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.