 h have been reported. The degree of instability varies between different manufacturers and between different fabrication processes at the same manufacturer. In fact, circuits from some manufactures do not exhibit any electrical parameter variations. It is therefore necessary to characterize the circuit performance over its early life to determine if electrical parametric variations occur. If they do occur, they must be eliminated before wafer acceptance, life testing, or product delivery can be made. If they are not eliminated, they will distort the life test results by shifting the "normal" operating parameters of the circuit; this will cause many circuits that are inherently good to appear defective.

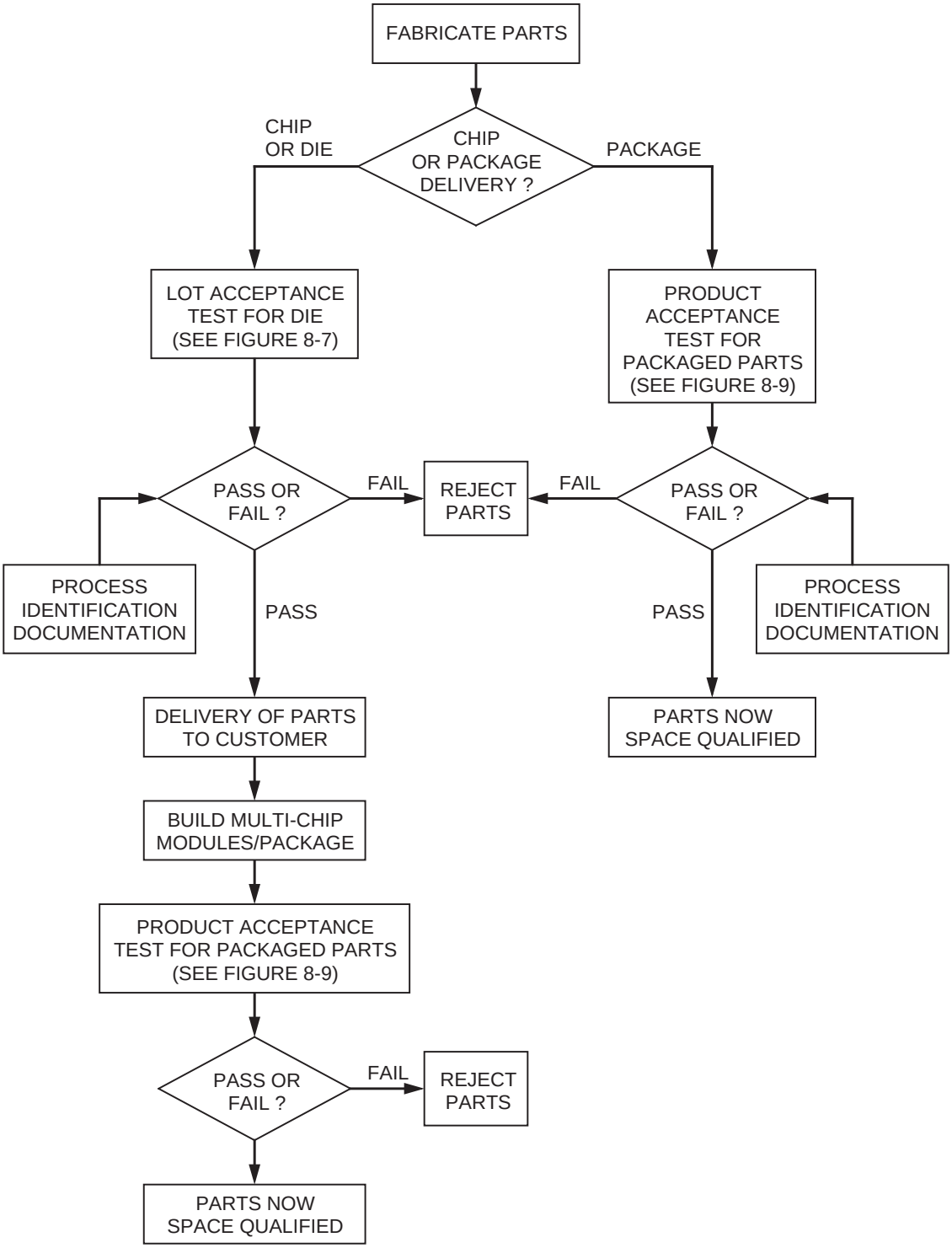

Figure 8-6. GaAs part qualification overview.

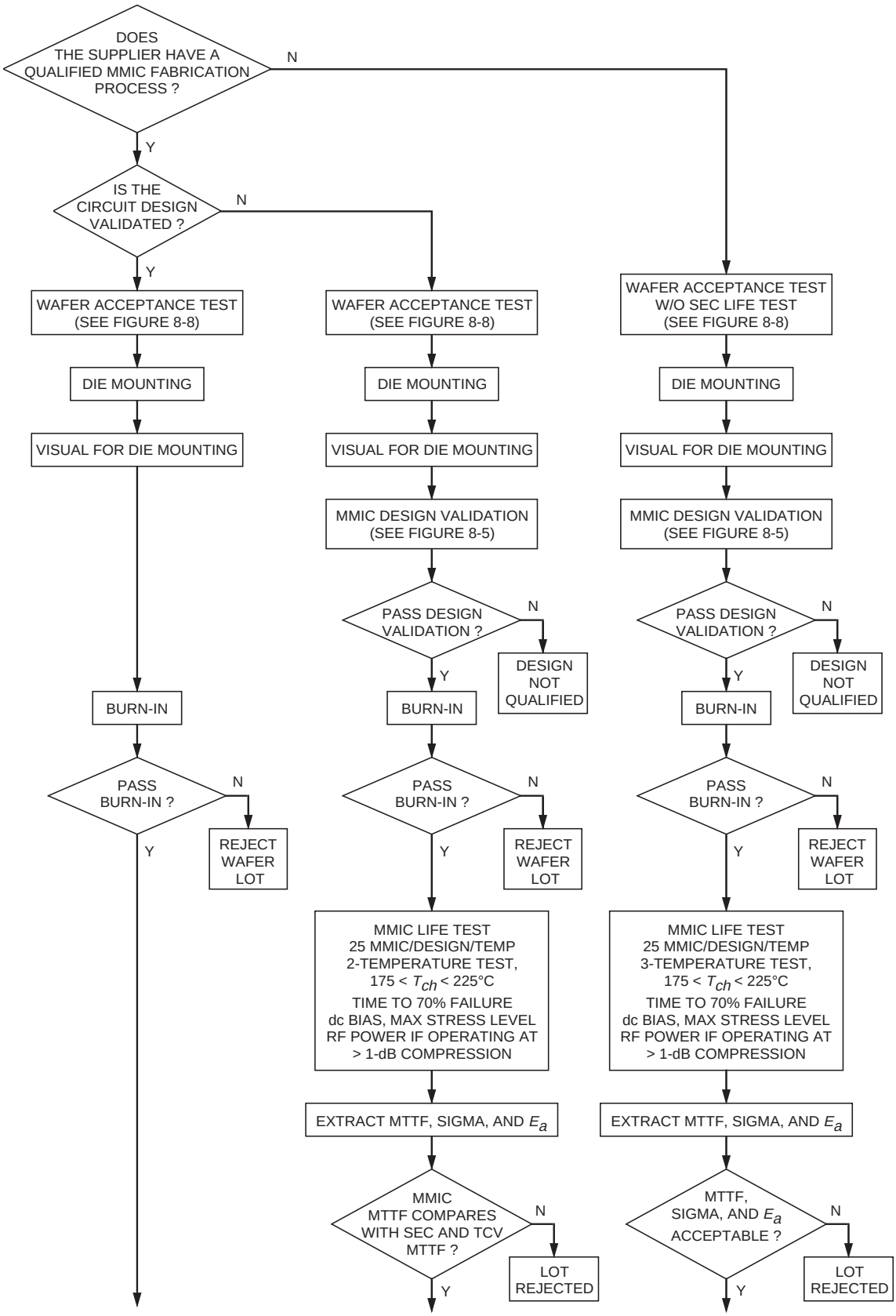

Figure 8-7. Lot acceptance test for die.

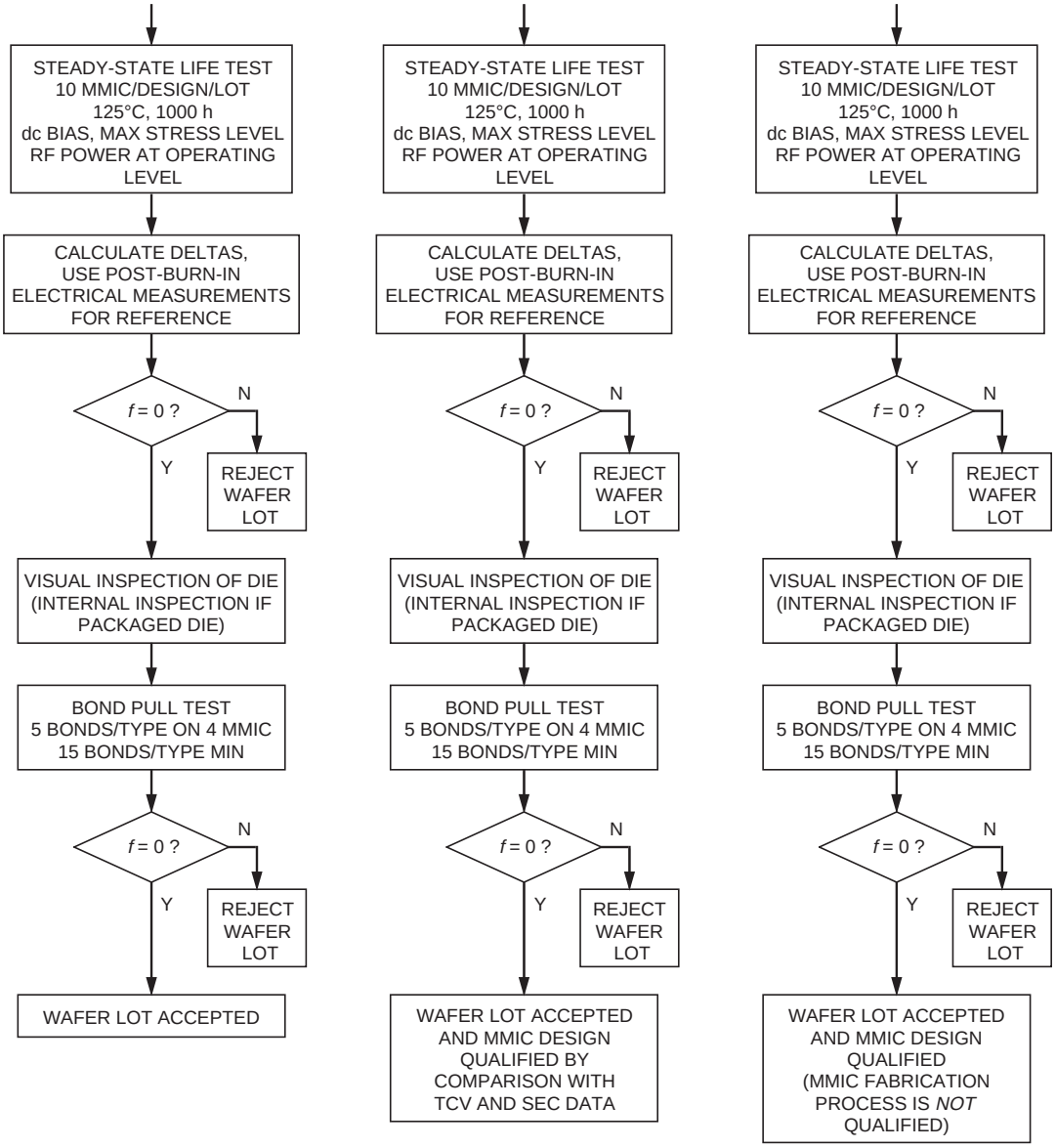

Figure 8-7. (continued)

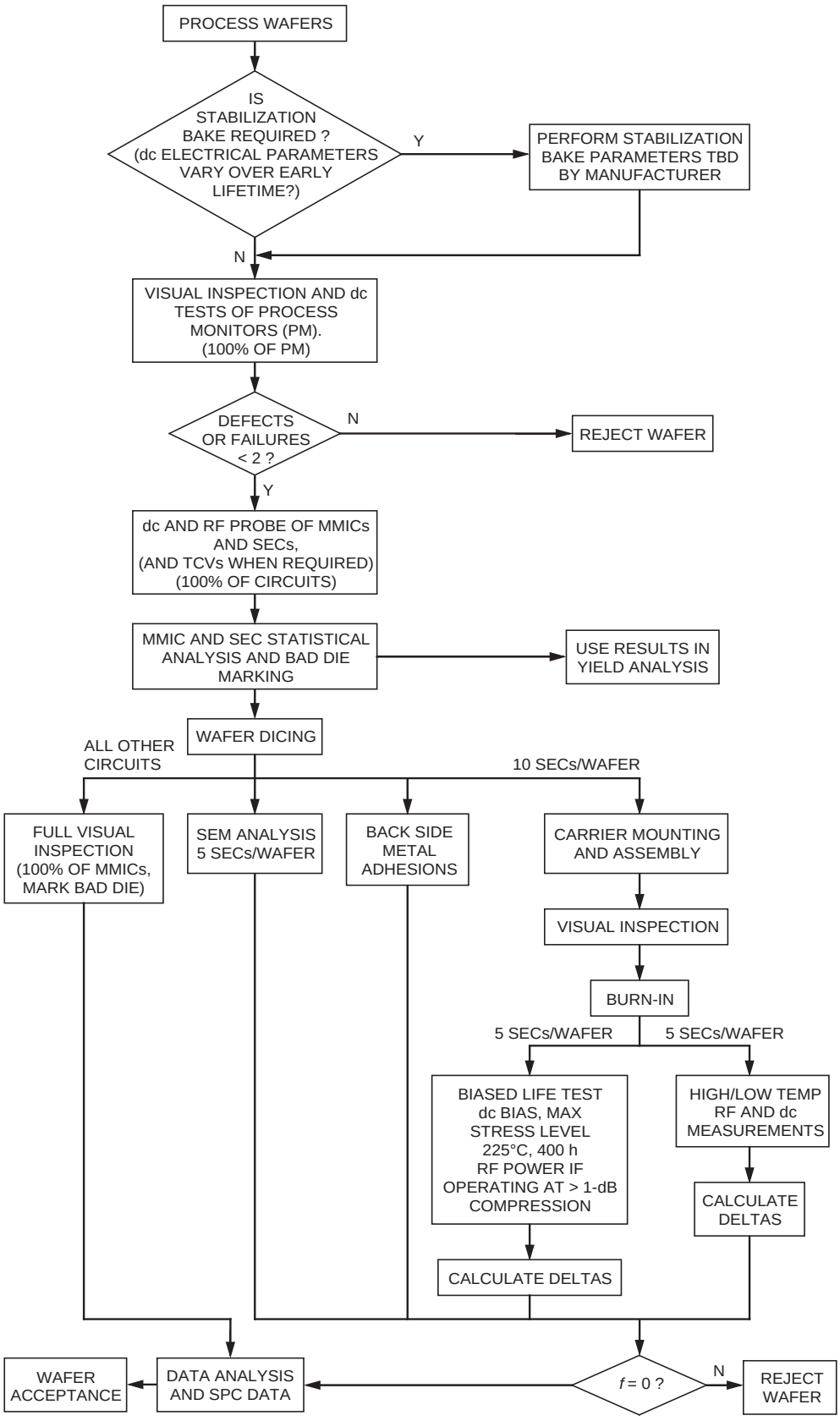

Figure 8-8. Wafer acceptance test.

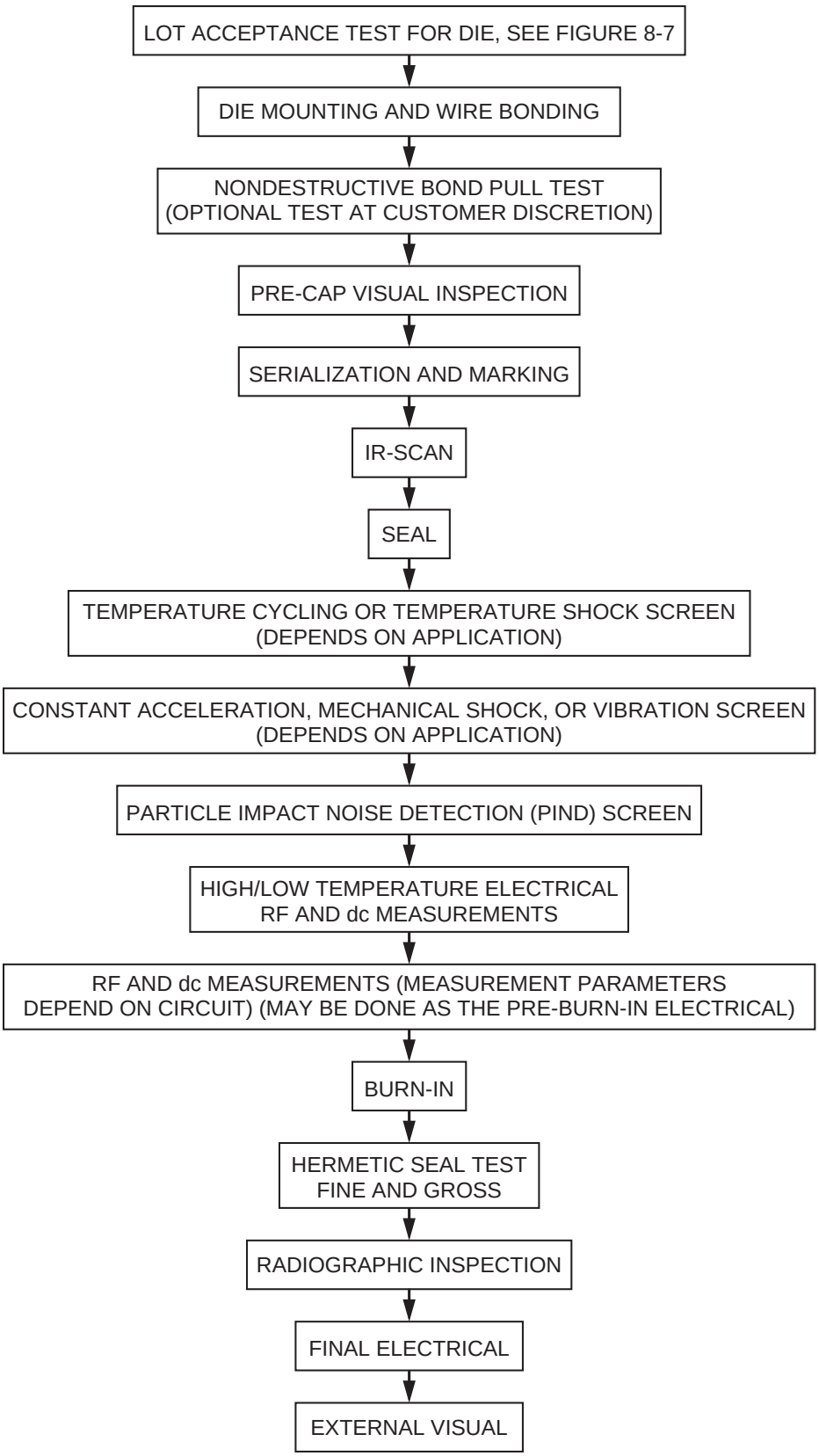

Figure 8-9. Screening process for packaged MMICs.

Table 8-1. Typical packaged device screening.

Test	Reference
Nondestructive bond pull	MIL-STD-883, Method 2023
Internal visual inspection	MIL-STD-883, Method 2017
IR-scan (prior to seal)	JEDEC Document JES2 [7]
Temperature cycling or Thermal shock	MIL-STD-883, Method 1010 MIL-STD-883, Method 1011
Mechanical shock or Constant acceleration	MIL-STD-883, Method 2002 MIL-STD-883 ,Method 2001
Particle impact noise detection	MIL-STD-883, Method 2020
Electrical	Customer's specification
Burn-in	MIL-STD-883, Method 1015
Electrical (high/low temp)	Customer's specification
Fine leak Gross leak	MIL-STD-883, Method 1014 MIL-STD-883, Method 1014
Radiographic	MIL-STD-883, Method 2012
External visual	MIL-STD-883, Method 2009

It has been shown that when parametric variations exist, the decay time is inversely proportional to the test temperature. In addition, it has been shown that a high-temperature bake may be used to stabilize the electrical parameters. These results may indicate that some of the fabrication processes, especially those that require bakes, are not adequately performed during fabrication. The alloying of ohmic contacts and the ion implantation activation bake are the two fabrication processes most often blamed. Another possibility is the development of mechanical stress in the GaAs lattice and in the metal deposited on the wafer during processing; this stress is relieved at high temperatures.

The bake performed to eliminate the parametric variations is called a stabilization bake. The stabilization bake is usually performed on the wafers immediately prior to dicing, but may be performed even before lapping and backside processing. The stabilization bake is an unbiased bake and should not be confused with the burn-in screen, which is a biased testing of the circuits at an elevated temperature. In addition, the stabilization bake is not the same as the Hi-Temperature Storage test, which some manufactures perform as part of the qualification process.

Although the stability of all electrical parameters is required before wafer acceptance, some manufacturers do not require a stabilization bake. Furthermore, some manufacturers who require stabilization bakes do not consider it a part of the wafer acceptance or reliability screening procedures, but rather a part of the fabrication process. Therefore, the stabilization bake may not appear in some manufacturers' reliability or product-acceptance procedures, while it does appear in others. Since the requirement for a stabilization bake is dependent on the manufacturer's processes, the bake temperature and time varies; typically, bake temperatures are between 200 and 300°C.

B. SEM Analysis

Scanning Electron Microscopy analysis can provide valuable information regarding the step coverage and quality of the metallization and passivation of GaAs devices. Thus, this tool is required as part of the wafer acceptance tests. Some accept/reject criteria are provided in MIL-STD-883, but they may need some modification to cover GaAs-device technology.

C. Nondestructive Bond Pull Test

The integrity of wire bonds cannot always be judged through visual and electrical tests. Therefore, some qualification procedures recommend the implementation of a nondestructive bond pull test of each bond. The pull force selected for this test is generally dependent on the material and wire diameter in question. MIL-STD-883, Method 2023, is normally used for this application. Obviously, selecting the required pull force is critical and must be decided by the manufacturer and the user.

Mechanical damage to good bonds as a result of this test is possible. Additionally, for microwave circuits, the wire bond's impedance can be changed when the shape of the wire loop is changed, which results in a change in the RF characteristics of the MMIC. Due to the problems associated with this test, some manufacturers have removed this step from their qualification and screen procedures and resorted to in-process controls and testing to provide the necessary information. The decision to require this test must be made by the MMIC user after careful consideration of the system application and the workmanship of the manufacturer.

D. Visual Inspection

Many defects in MMICs, such as metal voids, substrate cracks, poor wire bonds, and foreign materials, reduce the MMIC reliability. Small voids or cracks in the metallization will cause increased electrical resistance, increased current density, and an increased possibility of failure due to electromigration. Furthermore, microwave circuits radiate power at gaps and discontinuities in transmission lines. Edge chips and cracks created during wafer sawing or dicing easily propagate and cause circuit failures or die breakage during thermal cycling and wafer handling. This is especially true for GaAs monolithic circuits since GaAs wafers are more brittle than Si wafers and they are often thinned to 100 μm or less. The stray particles of GaAs created during wafer sawing or other byproducts of the circuit fabrication process may deposit themselves onto the wafer. Since GaAs is highly insulating, GaAs particles will usually not cause problems. However, other materials, especially metal particles, may adversely effect circuit performance. If the particles are on the gate of the transistors or on other circuit elements, the circuit performance will be degraded. This is especially true if the circuits have not been passivated. Since free particles may move during circuit testing, packaging, or in zero gravity space environments, even free particles away from the circuit elements may cause failures. During die attach, eutectic alloys and epoxies are used that may adhere to the sides or top of the circuit where it could short RF transmission lines and biasing pads to the ground plane. Lastly, the electrical connections between the package and the circuit must be made. These connections are usually made by ball or wedge bonds comprised of thin (typically 17 μm in diameter), gold wires attached to gold pads. The location and the quality of the bonds are critical for good MMIC performance and reliability.

These obvious defects and others not listed here in materials, construction, and workmanship must be eliminated since they degrade circuit performance and reliability. Furthermore, it is better to eliminate circuits with obvious defects before additional resources have been spent on them in bonding, packaging, and burn-in. Luckily, these defects are easily detected by performing a visual screen of every circuit with the aid of a microscope. The visual screen is performed during wafer acceptance tests for defects of the die and during the packaged MMIC screens for packaging and bonding defects.

E. IR Scan

Some defects such as substrate cracks and die-attach voids may not be visible, but they must be detected. Since these types of defects have a different thermal conductivity than the surrounding defect-free region, they may be detected through thermal mapping. The baseline for the comparison is the thermal profile of the MMIC that was made as part of the product or design verification step. For example, during design verification, it may have been determined that the final stage of an amplifier was the hottest part of the MMIC at 90°C, while the rest of the MMIC had a 15°C temperature variation. If a similar MMIC were thermally mapped and found to have a hot spot of 100°C or the wrong temperature variation across the die, a defect would be indicated. Typically, variations greater than 5°C are considered a reject. Thus, a simple comparison between the MMICs in the screening process and the MMIC thermal profile can be used to detect defects not visible to the eye.

Although infrared microscopes are expensive, require calibration, and have a minimum resolution of approximately 15 μm, they are the best method of mapping the MMIC's thermal characteristics since they do not damage the MMIC surface. Furthermore, the microscope can be computer controlled to scan the surface, make the required map, and perform the comparison to the thermal profile stored on file.

This screening step is not typically imposed as a requirement following MIL-PRF-38534. However, it is recommended for high-power devices and in applications that require good thermal stability. This step should be performed after die attach and before attachment of the package lid.

F. Temperature Cycling and Shock Screen

In the same way that electrical devices can be made to fail quicker at higher temperatures, mechanical devices can be made to fail quicker by applying thermal stress. These tests are used to detect flaws or weak points in the die attach, wire bonds, and package seals that would normally result in early failures. Temperature cycling consists of cycling the packaged MMICs between extreme temperatures many times. Typically, the temperatures used are -65 and 200°C, and the number of cycles is 15. The temperature shock screen is similar to the temperature cycle screen in that the test involves subjecting the packaged MMIC to extreme low and high temperatures (–65 and 150°C) over many cycles. The difference is a sudden change in temperature created by immersion of the parts into a bath, rather than the gradual change in air temperature used in the cycle test. Failure detection for both screens occurs during final electrical and visual inspections. Typically, only one of the screens is required, and the manufacturer and user decide on the appropriate screen for their application.

G. Mechanical Shock Screen

This screen is intended to detect weak parts that are required to undergo severe shocks during transportation, handling, satellite launch, or other operations. The test subjects the packaged MMIC to a number of short shock pulses with a defined peak. Failures are detected during final visual and electrical screens.

H. Constant Acceleration

This screen is intended to detect failures due to mechanical weakness by subjecting the packaged MMIC to a constant acceleration. Typical failures occur in the bonds and die attach, and these are detected during the final visual and electrical screens. Although this screen is typically required, it is not because of the forces caused during launch but rather as an effective tool to detect poor workmanship.

I. Particle Impact Noise Detection

During encapsulation, thermal stress screens, and mechanical stress screens, particles may break off the MMIC or package. These loose particles may mechanically damage the MMIC during handling, launch, or in operation, or they may cause short circuits. The particle impact noise detection screen is a nondestructive test used to find parts that have this defect. During the test, the part is vibrated and a sensor is used to detect anomalous noise. Failure criteria are given in the reference listed in Table 8-1.

J. Burn-In

Ideally, a well-controlled GaAs fabrication line, which employs proper wafer handling and fabrication procedures along with visual, dc, and RF screens, would eliminate circuits containing defects that result in the early failures that were discussed in Chapter 2. In fact, in some GaAs fabrication lines, the early failure rate is very small. However, in state-of-the-art circuits with 0.1- to 0.25-μm gate HEMTs, complex circuits with many air bridges, or packaged circuits with many wire bonds, latent defects may cause early failures at a higher rate. These are detected through the burn-in screen.

The burn-in screen stresses the circuits above their normal operating conditions to accelerate any early failures that would occur from latent defects. Although burn-in is often performed at elevated temperatures to shorten the time of the burn-in test, the temperature must be kept low enough so inherently good circuits do not fail due to failure mechanisms accelerated by the test. Also, since circuits that pass burn-in are used in either accelerated life testing or as flight deliverables, burn-in at too high of a temperature will distort the results of the accelerated life tests and reduce circuit lifetime during the mission. It is inevitable that the burn-in screen will use some circuit life, but if the circuit has an inherently long lifetime and the burn-in screen is not performed at too high of a temperature, only a few percent of the life will be lost. This small cost in circuit lifetime is accepted by the space industry, since the alternative is a failed mission or satellite.

To prevent creation of failures in inherently good circuits due to excessive stress conditions, burn-in should be performed only once on each circuit and appropriate test conditions should be selected. Circuits that fail burn-in should not be reworked and retested. If the circuits are to be delivered to another company for further processing or packaging, it is critical that the burn-in screen is coordinated to assure that it is not duplicated. An exception can be made if the system builder performs a burn-in on the

entire assembly, since assembly burn-ins are normally performed at lower temperatures and for shorter times than the GaAs die burn-in. Therefore, the total stress to the MMIC from the additional assembly burn-in is small and should not affect the circuit's lifetime.

It should be noted that only a small percentage of GaAs circuits fail the burn-in screen, and the burn-in screen increases the circuit cost. Furthermore, the increased handling of the circuits during the screening procedures increases the chance of creating failures in the circuits due to introduced mechanical, ESD, and handling defects. Therefore, most suppliers of commercial MMICs do not perform burn-in screens, but all satellite manufacturers require burn-in of all electronic parts.

The screen is typically performed at 125°C ambient temperature for 320 h with the circuits biased to their maximum stress levels. However, careful consideration of the resultant device channel temperature is recommended to avoid undue stress of the device during test and the introduction of thermally accelerated failure mechanisms. If the MMIC is classified as a large-signal (greater than 1-dB compression) device, RF energy should also be applied to the input port with the output port matched. Failure is usually specified as an electrical parametric drift from the initial conditions by a specified percentage. These conditions have been shown to be effective in removing weak MMICs.

K. Leak Test

There was considerable discussion in Chapter 4 about failure mechanisms that result from contamination and humidity. To eliminate these problems, MMICs, as well as all other electronic components intended for high-reliability applications, are sealed in hermetic packages, and the reliability of the MMICs is dependent on the integrity of these packages. To find weak packages that would result in loss of the hermetic seal, thermal and mechanical stress screens were performed. Although some gross package failures are visually detectable, most defects in the package require a leak test.

Fine leak tests consist of placing the packaged MMIC in a chamber pressurized with a known gas; some of the gas will enter cracks or defects in the package if they exist. Usually, helium or nitrogen gas with a small concentration of a radioactive isotope is used, since either is detectable in very small concentrations using standard, commercially available equipment. After a time, the chamber is cleansed by circulating air, and the packages are tested to determine if gas leaks from them. Although the use of radioactive isotopes sounds hazardous, it is the preferred method in high-volume production lines because it is easier to detect for a longer period of time. The disadvantage of fine leak testing is that the gas will leak from gross defects before it can be detected. Therefore, a gross leak test is required after the fine leak test. The principle of the test is the same except that a pressurized liquid bath is used instead of the gas.

L. Radiographic

The final screen is usually a radiographic “picture” of the inside of the sealed package taken in the same way that a doctor takes X-rays to image the skeletal structure of the human body. This nondestructive test uses radiation to penetrate the package walls and produce a shadow image on a photographic plate. It is useful for checking the location and position of wire bonds and for detecting loose particles that may have moved or broken off during the screening process. In some cases, this screen can also be useful in determining the presence of die-attach voids.

References

[1] MIL-I-38535B, *Integrated Circuits (Microcircuits) Manufacturing*, Department of Defense, Washington, DC, June 1, 1993.

[2] MIL-STD-45662, *Calibration System Requirements*, Department of Defense, Washington, DC, 1988.

[3] *Distributor Requirements for Handling Electrostatic Discharge Sensitive (ESDS) Devices*, JEDEC Publication EIA 625, EIA & JEDEC Standards & Engineering Publications, 1994.

[4] MIL-STD-1686, *Electrostatic Discharge Control Program for Protection of Electrical and Electronic Parts, Assemblies and Equipment (Excluding Electrically Initiated Explosive Devices)*, Department of Defense, Washington, DC, 1992.

[5] *General Standard For Statistical Process Control*, JEDEC Publication EIA 557A, EIA & JEDEC Standards & Engineering Publications, 1995.

[6] MIL-STD-883, *Test Methods and Procedures for Microelectronics*, Department of Defense, Washington, DC, November 15, 1991.

[7] *Transistor, Gallium Arsenide Power FET, Generic Specification*, JEDEC Publication JES2, July 1992, EIA & JEDEC Standards & Engineering Publications.

Additional Reading

Microelectronics Failure Analysis Techniques, A Procedural Guide, E. Doyle, Jr., and B. Morris, Editors, report written for Air Force Systems Command, Rome Air Development Center under contracts F30602-78-C-0339 and F30602-78-C-0281.

Guidelines for GaAs MMIC FET Life Testing, JEDEC Publication JEP118, EIA & JEDEC Standards & Engineering Publications, January 1993.

Chapter 9. GaAs MMIC Packaging

A. Immorlica, S. Kayali, G. E. Ponchak, and R. N. Simons

I. Introduction

G. E. Ponchak

Alone, an MMIC die sawed from a GaAs wafer is extremely fragile and must be protected from mechanical damage and hostile environments. In addition, it is electrically and thermally isolated and thus requires interfaces to electrical sources, other components, and thermal sinks. The broad technology focused on providing these functions is called packaging. Because of the many functions it must fulfill, packaging of MMICs, as with all ICs, is a technically challenging and critical step in the production of the product. The package provides mechanical support and protection, thermal heat sinks or paths to dissipate the heat generated by the IC, and electrical contact pads for both the RF and dc bias leads.

For microwave and millimeter-wave circuits, the package design must also provide electromagnetic shielding from the outside environment and between MMICs within a multichip package. Specific applications impose other unique design constraints. For phased-array antennas, the MMIC package should be smaller than a half of a wavelength to permit the proper antenna element spacing. Thus, very small packages such as the 20- to 40-GHz ceramic package shown in Figure 9-1 are required. Alternatively, the package may be designed with antenna elements on the package surface or inside the package; radiation from the latter must be through one of the package surfaces. For example, Figure 9-2 shows a 30-GHz, multichip package that contains four MMIC phase shifters, a power divider, control circuits, and four radiating elements protruding out of the package walls. For wireless applications, the package must be inexpensive and of low weight to be useful in a hand-held transmitter/receiver. Even optical interfaces may be required in packages designed to house microwave optical modulators or optically controlled MMICs. Each application and each MMIC or MMIC chip set represents new challenges and design constraints for the package designer.

Besides the electrical, thermal, size, and cost constraints imposed on the package design, the reliability of the package itself must be considered since the package must have a lifetime greater than or equal to the MMICs it is protecting from hostile conditions. In addition, the package design, materials, and fabrication must not degrade the MMICs performance or reliability. Unfortunately, it is common for the MMIC characteristics to change and new failure mechanisms to develop as a result of MMIC packaging, this due to the presence of the lid, sidewalls, coupling between components, mechanical stresses, and chemical interactions between materials. Therefore, since the reliability of the packaged part is the end user's ultimate concern, the user must consider the total packaged MMIC assembly in the reliability specifications. It is not sufficient to assume reliability of a packaged MMIC because the MMIC and the package have individually passed acceptance tests or have been qualified previously.

This chapter will first discuss the functions of the package in detail, then introduce different types of microwave packages and the advantages and disadvantages of each, and, finally, present reliability issues of several packages or packaging technologies. Within the context of this guide, this chapter should be used to gain understanding of the reliability issues; a more in-depth investigation with the aid of the reference lists is left to the reader.

Figure 9-1. 20- to 40-GHz ceramic MMIC package. (Fabricated by Hughes Aircraft Company under contract to NASA Lewis Research Center.)

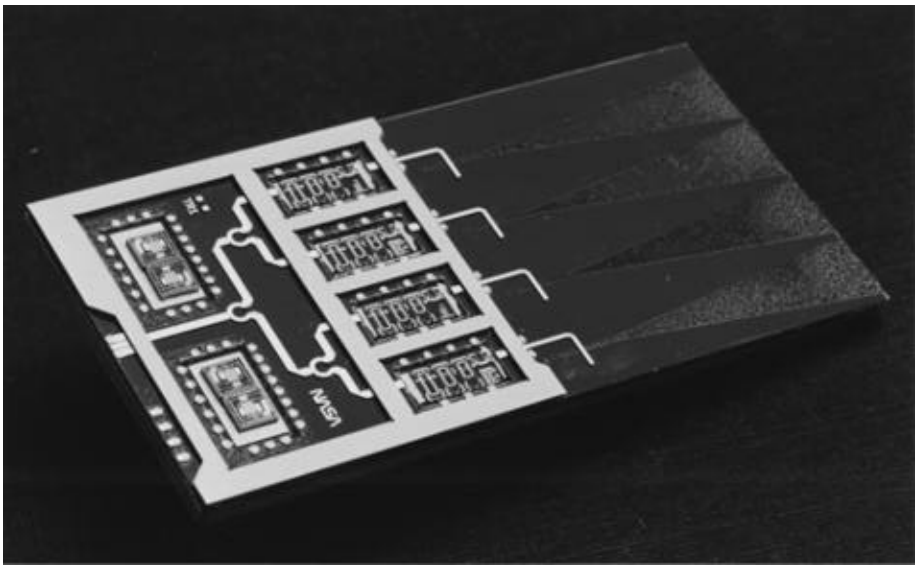

Figure 9-2. Prototype of a four-element antenna package with MMIC phase shifters, power divider, and control circuits. (Fabricated under SBIR Contract NAS3-25870 for NASA Lewis Research Center.)

A. Functions of Microwave Packages

Summarized in a single statement, the package serves to integrate all of the components required for a system application in a manner that minimizes size, cost, mass, and complexity, provide electrical and thermal interfaces between the components and the system, and ensure the reliability of the individual components and the overall package. The following subsections present the four main functions of the package: mechanical support and protection, protection from the environment, power and signal distribution, and thermal stabilization.

1. Mechanical Support and Protection

With the reduction of the MMIC substrate thickness to 25 to 100 μm to facilitate heat dissipation and lower via-hole inductance and 0.1-μm feature sizes, the requirement to support and protect the MMIC from thermal and mechanical shock, vibration, high acceleration, particles, and other physical damage during the storage, launch, and operation of the parts in space becomes critical. The mechanical stress endured depends on the mission or application. For example, landing a spacecraft on a planet's surface creates greater mechanical shock than experienced by a communication satellite. There is also a difference between space and terrestrial applications: In space, particulates are suspended and can damage the MMIC if they impact or land on electrically sensitive areas.

In a typical MMIC package, the three components shown in Figure 9-3 are integrated to protect the MMIC. A carrier or the package base supports the MMIC, the ring (or sidewalls) encloses all of the components in the package, and a lid seals the top. The base or carrier may be the most critical part of the package, because it is the only part in contact with the MMIC. The base or carrier may be designed with raised areas and wells to accommodate the MMIC, other ICs, signal distribution networks, and chip capacitors and resistors, but the most important decision made in the design of the base is the choice of materials. The coefficient of thermal expansion (CTE) of the carrier should be equal to or slightly greater than the CTE of GaAs for reliability, since thermal shock or thermal cycling may cause die cracking and delamination if the materials are unmatched or if the GaAs is subject to tensile stress. Other important parameters are thermal resistance of the carrier, the material's electrical properties, and its chemical properties, or resistance to corrosion. Typical materials used for the base of MMIC packages are metal alloys such as CuMo and CuW, metal composites such as Kovar™ and Silvar™, ceramics such as alumina and low-temperature cofired ceramic (LTCC), and glass, quartz, and diamond. If the base provides only mechanical support and thermal dissipation, metal and metal composites are preferred because of their high thermal and electrical conductivities and low manufacturing costs. However, if the base is also used as a substrate for electrical transmission lines, electrically insulating materials are required. Also, the method used to attach the MMIC die to the carrier will have a major effect on reliability; this subject is covered separately in Sections 9-II and 9-III.

Once the MMIC is supported on a carrier, the other components have been added, and the wire bonds or other electrical connections are made, the assembly must be protected from scratches, particulates, and other physical damage. This is accomplished either by adding walls and a cover to the base or by encapsulating the assembly in plastic or other material. Since the electrical connections to the package are usually made through the walls, the walls are typically made from glass or ceramic. Although the CTE

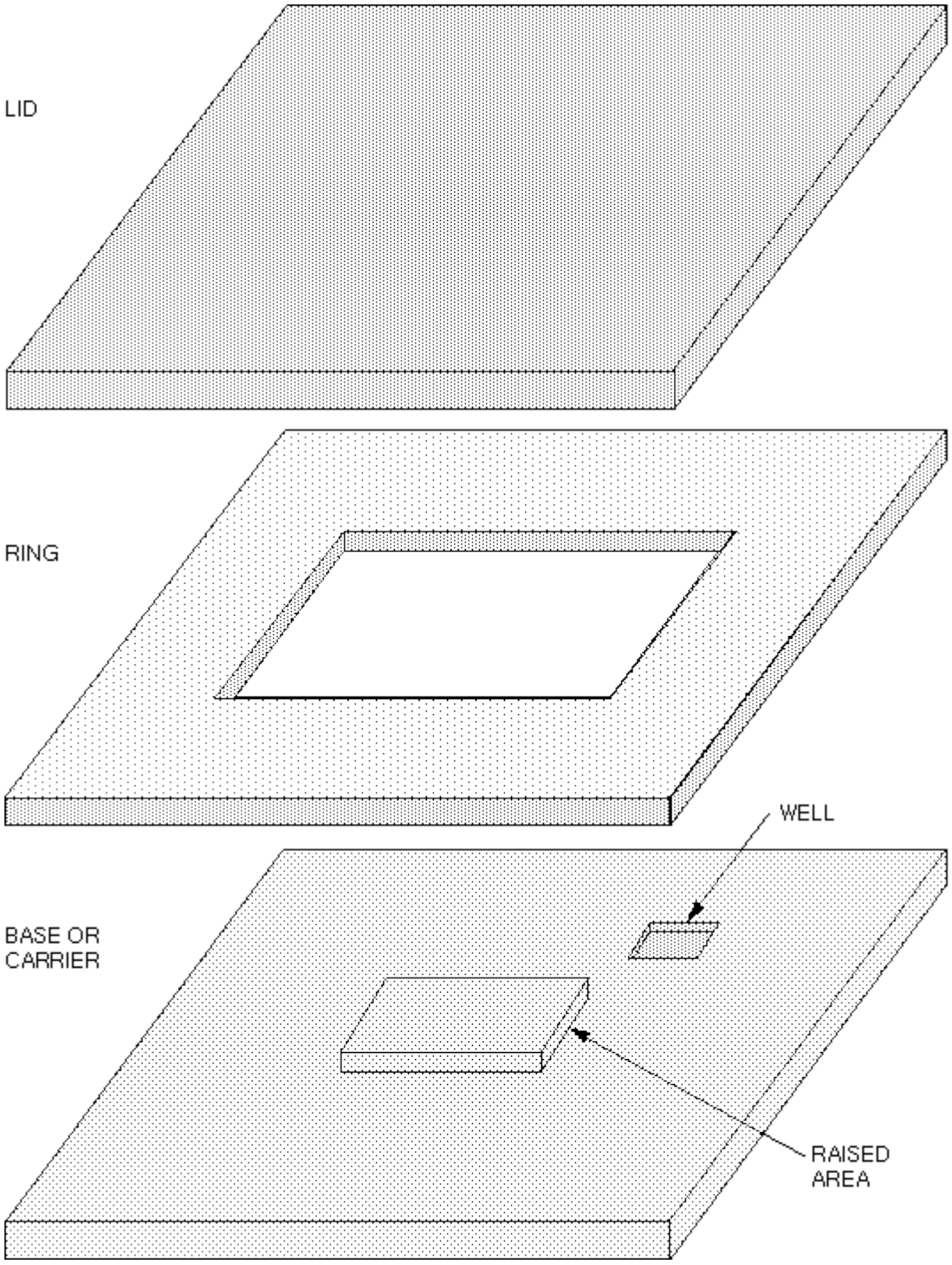

Figure 9-3. MMIC package.

of the walls and lid does not have to match the CTE of GaAs since they are not in contact, it should match the CTE of the carrier or base to which they are connected.

2. Protection From Environment

Many elements in the environment can cause corrosion or physical damage to the metal lines of the MMIC and other components in the package. Although there is no moisture in space, moisture remains a concern for MMIC space applications since it may be introduced into the package during fabrication and before sealing. The susceptibility of the MMIC to moisture damage is dependent on the materials used in its manufacture. For example, Al transmission lines and gate fingers can corrode quickly in the presence of moisture, whereas Au lines degrade slowly, if at all, in moisture. Also, junctions of dissimilar metals can corrode in the presence of moisture. Moisture is readily absorbed by some materials used in the MMIC fabrication, die attachment, or within the package; this absorption causes swelling, stress, and possibly delamination.

To minimize these failure mechanisms, MMIC packages for high reliability applications are normally hermetic with the base, sidewalls, and lid constructed from materials that are good barriers to liquids and gases and do not trap gasses that are later released. In addition, the bonds used in fabricating the package are compatible with hermetic package construction.

A recently discovered failure mechanism in GaAs MMICs produces a sudden change in transistor current when the MMIC is operated in a hydrogen-rich atmosphere. Unfortunately, hydrogen is easily trapped and released by many materials, and within a hermetic package, the hydrogen concentration can be high. This new failure mechanism is not well understood, but of enough importance to be discussed separately in Section 9-VII.

3. Power and Signal Distribution

Because the package is the primary interface between the MMIC and the system, it must provide the transfer of dc power and RF signals between the two. In addition, the package must distribute the dc and RF power to the components inside the package. The drive to reduce costs and system size by integrating more MMICs and other components into a single package increases the electrical distribution problems since the number of interconnects and transmission lines within the package increases. Furthermore, the current carried by the dc lines increases because more power is required, but the bias voltage cannot be raised because of system constraints. The dc power is usually fed into the package along metal lines that pass through the package walls, called feedthroughs, and it is then routed between different circuits along metal lines that may have vertical interconnects through via holes. These features may be seen in Figures 9-1 and 9-2.

RF signals can also be introduced into the package along metal lines passing through the package walls, or they may be electromagnetically coupled into the package through apertures in the package walls. Ideally, RF energy is coupled between the system and the MMIC without any loss in power, but in practice, this is not possible since perfect conductors and insulators are not available. In addition, power may be lost to radiation, by reflection from components that are not impedance matched, or from discontinuities in the transmission lines. Reflections also create standing waves since reflected signals add constructively and destructively at different points along the line. The final connection between the MMIC and the dc and RF lines is usually made with wire bonds, although flip-chip die attachment, discussed in Section 9-III, and multilayer interconnects using thin dielectric layers over the MMIC with via-hole interconnects are

gaining acceptance. Wire bonds add significant inductance to the transmission line and therefore limit high frequency operation and change the matching impedance. Therefore, short, flat, ribbon bonds are preferred when they are compatible with the application.

Within the package, undesired coupling between different parts of the circuit results in energy transfer from one line to another. Typically, coupling is stronger between adjacent transmission lines with discontinuities, but severe radiation from a discontinuity may travel across the substrate and couple to nonadjacent transmission lines. Examples of circuit elements especially prone to radiation, and therefore coupling, are wire bonds with large loops, open-circuit microstrip stubs, and apertures in ground planes. Coupling can be reduced by partitioning the package into smaller areas with metal walls or via-hole fences. Unfortunately for the package designer, models and design rules for packages and microwave components placed in packages are not available. Microstrip, coplanar waveguide, and stripline are usually used for the microwave transmission lines within the package and for the feedthroughs, but the presence of sidewalls, cover plates, and adjacent lines changes the line characteristics from their normal, open characteristics and complicates the design. Therefore, the perfection of package design usually requires more complex electromagnetic simulators and several iterations. Further discussion of problems resulting from electromagnetic effects is presented in Section 9-VI.

4. Thermal Stabilization

It was shown in Chapter 4 that the reliability of GaAs devices is inversely proportional to the junction temperature of the devices. For small signal circuits, the temperature of the device junction does not increase substantially during operation, and thermal dissipation from the MMIC is not a problem. However, with the push to increase the power from amplifiers, coupled with their poor efficiency, and the increased level of integration within a package, the temperature rise in the device junctions can be substantial and cause the device to operate in an unsafe region. Therefore, thermal dissipation requirements for power amplifiers, other large signal circuits, and highly integrated packages can place severe design constraints on the package design.

The junction temperature of an isolated device can be determined by

$$T_j = Q * R_t + T_{case}$$

where Q is the heat generated by the junction and is dependent on the output power of the device and its efficiency, R_t is the thermal resistance between the junction and the case, and T_{case} is the temperature of the case. Normally, the package designer has no control over Q and the case temperature, and therefore, it is the thermal resistance of the package that must be minimized. Figure 9-4 is a schematic representation of the thermal circuit for a typical package, where it is assumed that the package base is in contact with a heat sink or case. It is seen that there are three thermal resistances that must be minimized: the resistance through the GaAs substrate, the resistance through the die-attach material, and the resistance through the carrier or package base. Furthermore, the thermal resistance of each is dependent on the thermal resistivity and the thickness of the material. A package base made of metal or metal composites has very low thermal resistance and therefore does not add substantially to the total resistance, but electrically insulating materials used for bases, with the exception of diamond, have less thermal conductance than metal. To maintain high thermal dissipation through these materials,

metal-filled via holes are routinely used under the MMICs to provide a thermal path to the heat sink. Although thermal resistance is a consideration in the choice of the die-

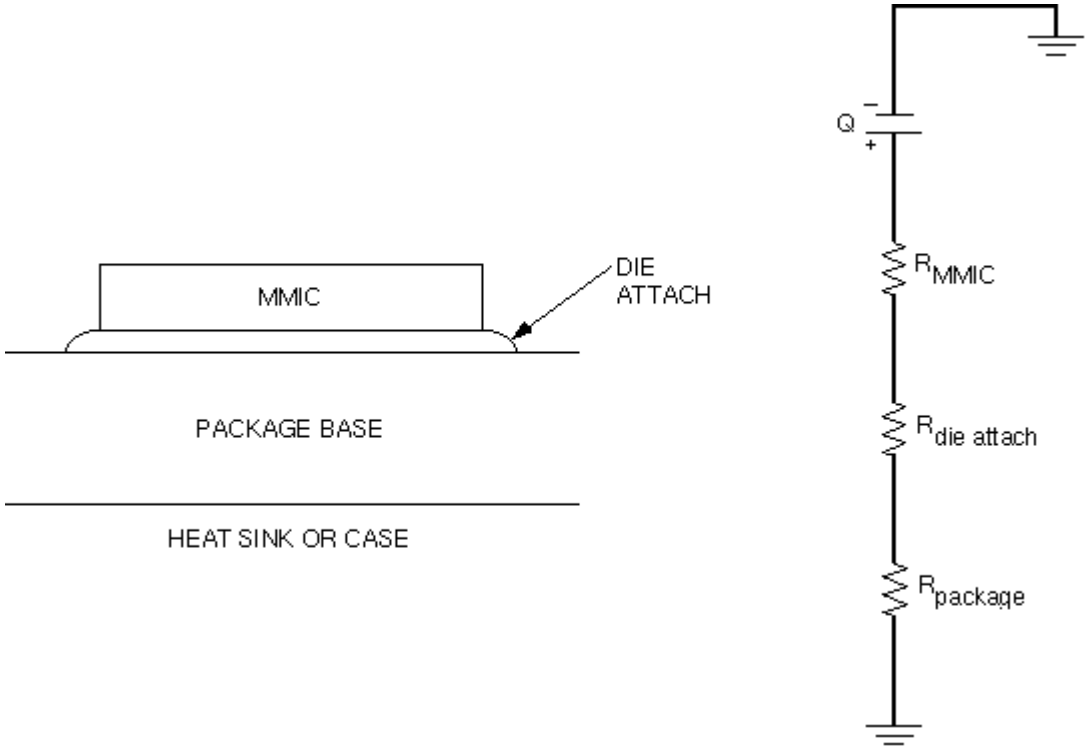

Figure 9-4. Cross section of MMIC attached to a package and its equivalent thermal circuit.

attach material, adhesion and bond strength are even more important. To minimize the thermal resistance through the die-attach material, the material must be thin, there can be no voids, and the two surfaces to be bonded should be smooth. Ideally, the thermal resistance of the GaAs determines the total thermal resistance. To minimize R_{MMIC}, substrate thicknesses have been reduced from a standard of 100 μm to 25 μm.

B. Types of Microwave Packages

As discussed earlier in this section, each application or MMIC usually requires a new package design to optimize the performance of the circuit or to meet the needs of the system. The packages shown in Figures 9-1 and 9-2 clearly illustrate the differences in the size, construction, and features that microwave packages may have, but it is possible to loosely group packages into several categories. Four of these categories—all metal packages, ceramic packages such as those shown in Figures 9-1 and 9-2, plastic packages, and thin-film multilayer packages—are presented below.

1. Metal Packages

The metal packages shown in Figures 9-5 and 9-6 are often used for microwave multichip modules, passive circuits such as filters and power dividers, and hybrid circuits because they provide excellent thermal dissipation, excellent electromagnetic shielding, and they can have a large internal volume while still maintaining mechanical reliability.

The package can use either an integrated base and sidewalls with a lid, as the two shown in Figures 9-5 and 9-6, or it can have a separate base, sidewalls, and lid. Inside the

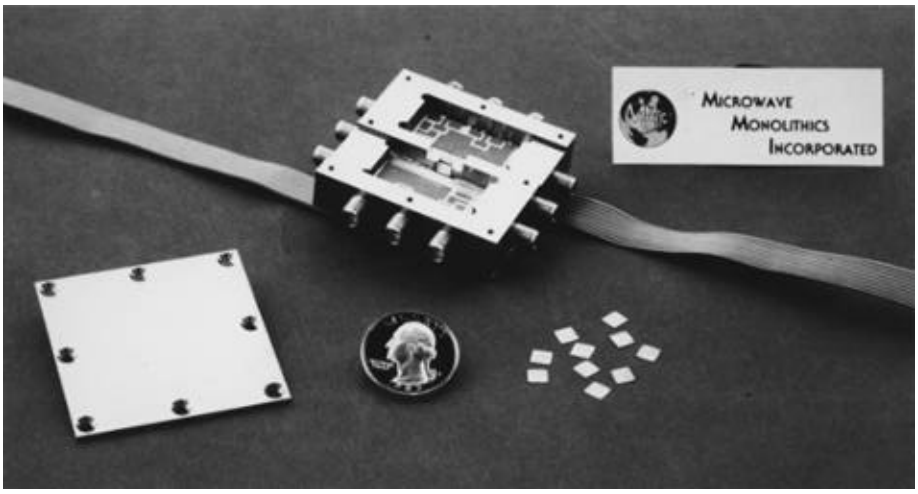

Figure 9-5. GaAs MMIC switch matrix in a metal package. (Fabricated under contract to NASA Lewis Research Center.)

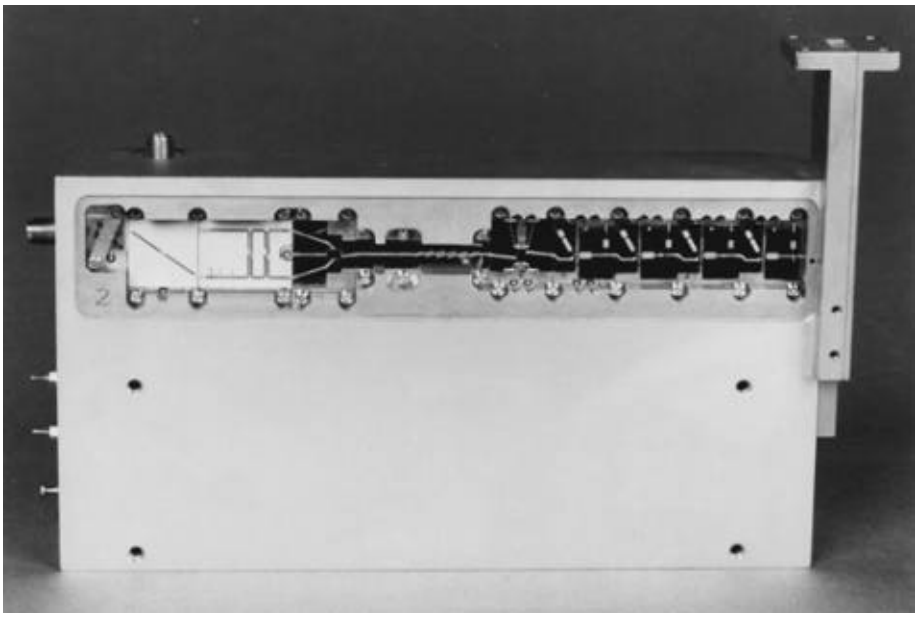

Figure 9-6. 20-GHz receiver in a metal package. (Fabricated by Harris Corporation under contract to NASA Lewis Research Center.)

package, metal partition walls are often added to decrease coupling between MMICs and to eliminate waveguide resonance in the package. Waveguide resonance occurs when the package becomes equal to $\lambda/2$ in any direction. The partition walls also act as ribs to strengthen the package. Lastly, ceramic substrates or chip carriers are required for use with the MMICs and feedthroughs.

The selection of the proper metal is critical. CuW/ 10–90, Silvar™ (a Ni–Fe alloy), CuMo/ 15–85, and CuW/ 15–85 have good thermal conductivity and a slightly higher CTE than GaAs, which makes them good choices. Kovar™, a Fe–Ni–Co alloy commonly used for sidewalls and lids, is not recommended for the base since its CTE is slightly less than the CTE of GaAs. All of the above materials, in addition to Alloy-46, may be used for the sidewalls and lid. Cu, Ag, or Au plating of the packages is commonly done.

Before final assembly, a bake may be performed to drive out any trapped gas or moisture. This reduces the onset of the hydrogen-related failures. During assembly, the highest curing epoxies or solders should be used first and the processing temperature should decrease until the final lid seal is done at the lowest temperature to avoid later steps damaging earlier steps. Au–Sn is a commonly used solder that works well when the two materials to be bonded have similar CTEs. Au–Sn solder joints of materials with a large CTE mismatch are susceptible to fatigue failures after temperature cycling, and Au–Sn intermetallics may form that have unfavorable mechanical properties. Welding using lasers to locally heat the joint between the two parts without raising the temperature of the entire part is a commonly used alternative to solders. Regardless of the seal technology, no voids or misalignments should be tolerated since they may cause the package to fail hermetic tests.

Feedthroughs or dc and RF interconnects can be coax-to-microstrip launchers, rectangular waveguide-to-microstrip transitions, and planar ceramic lines. These are illustrated in Figures 9-5, 9-6, and 9-1, respectively. Significant reflections can result within the package at the connections between the feedthroughs and the transmission lines, and undesired modes can be launched on the transmission lines. Some of these issues are discussed in Section 9-VI.

2. Ceramic Packages

Ceramic packages have several features that make them especially useful for MMICs: low mass, reduced waveguide box resonance, mass-production compatibility and therefore low cost, they can be made hermetic, and can more easily integrate signal distribution lines and feedthroughs. As illustrated in Figures 9-1 and 9-2, they can be machined to perform many different functions, and, by incorporating multiple layers of ceramics and interconnect lines, the interconnect line loss and parasitic effects are reduced. Multilayer ceramic packages also allow reduced size and cost of the total microwave system by integrating multiple MMICs and other components into a single, hermetic package. These multilayer packages offer significant size and mass reduction over metal-walled packages. Most of that advantage is derived from the close spacing of MMICs that is possible in ceramic packages and the use of three dimensions instead of two for interconnect lines.

The most widely used ceramics for MMIC packages belong to the class known as “low temperature cofired ceramics” (LTCC). These materials are based on Pb-B-Si-O glass with alumina fillers. The material properties, including the dielectric constant and loss tangent, are dependent on the ceramic composition; generally, the relative dielectric

constant is in the range of 6 to 9.9 and the loss tangent is acceptable for microwave applications. Other material properties that must be considered are the CTE, the processing temperature of the ceramics, the processing temperature of the metals used for interconnect lines, and interactions between the various materials.

LTCC packages are constructed from individual pieces of ceramic in the "green" or unfired state. These materials are thin, pliable films. During a typical process, the films are stretched across a frame in a way similar to that used by an artist to stretch a canvas across a frame. On each layer, metal lines are deposited using thick-film processing (usually screen printing), and via holes for interlayer interconnects are drilled or punched. If larger holes or cutouts are required in the layer to form wells for MMICs, they are also drilled, usually by a laser. After all of the layers have been fabricated, the unfired pieces are stacked and aligned using registration holes and laminated together. Finally, the part is fired at a high temperature of 800 to 1000°C. The MMICs and other components are then epoxied into place, and wire bonds are made the same as those used for metal packages.

Several problems that can affect the reliability of the MMIC arise from the fabrication procedure. First, the green-state ceramic shrinks during the firing step. The amount of shrinkage is dependent on the number and position of via holes and wells cut into each layer. Therefore, different layers may shrink more than others creating stress in the final package. Second, because ceramic-to-metal adhesion is not as strong as ceramic-to-ceramic adhesion, sufficient ceramic surface area must be available to assure a good bond between layers. This eliminates the possibility of continuous ground planes for power distribution and shielding. Instead, metal grids are used for these purposes. For microwave transmission lines, the ground planes are reduced by design to three times the strip width; this reduction increases the conductor loss of the lines. Third, the choice of metal lines is limited by the processing temperature and ceramic properties. To eliminate warping, the shrinkage rate of the metal and ceramic must be matched. Also, the metal must not react chemically with the ceramic during the firing process. The metals most frequently used for LTCC packages are Ag, AgPd, Au, and AuPt. Ag migration has been reported to occur at high temperatures, high humidity, and along faults in the ceramic. The microwave design issues that arise from ceramic packages are covered in depth in Section 9-VI.

3. Thin-Film Multilayer Packages

The disadvantages of ceramic packages need to be addressed. First, wire bonds are required for the final connection to the MMIC along with all of the parasitics associated with them (assuming flip-chip technology is not used). Second, the space between lines to reduce coupling is relatively large .

Thin-film multilayer packages solve both of these problems. Within the broad subject of thin-film multilayer packages, two general technologies are used. One uses sheets of polyimide laminated together in a way similar to that used for the LTCC packages described above, except a final firing is not required. Each individual sheet is typically 25 μm and is processed separately using thin-film metal processing. The second technique also uses polyimide, but each layer is spun onto and baked on the carrier or substrate to form 1- to 20-μm-thick layers. In this method, via holes are either wet etched or reactive ion etched (RIE). The polyimide for both methods has a relative permittivity of 2.8 to 3.2. Since the permittivity is low and the layers are thin, the same characteristic impedance lines can be fabricated with less line-to-line coupling; therefore, closer spacing of lines is possible. In addition, the low permittivity results in low line capacitance and therefore faster circuits. The wire bonds can be eliminated by depositing

the polyimide over the MMIC and using via holes to couple the RF signal, dc power lines, and the MMIC. These features of thin-film mutilayer packages are illustrated in Figure 9-7.

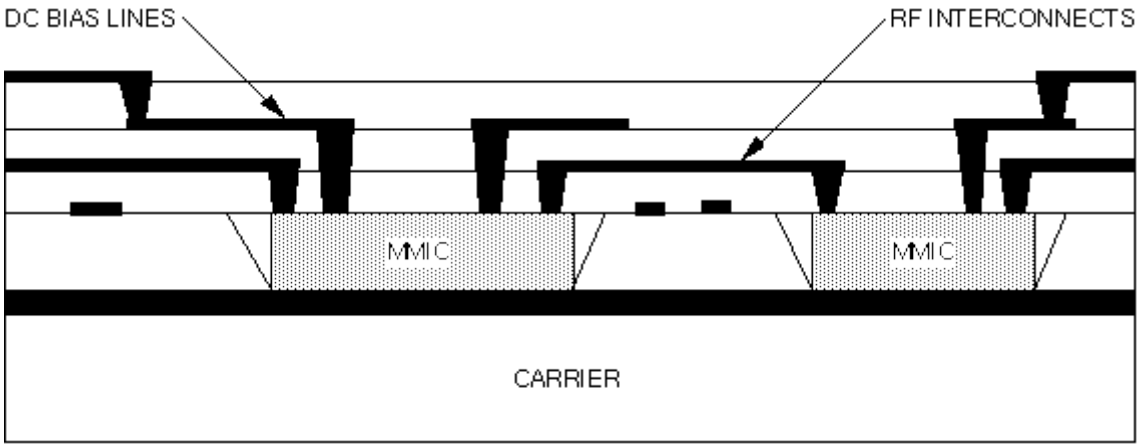

Figure 9-7. Schematic of thin-film multilayer package with integrated MMICs.

As with other technologies, there are problems with thin-film multilayer packages. The conductor loss per unit length can be as high as 10 dB/cm at 110 GHz, but because the line lengths are short, the loss is acceptable. Other problems relating to reliability are covered in depth in Section 9-IV.

4. Plastic Packages

Plastic packages have been widely used by the electronics industry for many years and for almost every application because of their low manufacturing cost. High-reliability applications are an exception because serious reliability questions have been raised. Plastic packages are not hermetic, and hermetic seals are required for high-reliability applications. The packages are also susceptible to cracking during temperature cycling in humid environments or where the plastic has absorbed moisture. The packaging of GaAs MMICs in plastic for space applications may gain acceptability if one of the proposed LEO or GEO satellite constellations for personal communications makes successful use of it. The reliability of plastic packages is presented in Section 9-V.

Additional Reading

Electronic Packaging and Interconnection Handbook, C. A. Harper, Editor, McGraw–Hill, Inc., New York, 1991.

Microelectronics Packaging Handbook, R. R. Tummala and E. J. Rymaszewski, Editors, Van Nostrand Reinhold, New York, 1989.

T. A. Midford, J. J. Wooldridge, and R. L. Sturdivant, "The Evolution of Packages for Monolithic Microwave and Millimeter-Wave Circuits," *IEEE Trans. on Ant. and Prop.*, Vol. 43, No. 9, pp. 983–991, September 1995.

D. S. Wein, "Advanced Ceramic Packaging for Microwave and Millimeter Wave Applications," *IEEE Trans. on Ant. and Prop.,* Vol. 43, No. 9, pp. 940–948, September 1995.

II. Die Attachment

G. E. Ponchak

How the GaAs MMIC is attached to the package (die attach) is a general technology applicable to all of the package configurations to be discussed in this section. Die attachment performs several critical functions. It must provide a good thermal path between the MMIC and the package base, which is itself usually attached to a heat sink to remove the heat generated by the MMIC. It must provide a good electrical contact between the backside metal of the MMIC, or its ground plane, and the package base that usually serves as the ground plane for the microwave interconnect lines within the package. Lastly, it must perform these two critical roles over the lifetime of the MMIC and through the environmental conditions required for the mission.

The stability and reliability of the die attach is largely determined by the ability of the structure to withstand the thermomechanical stresses created by the difference in the CTE between the GaAs and the package base material. These stresses are concentrated at the interface between the MMIC ground plane and the die-attach material and the interface between the die-attach material and the package. GaAs has a CTE of 5.8 ppm/K and most packages have a slightly higher CTE (6 to 10 ppm/K [1]); this puts the GaAs MMIC in compression as shown in Figure 9-8. An expression has been developed to relate the number of thermal cycles a die attachment can withstand before failure to the properties of the system [2]. This expression, the Coffin–Manson relation [2], is

$$Nf \propto \gamma^m \left(\frac{2 * t}{L * \Delta CTE * \Delta T} \right)$$

where

γ = shear strain for failure

m = constant dependent on the material

L = diagonal length of the die

t = die-attach material thickness

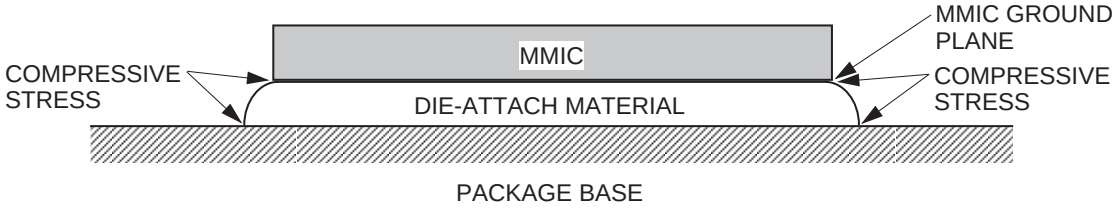

Figure 9-8. GaAs MMIC in compression.

The number of thermal cycles before failure can be significantly reduced by the presence of voids in the die-attach material as shown in Figure 9-9, since voids cause areas of concentrated localized stress concentration that can lead to premature die delamination. The void density tends to increase as the package assembly is thermally cycled. Also, voids cause localized heating of the MMIC, since the void is not a good thermal conductor. Therefore, the thermal resistance of the die-attach material increases as the

system is thermally cycled [3]. To minimize these effects, the CTE of the GaAs MMIC and the package base material should be matched as closely as possible, the die-attach material should have a high shear strain before failure, and the presence of voids in the die-attach material must be avoided. This is best accomplished by cleaning the GaAs MMIC and the package before assembly, the removal of all oxides from the two surfaces to be attached, performing a dehydration bake before die attach, and storage of the parts in an inert atmosphere before die attach.

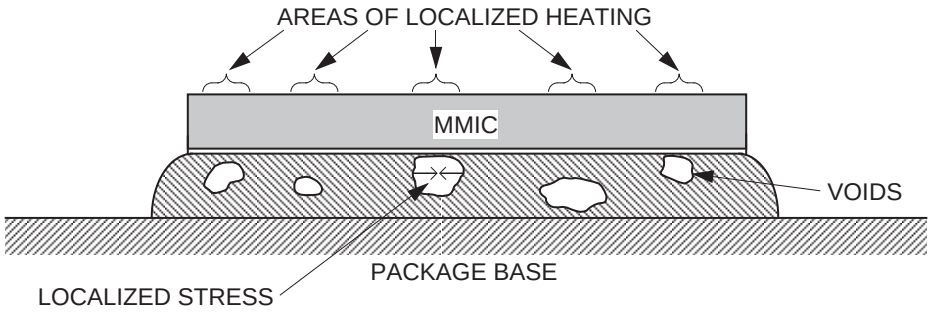

Figure 9-9. The presence of voids in the die-attach material.

Another potential failure mechanism is created if the die-attach material fills any unfilled via holes of the MMIC since the CTE of the die-attach material probably will not match the CTE of the GaAs. During thermal cycling, the die-attach material will expand at a different rate than the GaAs and cause cracks in the via holes [4]. This can be avoided by proper visual inspection of the backside metallization of the MMIC before packaging to reject any MMICs having unfilled via holes, although the same problem can occur from a CTE mismatch between the plated Au in the via hole and the GaAs substrate.

Presently, microwave packages use either hard solders, soft adhesives, or epoxies for die attach. Each method has advantages and disadvantages that affect MMIC reliability. Consider first the solders. Typically, a Au-Sn (80/20) solder is used for GaAs MMICs since it works with the Au ground plane of the MMIC and the package base material to form a Au-Sn eutectic when the assembly is heated to approximately 250°C in the presence of forming gas. Thus, a single, rigid assembled part with low thermal and electrical resistances between the MMIC and the package base is fabricated. Furthermore, these desirable characteristics are stable when the packaged MMIC is thermally cycled [5]. Since the solder die attach is rigid, it is even more critical that the CTE of the MMIC and the package be matched since the solder cannot absorb stresses created by thermal cycling or the die-attach process, and die cracking can result. Regardless of the solder used, flux, a commonly used soldering agent to assure the two surfaces to be bonded are clean and wettable, should not be used for GaAs MMIC since flux degrades the MMIC reliability [1].

Adhesives and epoxies are comprised of a bonding material filled with metal flakes, as shown in Figure 9-10. Typically, Ag flakes are used as the metal filler since it has good electrical conductivity and has been shown not to migrate through the die-attach material [6,7], even under thermal stress. The advantages of these die-attach materials are the lower processing temperature, between 100 and 200°C, required to cure the material, their ease of application, and a lower built-in stress from the assembly process as compared to solder attachment. Furthermore, since the die attach does not create a rigid assembly, shear stresses caused by thermal cycling and mechanical forces are

relieved to some extent [2,4]. This characteristic of the die-attach material also leads to crack formation and delamination during temperature cycling [2]. Before a catastrophic die-delamination failure, the MMIC will probably fail parametrically, since the thermal resistance of the die attach increases as voids and cracks are formed. Increases in R_T of a factor of 6 after 1000 temperature cycles are possible [2]. The rate of change in R_T vs the number of temperature cycles is dependent on the contact materials and how well they match the CTE of the GaAs [2,5]. Other disadvantages of the soft die-attach materials are a significantly higher electrical resistivity, which is 10 to 50 times greater than solder, and thermal resistivity, which is 5 to 10 times greater than solder. Therefore, solder is probably required for power amplifiers because of the need for low thermal contacts. Lastly, humidity has been shown to increase the aging process of the die-attach material [7].

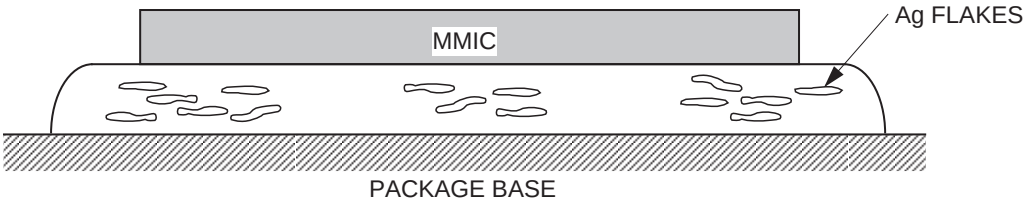

Figure 9-10. Bonding material.

References

[1] *Gallium Arsenide Products: Designers' Information*, Texas Instruments Application Note, Dallas, Texas, pp. 13–15, 1992.

[2] T. Tuhus and A. Bjorneklett, "Thermal Cycling Reliability of Die Bonding Adhesives," 1*993 IEEE Annual Inter. Reliability Physics Symp. Digest*, pp. 208, March 23–25, 1993.

[3] N. Strifas, P. Yalamanchili, and A. Christou, "Die Attach Void Formation and Delamination at the GaAs MMIC/Ceramic Substrate Interface," *1995 GaAs Reliability Workshop Programs and Abstracts*, pp. 55–60, October 29, 1995.

[4] P. Yalamanchili and A. Christou, "Finite Element Analysis of Millimeter Wave MMIC Packages Subjected to Temperature Cycling and Constant Acceleration," *1993 GaAs REL Workshop Programs and Abstracts*, October 10, 1993.

[5] M. A. Gaynes, R. H. Lewis, R. F. Saraf, and J. M. Roldan, " Evaluation of Contact Resistance for Isotropic Electrically Conductive Adhesives," *IEEE Trans. on Components, Packaging, and Manuf. Tech. - Part-B,* Vol. 18, No. 2, pp. 299–304, May 1995.

[6] H. L. Hvims, "Conductive Adhesives for SMT and Potential Applications," *IEEE Trans. on Components, Packaging, and Manuf. Tech. - Part-B*, Vol. 18, No. 2, pp. 284–291, May 1995.

[7] O. Rusanen and J. Lenkkeri, "Reliability Issues of Replacing Solder With Conductive Adhesives in Power Modules," *IEEE Trans. on Components, Packaging, and Manuf. Tech. - Part-B*, Vol. 18, No. 2, pp. 320–325, May 1995.

III. Flip-Chip Package

R. N. Simons

Finite element analysis as well as experimental studies have shown that chips with large edge length and small bump height tend to fail faster than chips with small edge length and large bump height. The reliability of flip-chip contacts is determined by the difference in the CTE between the chip and the ceramic substrate or the organic printed circuit board (PCB) [1,2]. For example, the CTE for GaAs is 5.8 ppm/K, for 96% alumina it is 6.4 ppm/K, and for PCB it is typically 20 to 25 ppm/K. The CTE mismatch between the chip and the carrier induces high thermal and mechanical stresses and strain at the contact bumps. The highest strain occurs at the corner joints, whose distance is the largest from the distance neutral point (DNP) on the chip [1]. For example, the DNP for a 2.5- $\times$ 2.5-mm chip is 1.7 mm. The thermomechanical stress and strain cause the joints to crack. When these cracks become large, the contact resistance increases, and the flow of current is inhibited. This ultimately leads to chip electrical failure. The failure criterion is an increase in resistance in excess of 30 mΩ over the zero time value [1]. The tradeoff in selecting the bump height is that large bumps introduce a series inductance that degrades high-frequency performance and increases the thermal resistance from the MMIC to the carrier, if that is the primary heat path.

The reliability of the bump joints is improved if, after reflow, a bead of encapsulating epoxy resin is dispensed near the chip and drawn by capillary action into the space between the chip and the carrier. The epoxy is then cured to provide the final flip-chip assembly. Figure 9-11 shows a typical flip-chip package. The epoxy-resin underfill mechanically couples the chip and the carrier and locally constrains the CTE mismatch, thus improving the reliability of the joints. The most essential characteristic of the encapsulant is a good CTE match with the *z*-expansion of the solder or the bump material. For example, if one uses 95 Pb/5 Sn solder having a CTE of 28 ppm/K, an encapsulant with a CTE of about 25 ppm/K is recommended. Underfilling also allows packaging of larger chips by increasing the allowable DNP. In some cases, the encapsulant acts as a protective layer on the active surface of the chip. Typical material properties of encapsulant used in flip-chip packaging are presented in [3,4].

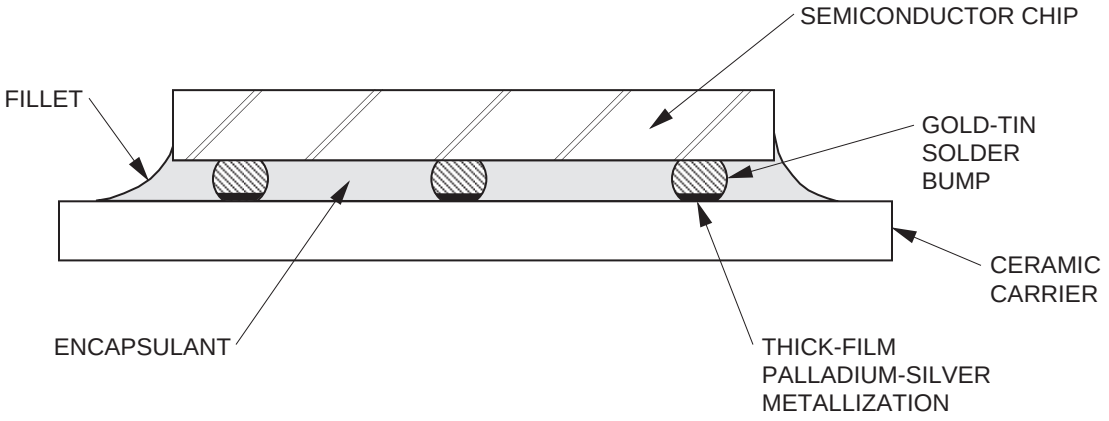

Figure 9-11. Flip-chip package.

Good adhesion between the underfill material, the carrier, and the chip surface is needed for stress compensation. The adhesion between the surfaces can be lost and

delamination can take place if contaminants, such as post-reflow flux residue, are present. For this reason, a fluxless process for flip-chip assembly is desirable [1]. Unfortunately, flip-chip bonding on PCB requires the use of flux [2]. However, on ceramic carriers with gold, silver, and palladium–silver thick-film patterns and via metallizations, fluxless flip-chip thermocompression bonding with gold–tin bumps has demonstrated high reliability [1]. The results of reliability testing [1] are summarized in Table 9-1 and may serve as a guideline for future work.

Table 9-1. Summary of reliability test conditions and results for fluxless flip-chip thermocompression-bonded bump contacts.

Parameter	Value
Bump height	30 to 70 μm
Chip size	A few mm
Chip carrier	Ceramic
Carrier camber	5 μm per cm
Camber compensation	By bump deformation
Underfill	Yes
Thermal cycling	After 6500 cycles (–55°C to +125°C), no contact failure and no change in contact resistance
High-temperature storage	After 1000 h, no increase in contact resistance
Temperature and humidity	After 1000 h (85°C and 85% RH), no change in contact resistance
Pressure-cooker test	After 1000 h (121°C and 29.7 psi), contact resistance increased slightly from 3 mW to 4 mW

Finally, care should be taken that the encapsulant or underfill covers the entire underside without air pockets or voids and forms complete edge fillets around all four sides of the chip. Voids create high-stress concentrations and may lead to early delamination of the encapsulant. After assembly, a scanning acoustic microscope can be used to locate voids in the encapsulant. The encapsulant should also be checked for microcracks or surface flaws, which have a tendency to propagate with thermal cycling and environmental attacks, eventually leading to chip failure [3].

References

[1] J. Kloeser, E. Zake. F. Bechtold, and H. Reichl, "Reliability Investigations of Fluxless Flip-Chip Interconnections on Green Tape Ceramic Substrates," *IEEE Trans. Components, Packaging, and Manufacturing Technology, Part A*, Vol. 19, No. 1, pp. 24–33, March 1996.

[2] J. Giesler, G. O'Malley, M. Williams, and S. Machuga, "Flip Chip on Board Connection Technology: Process Characterization and Reliability," *IEEE Trans. Components, Packaging, and Manufacturing Technology, Part B*, Vol. 17, No. 3, pp. 256–263, August 1994.

[3] D. Suryanarayana, T. Y. Wu, and J. A. Varcoe, "Encapsulant used in Flip-Chip Packages," *IEEE Trans. Components, Packaging, and Manufacturing Technology*, Vol. 16, No. 8, pp. 858–862, December 1993.

[4] G. O'Malley, J. Giealer, and S. Machuga, “The Importance of Material Selection for Flip Chip on Board Assemblies,” *IEEE Trans. Components, Packaging, and Manufacturing Technology, Part B*, Vol. 17, No. 3, pp. 248–255, August 1994.

IV. Multichip Module–Dielectric Package

G. E. Ponchak

The most frequently used materials for multichip module–dielectric (MCM-D) packages are polyimides and other polymers. These materials absorb moisture to varying degrees. Materials with the lowest moisture uptake have a water content of 0.5% [1] while other materials can have a water content greater than 4% [2]. Life tests on die containing Al test structures that were coated by various polymers showed failure modes attributed to Al : H_2O interactions. For polyimides, the measured median life of the Al test structure was approximately 385 and 6950 h when subjected to 121°C, 99.6% RH, and 85°C, 85% RH, 40-V bias, respectively. Further results have shown that passivation increases the median life of the test structures and that the passivation is not harmed by the polymer coating [3]. Therefore, passivation of the die is required to minimize moisture-failure mechanisms, and polymer coatings are not a substitute for hermetic packaging.

Adhesion of the polymer to the substrate is a major failure concern. Stresses are created at the polymer/substrate interface due to differences in the CTE between the two materials. These stresses can be large at room temperature due to the high processing temperature required to cure the polymers; the processing temperatures can be greater than 300°C. The amount of stress is proportional to the CTE mismatch and the polymer film thickness [4]. If the stress is greater than the adhesion strength of the polymer/substrate interface, the polymer will delaminate from the substrate. Stress cracks at the corners of via holes are another potential problem since the cracks tend to grow with thermal cycling. Optimization of the processing steps can minimize this effect [5]. Even if the polyimide does not delaminate from the substrate or base material, the stress may cause the substrate to warp. As a rule of thumb, the thickness of a ceramic base should be 20 times the thickness of the polyimide. Semiconductor substrates require an even greater thickness to avoid warpage [6].

The metal system used in the MCM-D process must be optimized. Typically, Cu is used for all of the dc and RF lines because of its good electrical conductivity and low cost. Unfortunately, it has been shown that Cu diffuses into polyimide at a rapid rate. The diffusion mechanism is temperature dependent. At low temperature ($T < 185°C$), Cu atoms diffuse through the polyimide, as shown in Figure 9-12. The rate of diffusion can be high enough that Cu will diffuse through 1 μm of polyimide in 4 months. At higher temperatures, Cu atoms are self gettering. Therefore, line widths are reduced, and the metal profile changes as shown in Figure 9-13. Finally, at the glass-transition temperature of the polyimide, Cu cluster migration has been reported [7]. Also, Cu has been shown to have poor adhesion to polyimides [5]. To minimize these failure mechanisms, Cr or Ti is required as a diffusion barrier between the Cu and the polyimide, but Cr/Cu/Cr metal pads have poor mechanical properties due to metal interdiffusion. Therefore, Ni is required as a diffusion barrier between the Cr and Cu. Ni cannot be used as a diffusion barrier between the Cu and polyimide since it also diffuses into polyimide [8]. Lastly, Cu diffusion is greater if the Cu lines are on top of the polymer surface instead of imbedded in the polyimide. The increased diffusion rate is due to surface voids in the polymer and CuO formation that has a higher diffusion rate than Cu [7]. In addition, exposed Cu lines on the surface of an MCM-D will corrode. Therefore, Cr or Au capping is required for upper level metal lines [9].

Via-hole formation is critical for MCM-D technologies since they are used in large numbers for interlevel interconnects. The via holes are made either by laser

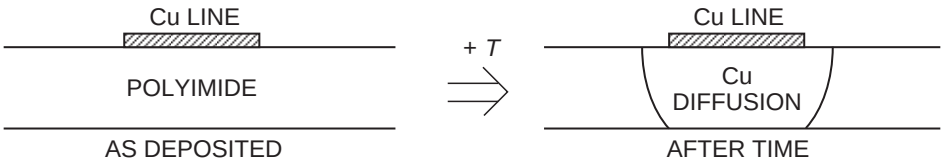

Figure 9-12. Copper diffusion in polyimide.

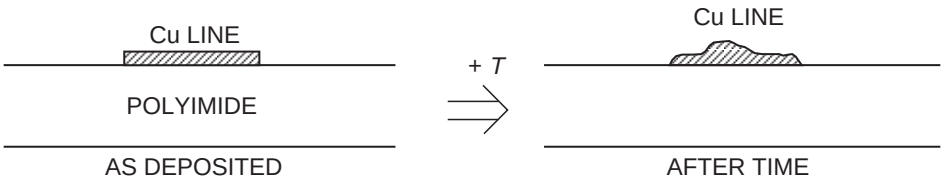

Figure 9-13. Self-gettering of Cu changes line geometry.

drilling, wet etching, or dry etching. Cleaning out the bottom of the via hole is critical. A residue of 400 Å of polymer at the bottom of the via is sufficient to create an open circuit [10]. Also, stress cracks at the corners of via holes, as shown in Figure 9-14, are a common problem.

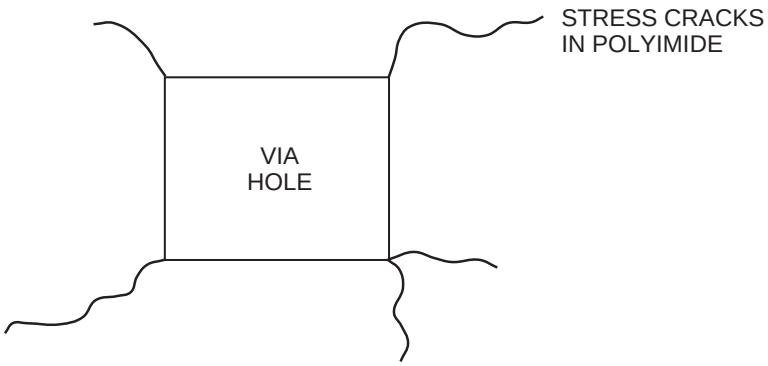

Figure 9-14. Stress cracks in polyimide at via holes.

Lastly, many of the MCM-D fabrication approaches, such as HDI [11], place polymers directly over the MMICs and other chips to be interconnected. Even though the polymers are thin relative to the die substrate and typically have permittivities of 3, they can have large effects on the microwave performance of the MMIC. Specifically, the polymer will increase the line capacitance, which decreases the guided wavelength. Therefore, distributed matching circuit elements will appear longer than the same structure without the polymer. Although the degree of circuit degradation is dependent on the transmission line type, substrate material and thickness, and characteristic impedance of the line, in general, coplanar waveguide (CPW) circuits will be effected more than microstrip circuits because of the greater field concentration at the substrate/polymer interface. If the MMIC is not designed to account for this frequency shift, the circuit performance will be degraded.

References

[1] Dupont Company product literature for Pyralin LX, data sheet # H-12287, April 1989.

[2] Dupont Company product literature for Pyralin PD, data sheet # H-32234, May 1991.

[3] Kodnani, D. Jaffe, and R. Jaccodine, "Evaluation of Polymeric Coatings for MCM Applications," *IEEE Trans. Comp., Packaging, and Manufact. Tech.–Part B*, Vol. 18, No. 4, pp. 658–665, November 1995.

[4] M. Pecht, X. Wu, K. W. Paik, and S. N. Bhandarkar, "To Cut or Not to Cut: a Thermomechanical Stress Analysis of Polyimide Thin-Film on Ceramic Structures," *IEEE Trans. Comp., Packaging, and Manufact. Tech.–Part B*, Vol. 18, No. 1, pp. 150–153, February 1995.

[5] V. Murali, T. Rucker, J. Fu, and R. Shukla, "Yield and Reliability Concerns in Polyimide Based Multi-Chip Modules," *1992 IEEE Multi-Chip Module Conf. Dig.*, pp. 98–101, March 18–20, 1992.

[6] K. P. Shambrook and P. A. Trask, "High-Density Multichip Interconnect (HDMI)," *Proc. 39th Electron. Components, Conf.*, pp. 656–662, 1989.

[7] J. H. Das and J. E. Morris, "Metal Diffusion in Polymers," *IEEE Trans. Comp., Packaging, and Manufact. Tech.–Part B*, Vol. 17, No. 4, pp. 620–625, November 1994.

[8] G. S. Frankel, S. Purushothaman, T. A. Petersen, S. Farooq, S. N. Reddy, and V. Brusic, "Corrosion and Adhesion of Multilayer Pad Structures for Packaging Applications," *IEEE Trans. Comp., Packaging, and Manufact. Tech.–Part B*, Vol. 18, No. 4, pp. 709–714, November 1995.

[9] R. S. Patel, T. A. Wassick, and C. Y. Ralston, " Metal Capping of MCM Thin Film Features Using a Laser," *IEEE Trans. Comp., Packaging, and Manufact. Tech.–Part B*, Vol. 17, No. 3, pp. 264–268, August 1994.

[10] G. White, E. Perfecto, D. McHerron, T. DeMercurio, T. Redmond, and M. Norcott, "Large Format Fabrication–a Practical Approach to Low Cost MCM-D," *IEEE Trans. Comp., Packaging, and Manufact. Tech.–Part B*, Vol. 18, No. 1, pp. 37–41, February 1995.

[11] R. A. Fillion, R. J. Wojnarowski, T. B. Gorcyzca, E. J. Wildi, and H. S. Cole, "Development of a Plastic Encapsulated Multichip Technology for High Volume, Low Cost Commercial Electronics," *IEEE Trans. Comp., Packaging, and Manufact. Tech.–Part B*, Vol. 18, No. 1, pp. 59–65, February 1995.

V. Plastic Package

R. N. Simons

Studies have shown that during the high-temperature soldering process encountered while mounting packaged semiconductor devices on circuit boards, moisture present in a plastic package can vaporize and exert stress on the package. This stress causes the package to crack and also causes delamination between the mold compound and the lead frame or die. These effects are most pronounced if the package has greater than 0.23% absorbed moisture before solder reflow [1]. Figure 9-15 shows a typical example of a package crack. The mismatch in thermal expansion coefficients of package's components also induces stresses. If these combined stresses are greater than the fracture strength of the plastic, cracks will develop. The cracks can provide a path for ionic contaminants to reach the die surface, and die delamination can cause wire-bond failure. Hence, these are reliability concerns.

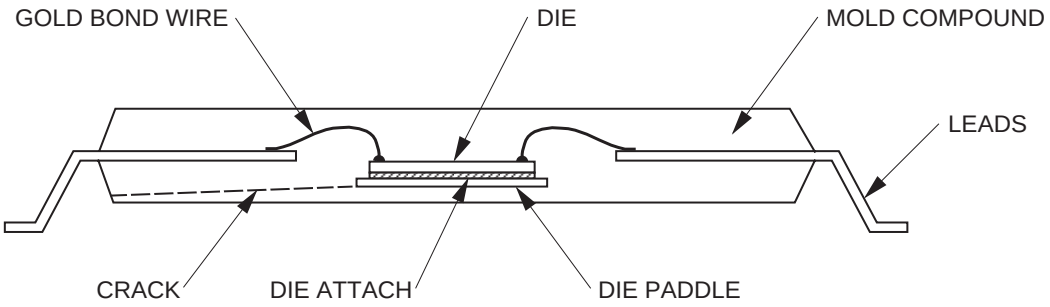

Figure 9-15. Typical plastic package showing the onset of a crack.

The following recommendations will ensure minimal package damage from moisture [2]:

(1) Complete board assembly 1 week after removal of the components from their dry packs, if the environmental conditions do not exceed 30°C and 60% RH.

(2) After 1 week, bake for a minimum of 12 h at 115°C; this will gradually drive out the moisture.

To overcome the delamination problem, results derived from numerical simulation and experimental data can serve as a guide in the selection of suitable molding compound properties [3]. The properties considered are the adhesion strength, *S*, and the coefficient of thermal expansion, α. These results are summarized in Figure 9-16. Also, it has been shown that polyimide die overcoat, or PIX, can reduce the percent of die or pad delamination by up to 30% on parts subjected to temperature cycling [1,4] as well as mechanically support air bridges during plastic encapsulation, provide a more uniform electrical environment for the die, and provide protection to the surface of the die. Figure 9-17 shows cross sections of three PIX-treated dies. It has been reported that the PIX shown in Figure 9-17(a) yields the best improvement in reliability [1]. The PIXs shown in Figures 9-17(b) and (c) are not as desirable, because, respectively, they cause wire-bond stress and do not protect the die surface.

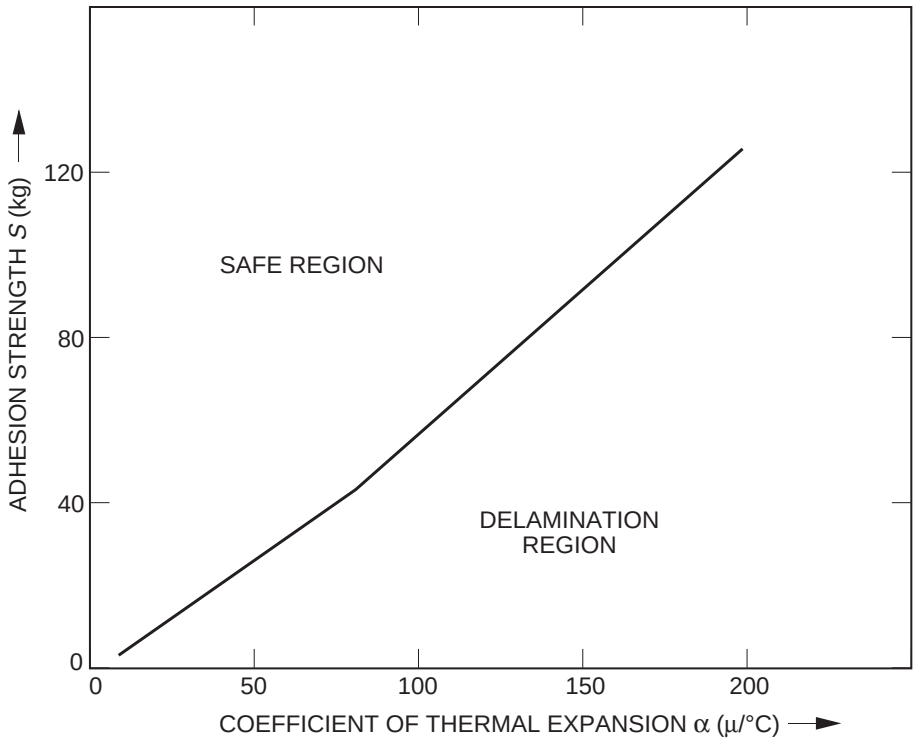

Figure 9-16. Mold compound properties.

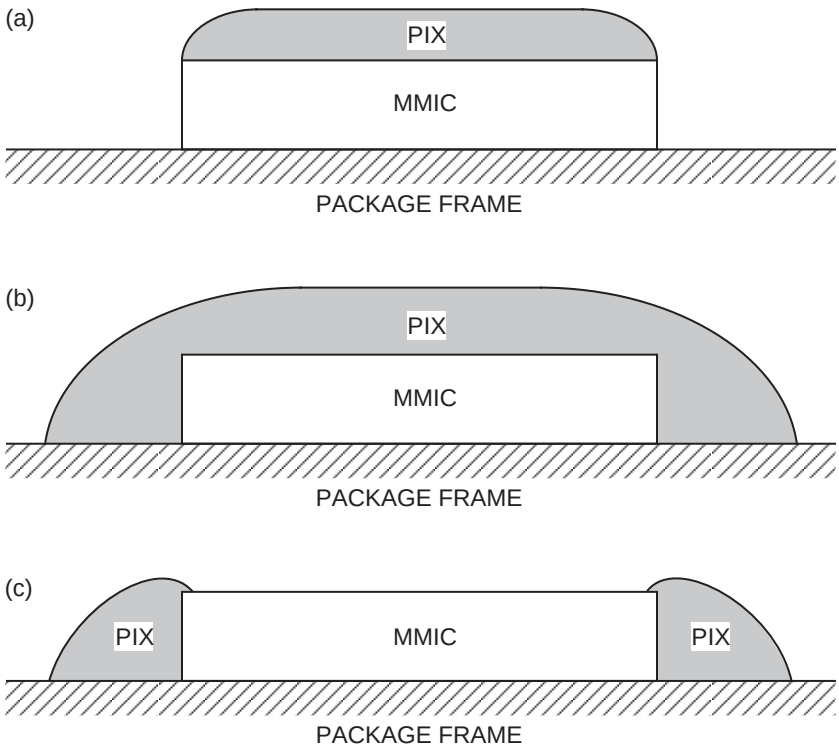

Figure 9-17. Polyimide die overcoat (PIX) on MMIC die: (a) PIX on MMIC top surface only, (b) PIX on MMIC and package frame, and (c) PIX on package frame and sides of MMIC only.

The last mechanisms by which a chip can fail in a plastic package are caused by bond-wire sweep and lift-off, which in turn are caused by the viscous flow of the molten-plastic mold compound. The viscosity of the molten plastic is a function of the filler particle size and concentration. Figure 9-18 shows the typical geometries of wire bonds with different die settings. Studies [5] show that of the three wire bonds, the one with the raised die experiences the largest maximum displacement. Further, the raised die and the downset die experience maximum stress at the ball bonds. In these cases, plastic deformation of the ball bonds is a major cause of failure. In contrast, the wire bond for the double-downset die suffers only elastic deformation. Thus, the double downset is the recommended device layout to minimize bond wire sweep.

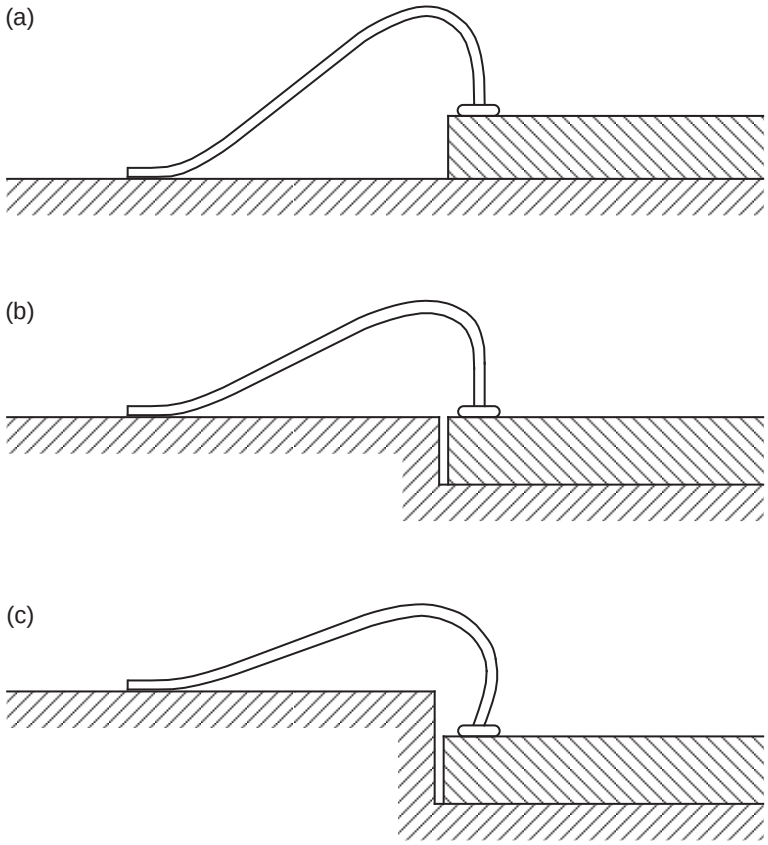

Figure 9-18. Typical geometry of wire bond with different die settings: (a) raised, (b) downset, (c) double downset.

References

[1] T. M. Moore and S. J. Kelsall, "The Impact of Delamination on Stress-Induced and Contamination-Related Failure in Surface Mount ICs," *1992 IEEE Annual Int. Reliability Physics Symp. Proc.*, pp. 169–176, March 31–April 2, 1992.

[2] L. Yip, "Moisture Sensitivity and Reliability of Plastic Thermally Enhanced QFP Packages," *IEEE Trans. Components, Packaging, and Manufacturing Technology-Part B*, Vol. 18, No. 3, pp. 485–490, August 1995.

[3] A. A. O. Tay, G. L. Tam, and T. B. Lim, “Predicting Delamination in Plastic IC Packages and Determining Suitable Mold Compound Properties,” *IEEE Trans. Components, Packaging, and Manufacturing Technology-Part B*, Vol. 17, No. 2, pp. 201–208, May 1994.

[4] W. J. Roesch and A. L. Rubalcava, “GaAs IC Reliability in Plastic Packages,” *1995 GaAs Reliability Workshop Programs and Abstracts*, pp. 25–30, October 29, 1995.

[5] A. A. O. Tay, K. S. Yeo, and J. H. Wu, “The Effect of Wirebond Geometry and Die Setting on Wire Sweep,” *IEEE Trans. Components, Packaging, and Manufacturing Technology-Part B*, Vol. 18, No. 1, pp. 201–209, February 1995.

VI. Package Resonance and Field Leakage

R. N. Simons

MMIC package performance degrades at millimeter-wave frequencies mainly as the result of ring resonances and cavity resonances. Ring resonances occur when stray electromagnetic fields couple to the ceramic frame of the package [1,2]. Cavity resonances occur when the volume enclosed by the package behaves as a rectangular metal cavity [3]. These resonances are observed as large spikes in the insertion-loss-versus-frequency characteristic of the packaged MMIC. The frame resonances are also responsible for the poor isolation between the input and output RF ports.

Ring resonances can be eliminated by fabricating the frame from metal. However, in general, intricate metal-frame shapes are difficult to fabricate and, therefore, very expensive. A low-cost approach is fabrication from a ceramic material, such as alumina. In this approach, several thin, punched, metallized green-ceramic layers are first stacked to form a frame. Second, the inside and outside vertical walls of the frame, except those areas around the RF signal and dc bias lines, are metallized. Third, the multilayer green frame is co-fired. Fourth, the frame is attached to a metal base and all conductor surfaces are electroplated. This approach is not only cost effective, but ensures that the frame is grounded to the metal base. Grounding the frame significantly reduces the stray coupling between the input and output RF ports. A package with five ceramic layers of this type reportedly has had resonance-free operation with frequencies up to 33 GHz [1]. A schematic of the package is shown in Figure 9-19. However, such packages can support strong cavity resonances if the inside dimensions are not properly chosen. A simple equation to predict cavity resonances is presented at the end of this section.

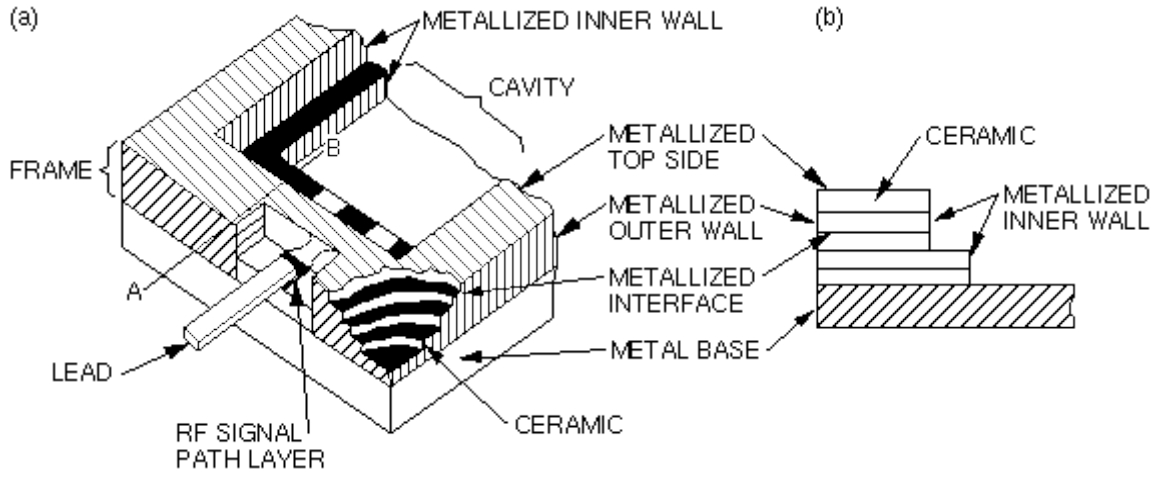

Figure 9-19. Multilayer ceramic package with metallized frame walls: (a) structure around terminal, (b) cross section A–B. (From [1]; ©1988 IEEE.)

An alternate approach in suppressing ceramic frame ring resonances is the periodic placement of metal-filled vias in the frame walls [4]; this replaces metallized surfaces. Figure 9-20 illustrates a Ka-band package with filled metal vias. This package, when experimentally characterized for return loss and insertion loss, shows that the frame resonances are not fully suppressed and they do occur over a narrow frequency band around 20 GHz, this because the metal filled vias are not as effective a shield as the

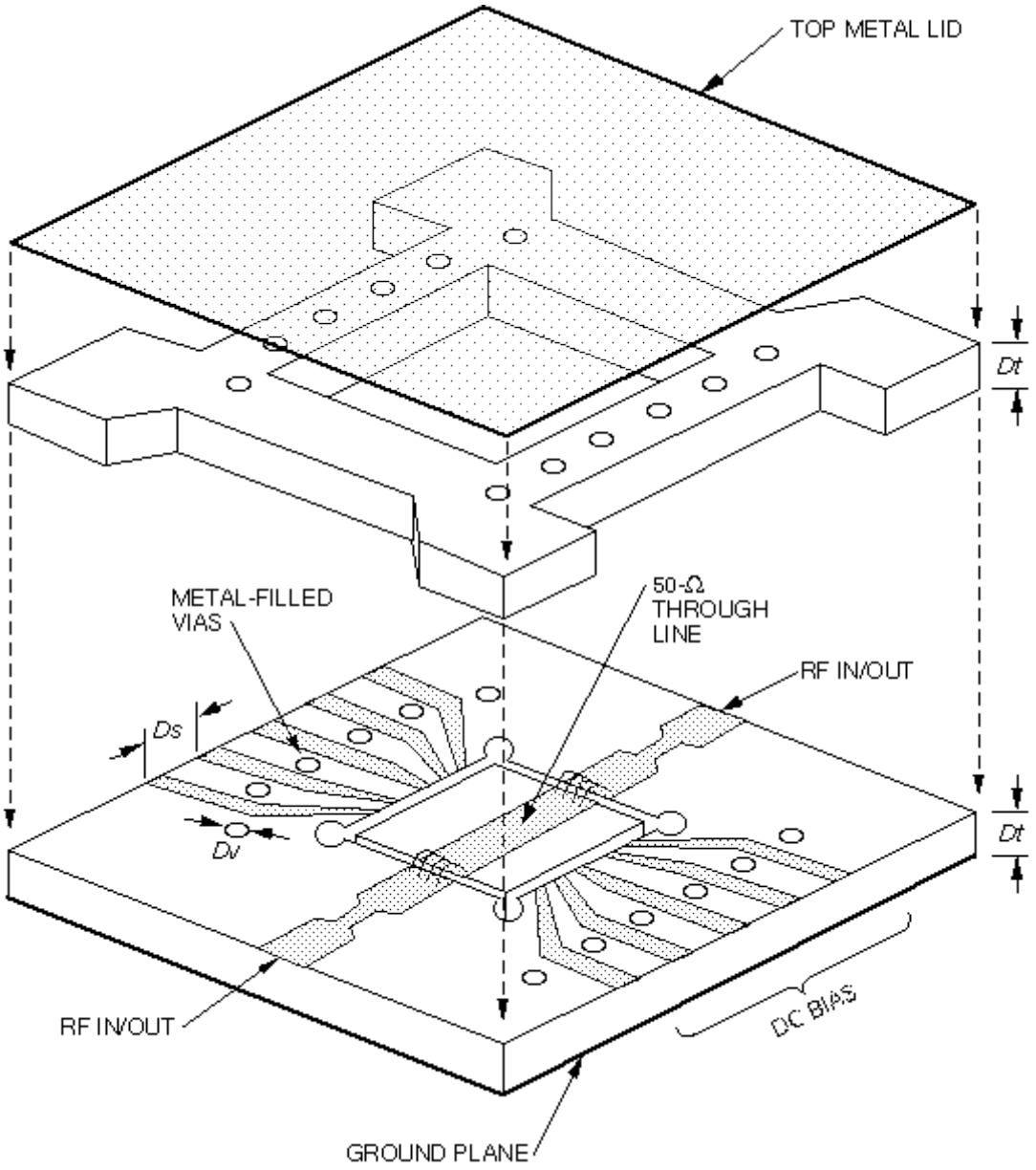

Figure 9-20. Ceramic package with metal-filled vias. (From [5]; ©1996 IEEE.)

metallized wall in the package shown in Figure 9-19. The measured and modeled return loss and insertion loss of the package is shown in Figure 9-21. A full-wave finite-element method (FEM) and numerical simulation of the package shows that a significant amount of RF energy leaks through the walls when the frame resonates [5]. Figure 9-22 shows the computed vertical electric-field distribution at about 20 GHz, as viewed from the top. The plot shows RF energy leaks on all sides through the gaps between the metal-filled vias. It is interesting to note that cavity resonances are not easily supported in this type of package because the side walls are not perfectly conducting—allowing RF energy to leak through the walls. The fields inside the package interact strongly with the ceramic material of the walls, and these fields are attenuated if the dielectric loss tangent is large.

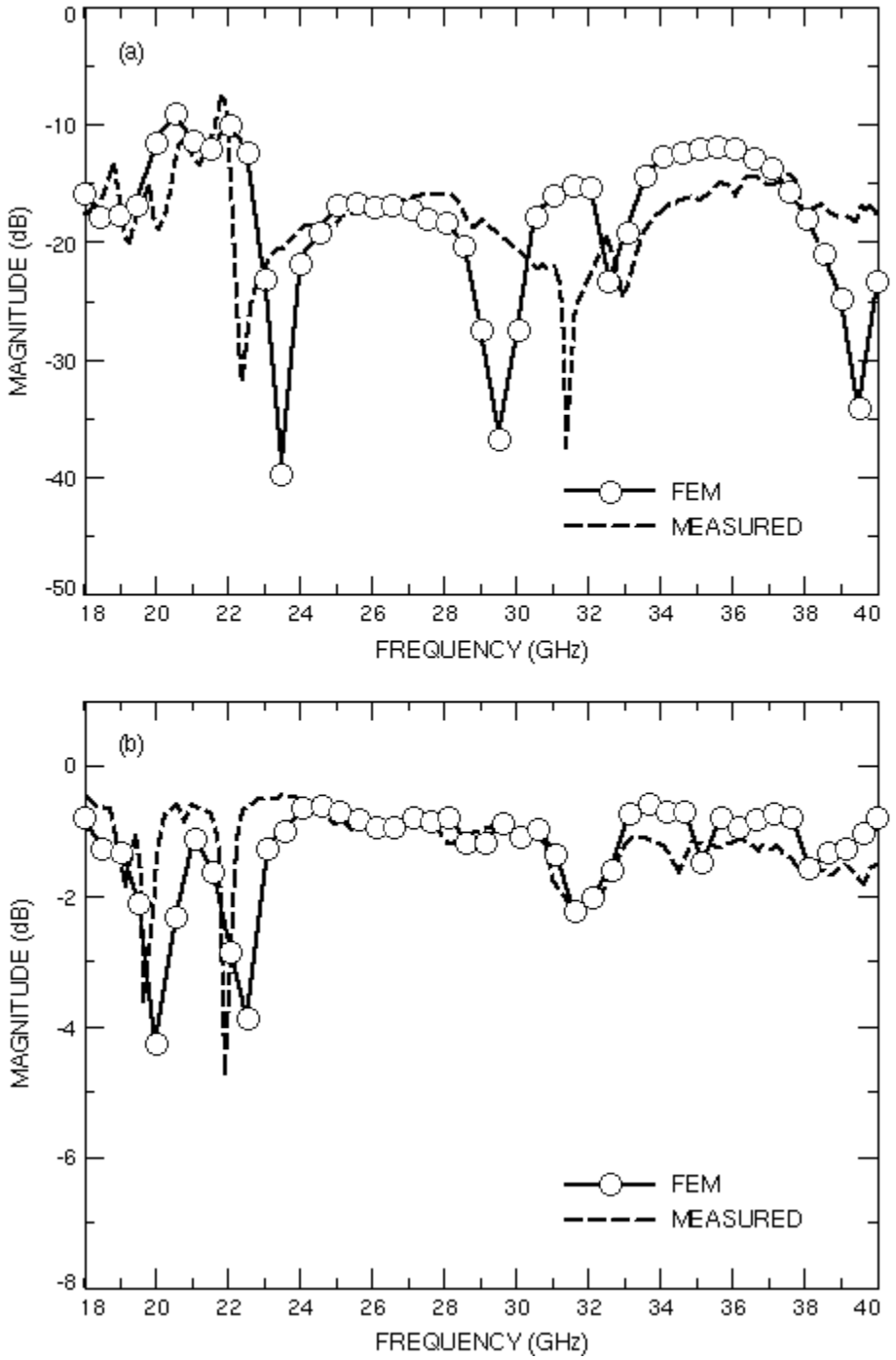

Figure 9-21. Measured and modeled *S*-parameters: (a) S11, (b) S21. (From [5]; ©1996 IEEE.)

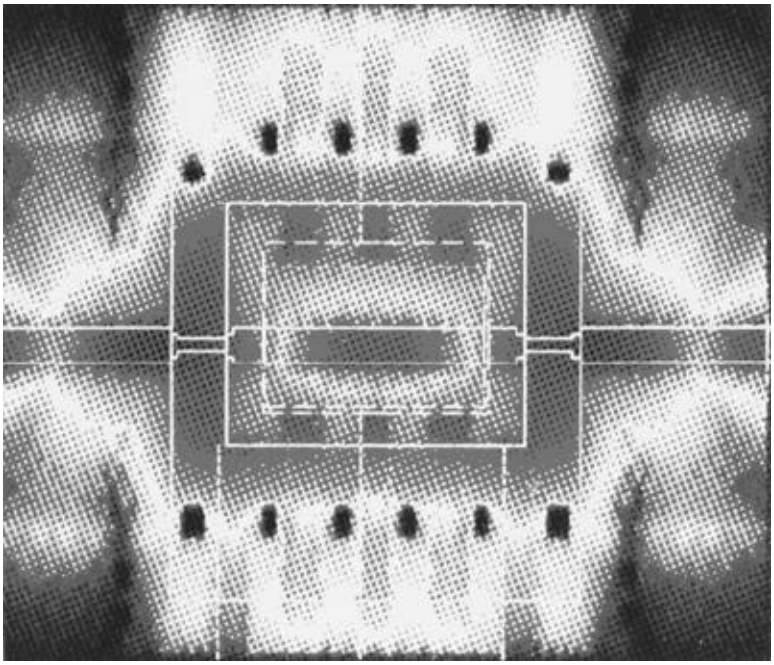

Figure 9-22. Computed vertical electric-field distribution.

The cavity resonances are predicted from a model [3] that considers the package as a rectangular metal cavity loaded with a H-plane dielectric slab, as shown in Figure 9-23. In this model, the length, width, and height of the cavity are represented as L, W, and H, respectively. The dielectric slab is of thickness d and relative permittivity ε_r. The cavity is excited by a microstrip line at the input and output ports. The resonance frequency f_r is approximately given by

$$f_r = f_c \sqrt{1 - \left(\frac{d}{H}\right) * \left(\frac{\varepsilon_r - 1}{\varepsilon_r}\right)}$$

where f_c is the cut-off frequency of the TE_{101} mode in the empty cavity. As an example, if $\varepsilon_r = 13$, $d = 0.01016$ cm (0.004 in.), $L = 0.4064$ cm (0.16 in.), $W = 0.254$ cm (0.1 in.), and $H = 0.04826$ cm (0.019 in.), then f_c and f_r are 69.641 GHz and 62.508 GHz, respectively.

Even when the package is designed to avoid in-band resonances and to suppress field leakage, electromagnetic effects may still degrade the circuit performance. To understand this MMIC performance degradation, first recall that microwave transmission lines such as microstrip and coplanar waveguide (CPW) do not confine the electromagnetic energy to a finite volume, but simply guide it along the path of the line. The amount of energy that spreads beyond the microstrip or slots of the CPW is dependent on the substrate thickness, permittivity, and the geometry of the transmission line. When a MMIC is tested on a wafer probe station, this energy that is not well confined to the guide tends to radiate outward from the line and dissipate, but if that same MMIC is placed in a package, the unconfined energy reflects off of the package walls and

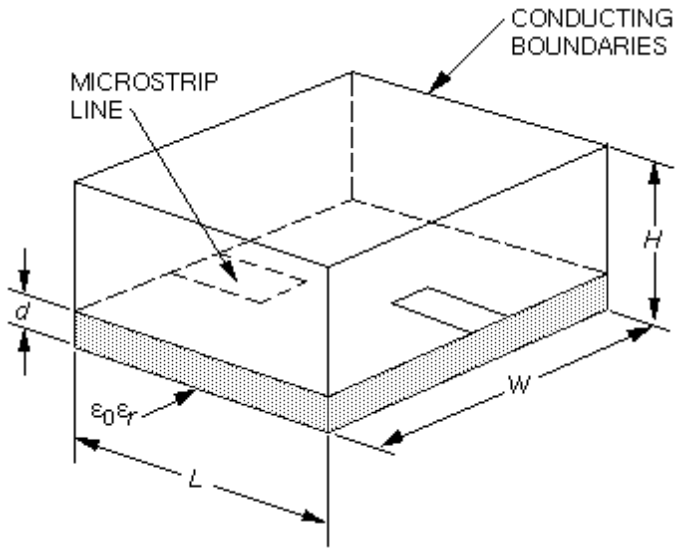

Figure 9-23. Partially filled metal cavity with microstrip input/output ports.

recombines with the energy traveling down the transmission line. The result is a distortion of the transmitted signal [6]. Therefore, it is recommended that MMICs be tested in their package as well as subjected to the on-wafer RF characterization.

References

[1] F. Ishitsuka and N. Sato, "Low Cost, High-Performance Package for a Multi-Chip MMIC Module," *Technical Dig. of IEEE GaAs IC Symposium Dig.*, pp. 221–224, October 1988.

[2] F. Ishitsuka, N. Sato, and H. Kato, "A New Multilayer Ceramic-Frame Package for High-Frequency MMIC Modules," *IEICE Trans.* (Japan), (special issue on *Advanced Packaging Technology for Microelectronics Manufacturing*) Vol. E 74, No. 8, pp. 2344–2347.

[3] G. H. Robinson, "Resonance Frequency Calculations for Microstrip Cavities," *IEEE Trans. Microwave Theory Tech.*, Vol. MTT 19, No. 7, pp. 665–666, July 1971.

[4] K. Li and S. Yuan, *MMIC Packaging Technology Development*, Final Technical Report for NASA Contract No. NAS3-25864, Hughes Aircraft Co., March 15, 1993.

[5] J-G. Yook, L. P. B. Katehi, R. N. Simons, and K. A. Shalkhauser, "Experimental and Theoretical Study of Parasitic Leakage/Resonance in a K/Ka-Band MMIC Package," 1996 IEEE MTT-S Inter. Microwave Symp. Dig., June 1996.

[6] M. Tsuji and H. Shigesawa, "Packaging of Printed-Circuit Lines: a Dangerous Cause for Narrow Pulse Distortion," IEEE Trans. Microwave Theory Tech., Vol. 42, No. 9, pp. 1784–1790, September 1994.

Additional Reading

Johnson, E. F., "Technique Engineers the Cavity Resonance in Microstrip Housing Design," *MSN & CT*, Vol. 17, pp. 100–102, 107–108, and 126, February 1987.

Williams, D. F., "Damping of the Resonant Modes of a Rectangular Metal Package," *IEEE Trans. Microwave Theory Tech.*, Vol. 37, No. 1, pp. 253–256, January 1989.

VII. Hydrogen Poisoning of GaAs MMICs in Hermetic Packages

A. Immorlica and S. Kayali

Microwave packages and modules typically employ iron- and nickel-based alloys and plated layers that contain or use hydrogen in the manufacturing process. Usually, the hydrogen will outgas and not cause a problem, but in a hermetically sealed package, hydrogen can reach partial pressures as high as a few percent [1]. Hydrogen is known to cause degradation of some types of GaAs devices, and hydrogen-induced degradation has been reported over large ranges of hydrogen partial pressures, even down to a few hundred millitorr [2,3].

The degradation of GaAs MESFETs and PHEMTs in hydrogen atmospheres has commanded significant attention over the past several years [4]. The effect was first reported in MESFETs by Camp et al [5] in 1989, and has been observed more recently by others in MESFETs [6,7,8], PHEMTs [2,3,9], and InP HEMTs [10]. This is now a recognized industry-wide problem, particularly for devices incorporating Schottky barrier gates having Pt or Pd, which are widely used.

The poisoning of GaAs devices is manifested by a sudden and dramatic change in device electrical properties, which can occur after several hundred to several thousand hours of hydrogen exposure at elevated temperatures. This is illustrated in Figure 9-24, which shows the percent change in pinch-off voltage of MESFETs over 700 hours [6]. The time for onset of this degradation is dependent on the device technology, which varies greatly among suppliers, and the partial pressure of hydrogen to which the device is exposed.

Several degradation mechanisms have been proposed. Camp et. al. [5] suggested a compensation model in which silicon atoms, a typical n-type dopant in GaAs, are neutralized by hydrogen, causing a loss of channel conductivity and current. Others have proposed models in which the Schottky barrier contact potential is changed, causing a shift in the device pinch-off voltage and transfer characteristics. The latter model appears to be particularly appropriate to PHEMTs. Regardless of the model, almost all reported degradation has been associated with devices having refractory metal gates containing Pt or Pd. It has been theorized that these commonly used gate metals act as catalysts, converting molecular hydrogen to atomic hydrogen, greatly enhancing the degradation reactions. It is interesting to note that there has been no reported degradation for non-refractory gates, such as those made of Al.

The activation energy associated with hydrogen poisoning has been reported to be as low as 0.4 eV [2,7]. This is relatively low compared to the 1.2- to 1.8-eV reported mechanisms in GaAs. Thus, care must be exercised when interpreting or extrapolating accelerated life-test data taken at high temperatures.

A number of options for dealing with this problem have been suggested in the literature. These include

(1) Elimination or minimization of the hydrogen source.

(2) Changing device technology.

(3) Use of an in-package hydrogen getter.

(4) Circuit compensation for device electrical changes.

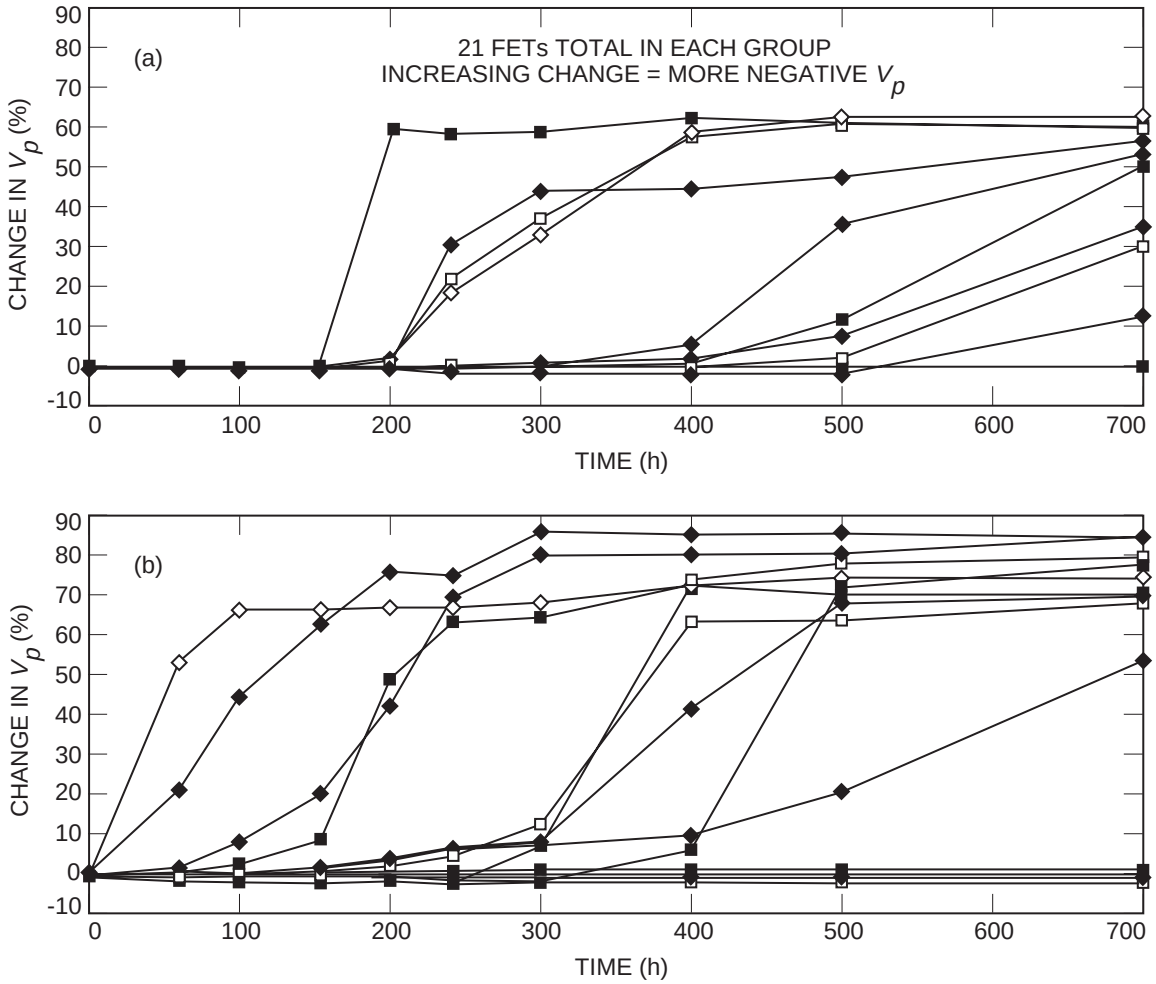

Figure 9-24. Percent change in pinch-off voltage for (a) palladium gate and (b) platinum gate FETs in high hydrogen at 225°C. (From [6].)

Hydrogen is present in many of the common packaging materials used in the microwave industry today. Materials known to outgas hydrogen include Kovar™, Ni and Au plating, ferrite circulators, other iron based materials, and even some RF-absorber materials used for circuit stabilization. Hydrogen can be reduced to some extent by vacuum baking the package parts; however, care must be taken to not impact other package properties, such as solderability. Alternatively, hydrogen can be reduced by the judicious choice of materials that have low hydrogen solubility, such as aluminum, but it is difficult to completely eliminate all hydrogen-bearing materials. (Note that non-hermetic packaging does not pose a poisoning threat, as the leak rate of hydrogen would be sufficient to keep concentrations down to safe levels).

A second approach modifies the device technology with the use of a gate metal that is hydrogen insensitive. As mentioned above, aluminum is one such candidate, and GaAs laboratories are working on other hydrogen-insensitive schemes. While this may be an adequate approach for a relatively new technology such as InP HEMTs, changing processes in relatively mature technologies, which have a heritage of use and field experience, is not usually favorable.

A third approach uses an in-package hydrogen getter to reduce hydrogen partial pressures to safe levels. The use of getters in semiconductor packaging for substances such as water vapor and particulates is common, and there are several commercial as well as proprietary hydrogen getters that can be employed in microwave packaging. To be effective, one must determine the sensitivity of the device technology to hydrogen, and ensure that the getter has adequate capacity to maintain a safe hydrogen partial pressure during the expected mission lifetime of the device.

Finally, the hydrogen sensitivity problem might be circumvented by appropriate circuit design. One generally defines device "failure" as a given percentage change in one or several device parameters. Depending on the circuit design and application, this change may or may not cause failure of the device to meet its intended function. Alternatively, it is sometimes possible to re-bias the circuits to recover the initial conditions, if the device change is caused by a shift in parameters. The nature of the device changes would, of course, have to be well understood.

References:

[1] P. Schuessler and D. Feliciano-Welpe, "The Effects of Hydrogen on Device Reliability," *Hybrid Circuit Technology*, pp. 19–26, January 1991.

[2] W. W. Hu, E. P. Parks, T. H. Yu, P. C. Chao, and A. W. Swanson, "Reliability of GaAs PHEMTs under Hydrogen Containing Atmosphere," *Proceedings GaAs Integrated Circuits Symposium*, pp. 247–250, 1994.

[3] S. B. Adams, J. A. MacDonald, W. W. Hu, A. A. Immorlica, A. R. Reisinger, and F. W. Smith, "Reliability of GaAs PHEMT MMICs in Hydrogen Ambients," GaAs Reliability Workshop, Philadelphia, October 1994.

[4] S. Kayali, "Hydrogen Effects on GaAs, Status and Progress," GaAs Reliability Workshop, San Diego, October 1995.

[5] W. O. Camp Jr., R. Lasater, V. Genove, and R. Hume, "Hydrogen Effects on Reliability of GaAs MMICs," *Proceedings GaAs Integrated Circuits Symposium*, pp. 203–206, 1989.

[6] W. J. Roesch, "Accelerated Effects of Hydrogen on GaAs MESFETs," GaAs Reliability Workshop, Philadelphia, October 1994.

[7] M. J. Delaney, T. J. W. Tsey, M. Chiang, and K. K. Yu, "Reliability of 0.25 μm GaAs MESFET MMIC Process: Results of Accelerated Lifetests and Hydrogen Exposure," GaAs Reliability Workshop, Philadelphia, October 1994.

[8] K. Decker, "GaAs MMIC Hydrogen Degradation Study," GaAs Reliability Workshop, Philadelphia, October 1994.

[9] G. Kelley, M. Cobb, M. Welch, M. Weig, and D. Weir, "Effects of Temperature and Concentration on Hydrogen Degradation of Pseudomorphic GaAs HEMTs," GaAs Reliability Workshop, San Diego, October 1995.

[10] P. C. Chao, M. Y. Kao, K. Nordheden, and A. W. Swanson, "HEMT Degradation in Hydrogen Gas," *IEEE Electron Device Letters*, Vol. 15, pp. 151–153, May 1994.

Acronyms and Symbols

e0	vacuum permittivity
er	relative dielectric constant
h	efficiency
l	wavelength
m	mobility
v	electron drift velocity
*v*sat	saturated electron velocity
t	time or lifetime
fb	barrier potential
fm	metal work function
c	electron affinity
2DEG	two-dimensional electron gas
A	emitterbase junction or diode area; proportional multiplier
a	physical channel depth
ADC	analog-to-digital converter
Al	aluminum
As	arsenic
Au	gold
Be	beryllium
BER	bit error rate
BJT	bipolar junction transistor
b(*x*)	effective FET channel depth
C	carbon
CAD	computer aided design
CAE	computer aided engineering
*C*CS	collector-substrate capacitance
*C*DC	drain-channel capacitance
*C*F	total forward biased junction capacitance
*C*j	junction capacitance; depletion region capacitance
CMOS	complimentary metal oxide semiconductor (transistor)
Cr	chromium
CTE	coefficient of thermal expansion
*C*v	capacitance of insulating region
dB	decibels (dB = 10 log[power])
DBS	direct broadcast satellite
*D*n	electron diffusivity

DNP	distance neutral point
*D*p	hole diffusivity
*E*a	activation energy
E-B	emitterbase
*E*c	conduction band energy
EDM	empirical device model
*E*f	Fermi energy
*E*g	bandgap energy
*E*m	breakdown electric field
EOS	electrical overstress
ESD	electrostatic discharge
*E*v	valance band energy
eV	electron volt
*f*c	forward current cutoff frequency
fcc	face centered cubic
FEM	finite element method
FET	field effect transistor
FIT	failure in time (1 failure/1 ´ 109 device h)
*f*max	maximum frequency of oscillation
*f*t	cutoff frequency or frequency where unilateral power gain equals one
f(t)	failure rate
g	gram
GaAs	gallium arsenide
GCR	galactic cosmic rays
Ge	germanium
GEO	geostationary orbit
*g*m	transconductance
GPS	Global Positioning System
h	hour
HBT	heterojunction bipolar transistor
HDI	high density interconnect
HEMT	high-electron mobility transistor
*h*fe	current gain
I	current
*I*D	drain current
*I*DSS	drain-source saturation current
IF	intermediate frequency
IMPATT	impact ionization avalanche transit time (diode)

In	electron injection current
In	Indium
IR	infrared
J	current density
JFET	junction field effect transistor
J0	constant depending on doping concentration
K	Kelvin
k	Boltzman's constant
Ka	frequency band (26.540 GHz)
Ku	frequency band
L	length
LEC	Liquid Encapsulated Czochralski
LEO	low Earth orbit
LET	linear energy transfer
LNA	low-noise amplifier
LO	local oscillator
LTCC	low-temperature cofired ceramic
MBE	molecular beam epitaxy
MESFET	metal-semiconductor field effect transistor
MIC	microwave integrated circuit
MIM	metal-insulator-metal (capacitor)
MMIC	monolithic microwave integrated circuit
MOCVD	metal-organic chemical vapor deposition
MODFET	modulation doped field effect transistor
MTBF	mean time between failure
MTTF	mean time to failure
n	ideality factor
NA	acceptor impurity density
NB	base doping concentration
ND	donor impurity density
Nd	donor doping concentration
NE	emitter doping concentration
NF	noise figure
Ni	nickel
NPN	n-typep-typen-type (transistor)
$n(x)$	electron density
P(s)	probability of success
PBM	physically based model

PCS	personal communication system
Pd	palladium
PHEMT	pseudomorphic high-electron mobility transistor
PIN	p-typeinsulatorn-type (diode)
PIND	particle impact noise detection
PIX	polyimide die overcoat
PM	parametric monitor
PNP	p-typen-typep-type (transistor)
*P*out	output power
Pt	platinum
Q	quality factor
q	charge of an electron
*Q*c	critical charge
QML	Qualified Manufacturers Listing
r	rate of a process
*R*b	base resistance
*R*D	drain resistance
*R*DS	channel resistance between drain and source
RF	radio frequency (typically refers to highest frequency in the circuit)
*R*F	total forward biased series resistance
RHA	radiation hardness assurance
RIE	reactive ion etch
*R*j	junction resistance
*R*ohm	ohmic contact resistance
*R*S	source resistance
R(*t*)	reliability-the probability of a component surviving to time *t*
RTG	radioisotopic thermoelectric generator
*R*v	resistance of insulating region
SAA	South Atlantic Anomaly
SCR	silicon-controlled rectifier
SEB	single event burnout
SEC	standard evaluation circuit
SEE	single event effect
SEGR	single event gate rupture
SEL	single event latchup
SEM	scanning electron microscope
SEU	single event upset
Si	silicon

SOI	silicon on insulator
SOS	silicon on sapphire
SPC	statistical process control
T	absolute temperature in Kelvin
t	time or carrier transit time
TCV	technology characterization vehicle
TEGFET	two-dimensional electron gas field effect transistor
TID	total ionizing dose
TRB	Technology Review Board
V	voltage
*V*a	early voltage
*V*BE	baseemitter voltage
*V*bi	built-in voltage
*V*CF	potential between conduction band and Fermi level
VCO	voltage-controlled oscillators
*V*D	drain bias voltage
*V*DS	potential between drain and source
*V*G	gate bias voltage
VLSI	very large scale integration
*V*p	pinch-off voltage
*V*T	thermal potential
W	width
W	frequency band (75110 GHz); tungsten
WLAN	Wireless Local Area Network
X	frequency band (812 GHz)
Z	gate width; width of channel