Electrotehnica, Electronică, Automatică (EEA), 70 (2022), nr. 3, pp.35-46
https://doi.org/10.46904/eea.2022.70.3.1108004

**Power FETs: Failure Analysis (FA) and Reliability Analysis (RA)**

**Titu-Marius I. BĂJENESCU**

1 Prof. Eng., Doctor Honoris Causa of Military Technical Academy of Romania and of Technical University of Republic of Moldova (Chișinău); Romanian Academy „Tudor Tănăsescu” Prize Laureate

**Abstract**

FET’s have become increasingly popular as possible replacements for microwave electron tubes and solid-state active diodes. Reliability is extremely important for these devices when used in communication systems, especially for space applications. The reliability of small, signal GaAs FET’s has been extensively investigated. As a result, the reliability of small-signal or low-noise devices is fairly well understood, and the essential problems are practically solved. However, the reliability of power devices is much more complicated because they are expected to operate in the vicinity of their maximum capability of voltage, current, or power dissipation, in the presence of large RF signals.

**Keywords:** Power GaAs, SiC, GaN FETs, GaAs, HEMT, failure mechanisms, failure analysis, reliability analysis, IRTS, computed tomography, space-domain reflectometry (SDR), security.

**Received**: day Month 2019

To cite this paper:


**Introduction**

The explosive development of the semiconductor technology imposes greater demands on the quality and reliability of the components.

The problem of the failure rate during the life of a device or system is more and more important and the failure analysis helps to the improvement of the product quality. The power transistor is an important element of the interface between the command electronics and the elements of the power electronics. The greater currents and voltages led to new absolute limit values for the dissipated power of these components.

Outside of the well-known thermal power dissipation corresponding to the operating state, there are some specific limits for bipolar transistors: a limit for pulse operating and another for second breakdown. At first sight, it seems that the quality is a relative notion for a power transistor, but this relativism can be overcome by defining the electrical properties.

In practice, the data sheet is the simplest quality certificate: smaller is the number of bad components, higher is the quality level. Usually, the manufacturer establishes a maximum value that can be reached by the silicon crystal in the operating life. This value arises from the quality and reliability criteria, and it is justified by the risk of contamination for the oxide or for the protecting layers of the active junction area and it takes into account the solder “fatigue”, produced by the mechanical dilatation stresses.

Generally, the value specified for this temperature varies from 200 °C for metal encapsulated transistors, to +150 °C for plastic encapsulated ones [1], [2].

Because the crystal surface has some tens of mm2, one must admit that the specified limit value for the junction temperature is overreached in some crystal points, even if the average value is bellowing this limit.

Experimentally, it has been shown that the transistors can be destroyed in some operating points, even if the limit given by the dissipation hyperbole is not surpassed. This means that the concept of maximum junction temperature is insufficient for safety using the power transistors.

Consequently, the producers perform testing at power limit, the transistors being commanded through the base, in the purpose to attain an operating point at the limit specified by second breakdown.

Recent advancements in the manufacturing of power transistors allow extending their application field. Due to the values obtained for currents, voltages and switching speed, electronic systems operating in the power field of some kW were realised. The technical problems raised by the
circuits with power transistors are simple and allow small manufacturing costs, small dimensions, and small weight. Moreover, high frequency operation produces fewer disturbances to the power supply than classical device [3].

Generally, these components work in supply circuits having small source impedance.

Therefore, overvoltage with multiple causes can come out [4]. For instance, voltage peaks arise in the inductive circuit at a current-off or due to the disturbances transmitted on-line. These overvoltages represent a great danger for the semiconductor device in an off state, because in this case the device acts like a dielectric. This phenomenon occurs mainly for the components without a “controlled avalanche” characteristic and, particularly, for transistors sensitive to second breakdown. One must realise that the energies implied are very important and their suppression is difficult because of the reduced source impedance. The protection device, preventing the reaching of a dangerous value of the circuit voltage, must satisfy the following conditions:

- Do not cause losses in normal operating,
- To be an effective limiter for the overvoltages with a rapid on-characteristic,
- To swallow-up, without being destroyed, the delivered energies.

Today, one knows that the difficulties encountered with power transistors arise from bad using conditions. That means that the most important technological problems linked to the component reliability have been solved. Researches on the operating conditions (detailed definition of the specifications on protection, correct choice of components and the avoidance of the errors in use) must be done. Because their essential function is linked to high energy levels, the practical using conditions for power transistors are a peremptory factor in the quality of the systems using power transistors.

Experimentally, it has been shown that the lack of this information always leads to failures.

Consequently, the user must ask as follows:

- Has the transistor complete and correct specifications? Is it correctly mounted, compiling with these specifications?
- If yes, an inherent reliability for the specified using mode can be obtain?

Power GaAs (FETs)

In our experience, the dominant failure mode of power GaAs FET’s has been catastrophic burnout. This burnout is defined as a sudden failure process leading to the melting of the metallization of one or more cells of a device. It has recently been shown that a large group of censored data for a failure process lacking a precursor can be statistically treated. This treatment has demonstrated that the burnout obeys the lognormal failure distribution. The burnout mechanism has been investigated by several workers presenting various explanations.

In accordance with these investigations, the structure, material, and fabrication process of power GaAs FET’s have been modified toward possible elimination of burnout. There are two types of burnout.

First, burnout takes place during the application of bias voltages which is called “instantaneous burnout”. A collaborative effort was made to raise the drain-source instantaneous burnout voltage above 35 V at normal operating current.

Secondly, burnout occurs in the course of aging, which is designated as “long-term burnout”. This burnout is associated with gradual degradation in the burnout voltage.

The present generation of our power GaAs FET’s has been improved to reduce this mechanism. Along with the development of this improved model of power GaAs FET’s, a large-scale experiment on the reliability of these devices was initiated.

The proverbial needle in the haystack – locating a minute process defect or subtle ESD strike in a large sea of analogue output power FETs can be just that.

We discuss failure analysis techniques used to identify these elusive “needles”, specifically in large array power FET structures.

ICs, with a significant portion of the die area (>50 %) populated with power FET structures, can serve as submicron defect monitors for the wafer fab. Statistically, if a FET related process defect is to occur, there is a high probability it will occur in the dominant output FETs.

Inline defect screening and comprehensive test coverage contain the majority of manufacturing defects. However, latent defects resulting in device failure post final test do occur and require subsequent product analysis. Conveying accurate and timely failure analysis information back to the wafer fab for process improvement is critical.

Analogue output power FET structures are often realized as an array of power FETs connected in parallel, functioning collectively as a single three terminal device.

Common configurations are high-side and low-side output drivers, which are used to switch electric loads.

The output driver of an IC fails in an application when the output loses the ability to fully turn-on (high-side) or fully turn-off (low-side). This inability is commonly diagnosed during ATE (Automated Test Equipment) testing as a high RDS_{on} failure.
Often, a high RDS$_{on}$ failure is attributed to some form of gate leakage, either gate-source or gate-drain.

Traditional fault isolation techniques are typically effective for G-S and G-D leakage failures, however precise fault isolation becomes more challenging in FET arrays because there are multiple power FETs connected in parallel, each power FET with numerous gate fingers. A gate-to-source leakage path for example, can exist on any one of the many gate fingers on any one of the power FETs in the output array. Accurate electrical localization becomes crucial for finding physical evidence of the failure mechanism. years or so).

**Failure analysis (FA)**

Today, FA is the key-method in reliability analysis. It is impossible to conceive of a serious investigation into the reliability of a product or process without FA.

The idea that failure acceleration by various stress factors (which is the key to accelerated testing) could be modelled only for the population affected by a single failure mechanism (FM) greatly promoted. FA as the only way to separate populations damaged by specific failure mechanisms (FMs).

A large range of methods are now used, from (classical) visual inspection to such expensive and sophisticated methods as atomic force microscopy and scanning near-field optical microscopy. Many others are still waiting to be created [1].

Recently, through device shrinking and complexity growing, it has become more and more difficult to carry out FA of semiconductor devices. A recent prediction [2] of the Semiconductor Industry Association (SIA), made in the International Technology Roadmap for Semiconductor (ITRS)$^1$, says that silicon technology will continue its historical rate of advancement predicted by the Moore’s law. So, the challenges for FA in the next few years will be extended to broad new aspects:

- design for analysis or design for test;
- physical limit-tools for chip, tools for package;
- chip-package co-design;
- organisational issues like FA cost and FA cycle time.

Above all, it is clear that failure diagnosis and failure position analysis have to increase in accuracy.

*Hitachi High Technologies, Ltd.* has developed an extremely fine SEM-type mechanical proving system [3].

For example, a high-precision probe and stage mechanism corresponding to the fine devices, a six-probe mechanism expandable to applications including inverter testing and a high-precision unit transistor testing were all investigated, together with an in-vacuum probing, sample-exchanging mechanism, and a computer aided design (CAD) navigation system.

As this system was applicable to 65 nm devices, it seems likely it will be possible to apply it to any device of such a scale in the future.

The microsystems are relatively new devices, containing, on a single chip, a mixture of components—a sensor, an actuator (a mechanical component) and the electronics, which creates new challenges for their reliability [4].

The package should protect the chip from an often harsh and demanding environment (as for the “classical” microelectronic devices: transistors, ICs, etc.), but is also an interface between the sensor and that environment. The small dimensions of the mechanical elements of the actuator produce new FMs and the interactions between mechanical, electrical and material reliability must be taken into account.

Moreover, the third dimension (the depth) of the structure cannot be ignored, as occurred for microelectronic devices, where all the simulations are basically two-dimensional. This collection of reliability risks has to be taken into account when FA is performed for any type of microsystem.

It should be noted that the subject is common in papers today, but we still have limited knowledge about how microsystems fail. The main explanation is a lack of specific tools for studying microsystems.

Fabricating multiple devices on the same chip will have to deal with more failure modes (FMos).

Complex interactions of cross-domain signals, interference and substances induce new FMos and FMs.

Another challenge for FA comes from a new domain, called nanotechnology.

Here, everything is new and the FMs for nanomaterials, which are different from those for the same materials at micro level, have to be studied. Supplementary issues are induced by organic materials, which is a new trend in this field. Also, at nano level, new techniques for FA have to be created [5].

As one can see, nano-reliability (study of the reliability of nano-devices) offers a huge range of subjects for FA.

---

$^1$ The ITRS is a forecast of the future development of semiconductor technology that is considered a red thread for chip and device manufacturers. It is updated every two years by a panel of global experts from the semiconductor industry. In addition, a small update takes place in the intermediate years.
The near future will see an important step forward in this field.

In conclusion, it is both easy and difficult to predict the future evolution of FA. Easy because everyone working in this domain can see the current trend. FA is still in a “romantic” period, with fabulous pictures and smart figures smashing customers, convinced by such a “scientific” approach.

Seldom, these users of electronic components understand the essence of the FA procedure, because the logic is frequently missing. But this situation is only a temporary one.

Very soon, the procedures for executing FA will be stabilised and standardised, allowing any user of an electronic component to verify the reliability of the purchased product.

But it is also difficult to predict the evolution of FA, because the continuous progress in microelectronics and microtechnology makes it almost impossible to foresee with good accuracy the types of electronic component that will be most successful on the future market. And FA must serve this development, being one step ahead and furnishing manufacturers with the necessary tools for their researches. However, with sufficiently high probability, one may say that nano-devices (or even nano-systems) will become a reality in the next few years, so we must be prepared to delve deeper into the matter, with more and more expensive investigation tools.

Commonly performed failure analysis “triage” activities including package optical inspection, pin curve-tracing, and X-ray inspection (both 2D and 3D) provided no insight into the failures, as might be expected given the subtlety of the observed failure.

The well-established best practices of component failure analysis require both non-destructive and destructive analysis techniques. Of course, we perform as much non-destructive analysis as could be meaningful, but generally, we end up having to perform analytical steps that are progressively more destructive to learn the information necessary to determine the root cause of failure, its scope, and corrective actions. With each destructive step, we analysts ask ourselves, “In pursuing this evidence, am I destroying other evidence that may be useful or vital?” - and then proceed to do as we must.

Failure analysis of electronic components is almost always destructive, and there is no going back once a destructive step is performed. Or that’s the way it used to be, before the development of some of the more sophisticated non-destructive techniques such as computed tomography (CT).

The trend to higher integration of electronic devices in order to include more and more functions into ever smaller devices, such as mobile phones or tablet computers, drives the development of new packaging technologies for semiconductor chips.

The goal is to reduce packaging size and power consumption while at the same time allowing more processing power.

One of the approaches is to stack multiple silicon chips on top of each other, i.e. pack the corresponding DRAM chip directly on top of an application processor.

The electrical connections between both chips are traditionally established by wire-bonding from the edge of the upper chip to the edge of the lower chip. However, this requires additional space on the chips or could be undesirable due to circuit design reasons.

An alternative approach is the utilization of through-silicon vias (TSV) to connect multiple chips to each other. The electrical connection is established through Cu-filled holes that run directly through the silicon substrate of one of the chips.

Since the integration of TSVs into semiconductor manufacturing flows has required additional process steps that may influence the performance or reliability of the devices, physical characterization is needed during development, qualification and manufacturing.

State of the art

The current issues related to FA of electronic components and systems could be structured around three main areas:

- Techniques of FA;
- Failure mechanisms (FMs);
- Models for the physics-of-failure (PoF).

Classification of failures

The classification of failures may be divided into three regions.

A. Infant Failures: are directly attributable to excessive mechanical stresses introduced in the mounting operation and the assembly of the equipment. The decreasing failure-rate region contains mainly leakage and die-cracks failures.

B. Random Failures: are very difficult to analyse with great precision and the experience showed that they belong either to late infant failures or early wear out failures.

C. Wear out Failures are mainly caused by thermal-cycle or fatigue. Applications without such conditions show a decreasing failure-rate. Long-term life tests at high temperature have substantiated this fact.

Considering the relationship between Reliability and Failure Rate, it becomes obvious that for most of the applications we need to give
more information about all operational conditions the device will undergo in the field.

Experience has shown that lack of this information was one of the major reasons for repetitive field-failures.

Therefore, good field performance depends on at least two factors:
1. Is the device specified and used properly?
2. Is the device inherently reliable when used properly?

Open and short failures

Open and short failures are two common failure modes frequently seen in assembly back-end-of-line (BEOL) processing as well as in integrated circuit (IC) testing [6].

With continuously increasing memory density and shrinking semiconductor device sizes, failure analysis (FA) is becoming more and more challenging [7].

Fault isolation is a critical step in the FA flow because it helps narrow the focus to the suspected area. There are many fault detection tools available for a short failure, such as scanning super-conducting quantum interference device (SQUID) microscopy [8], infrared lock-in thermal camera [9], photon emission microscopy (PEM) [10], and thermal-induced voltage alteration (TIVA) [11].

However, with regard to open failure detection, for more than 10 years, time-domain reflectometry (TDR) has been the only technique available to the semiconductor industry. Its distance-to-defect accuracy is about 500 µm [12]; hence, to isolate faults in advanced flip-chip devices accurately using TDR is becoming more difficult.

The recently introduced electro-optical terahertz pulse reflectometry (EOTPR) system demonstrated its capabilities for open fault isolation with higher accuracy compared to TDR [13]; however, it also requires waveform comparison among the reject, golden and bare substrate units to determine the defect location.

In contrast, SQUID space-domain reflectometry (SDR) directly displays a physical 2D image of an open failure, which provides real visualization of the defect location. With application of the appropriate radio frequency (RF) scanning frequency and linear decay analysis, SDR is able to isolate open fault locations with better accuracy than TDR.

Backside techniques

Backside techniques are typically preferred when electrically localizing submicron defects or subtle ESD damage in multiple large power FETs.

However, depending on the ICs design, layout and packaging constraints, backside analysis is not always feasible.

Backside functional bias conditions are often inadequate to properly exercise the leakage source fault isolation. Furthermore, the electrical analysis of power FET gate nodes typically requires topside micro-probing, making a topside approach necessary.

Unfortunately, fault isolation techniques such as OBIRCH (Optical Beam Induced Resistance Change) can be challenging from the topside due to thick blanket metallization and multiple backend process layers that distort or obscure OBIRCH signals.

The prime purpose of our study is to determine the failure rate of the improved power GaAs FET’s at their maximum possible temperature in normal operation. This temperature can be specified in terms of the channel temperature to be 110 °C.

Under the assumption that all failure mechanisms obey the Arrhenius law, thermally accelerated life tests can be effective for this purpose. It is desirable to adopt as many test temperatures as possible to investigate the temperature dependence of the device life. However, there is a practical limit on choosing the number of test temperatures. This limit is related to the test facilities and the available number of sample devices.

Three channel temperatures of 250 °C, 210 °C, and 175 °C were chosen corresponding to equal separation on an inverse-temperature scale. As the temperature decreases, an increased number of sample devices were allocated to have statistically meaningful data as even as possible over the three temperatures.

As has been experimentally found, the thermal resistance of a GaAs FET under study increases with increasing channel temperature until approximately 270 °C. Beyond this temperature, however, the temperature dependence of thermal resistance becomes difficult to predict precisely.

Therefore, the highest temperature was selected to be 250 °C, providing some safety margin.

GaN FETs

GaN FETs are increasingly used as next-generation, high-power devices for power electronics systems. They can achieve low power loss operation due to high carrier mobility in the two-dimensional electron channel (2DEG) as well as high breakdown voltage. GaN FETs are majority-carrier devices; therefore, the absence of reverse charge recovery allows high-voltage operation.

GaN gallium nitride transistors can switch much faster than silicon MOSFETs, thus having the potential to achieve lower switching losses. At high speeds, however, certain types of packaging can limit the switching performance of GaN FETs. Integrating the GaN FET and driver in the same package reduces parasitic inductances and optimises switching performance.

The integration of the driver also enables the implementation of protection features.
Based on silicon and carbon, SiC has a bandgap of 3.3...3.4 eV. Silicon has a band gap of 1.1 eV. SiC devices are made on 100 mm or 150 mm substrates, which makes wafers somewhat expensive. Also, SiC FETs have a low channel mobility efficiency. Because of the extremely low gate charge and output capacitance, GaN power FETs can be switched at extremely high speeds with significantly reduced switching losses and improved efficiency - compared to silicon FETs.

**Reliability of GaN power transistors**

There has been tremendous progress in the industry’s understanding of GaN reliability in the last few years spurred by investments to deploy GaN power devices in IT infrastructure such as server, datacentre power supplies and consumer electronics such as AC/DC adaptors. The gallium nitride (GaN) power industry has invested considerable time and effort on reliability validation, formed a JEDEC Committee and published several key guidelines including one on switching reliability.

The overarching question today is not “Is GaN reliable?” anymore; but instead, it is “How did you validate GaN’s reliability and what is the data?”

GaN on SiC High Electron Mobility-Transistor (HEMT) is the most promising III-V semiconductor technology to deliver large RF power densities (5 W/mm and beyond) at high frequency (~10 GHz and beyond) [1]. This performance can be achieved [14]-[27] because of the large bandgap of GaN that allows for high critical fields and thus high voltage operation and high RF power densities at small drain source spacing, because of the existence of an even higher heterobarrier (AlGaN) with a reasonable lattice mismatch, capable of providing the conduction band discontinuity required for effective carrier confinement, and because of the existence of a semi insulating substrate (SiC) with excellent thermal and insulating properties on which GaN of reasonable quality can be grown.

**The meaning of reliability**

There are several meanings of the term reliability in common use.

At first, people are considered reliable when they can receive credit from strict banks.

Secondly, people are said to be reliable when their word can be trusted as truthful.

Third, the behaviour of reliable or responsible people matches what they say.

Finally, since no one would have much confidence in a very old or weak person’s ability to climb Mt. Everest—not because of poor planning, but because of the person’s physical ability – a person is reliable (or will be successful in climbing), because his health is excellent. This last is the sense in which we use the term reliability here. Reliability resides not in the software, but in the hardware itself; that is, it is trouble-free in its physical substance.

Applying this to hardware products, reliability means that there are no problems with performance resulting from constituent material fractures due to sudden environmental changes, such as shock, or from material degradation over time.

Note that software bugs are excluded, because they are not accompanied by physical change; they pertain to performance-related issues, not to reliability.

As NASA’s manual of reliability training says, “Software reliability is really a misnomer”, it actually means “a measure of software design adequacy”.

Good reliability does not mean good performance or better ease of use. For example, good reliability in an airplane means the hardware remains trouble-free during take-off, navigation, and landing, not that the cruise speed is high or the turning radius is small.

The reliability of a product can be described in terms of two indices: lifetime and annual failure rate within lifetime, based on the occurrence of physical product disorders, or failures.

To illustrate, consider human reliability indices. In this case, there are also two indices: death rate and life expectancy. The infant mortality rate, crude death rate, life expectancy at birth, and life expectancy at age 60 are common statistics reported by many countries.

However, in the case of hardware, the task of designing and manufacturing products is more critical because the product does not have a spontaneous healing system like living things. Awkward design or improper manufacture of products can induce many initial failures at an early stage and high random failures during use, resulting in a short lifetime due to wear-out failures.

The reliability target of a product should be set at zero, or nearly zero, failures for the lifetime generally expected by customers.

Therefore, product reliability regards hardware as successful or reliable when the product operates without any failures during its anticipated lifetime.

Since reliability is time related, the term future quality can be used to describe how quality changes over time.

Good reliability is good future quality. But just as every person in old age shows sign of decrepitude, every product also exhibits weaknesses over time due to degradation (wear out).

Future quality should ideally be the same as initial quality, although, of course, it cannot be better than initial quality. For instance, my friend’s car has been never repaired for 5 years
and my car, purchased at the same time, has been repaired a few times, so my friend’s car would have better future quality than mine.

But that does not mean that my friend’s car after 5 years of use would be better than it was when my friend purchased it. It is possible, however, to say that the future quality of a product can be better than that of competitors’ products.

Reliability indices always include the concept of time. The index for lifetime includes the time dimension, but the index of the failure rate does not. The failure rate is easier to understand if duration is included, using such terms as the annual failure rate or the hourly failure rate.

And the appropriate concept of time for the lifetime index also requires consideration. For example, is the mean time to failure the same as the actual lifetime? Finally, does good reliability mean a good product?

A good product has high value compared to its price. This concept appears in the “survival equation” derived by the theorist S. Yoon: “The product value should be greater than the product price, and the product cost should be less than the product price.”

**Dimensional differences between quality defects and failures**

Generally, quality-related trouble appears for one of three reasons:

1. Customer misuse of the product
2. Non-conformance to specifications during manufacture
3. Mistakes or incompleteness in design specifications or manufacturing.

An analogy of the first case would be a car crash due to the driver’s dozing off.

In the second instance, being out of specification or having a quality defect would obviously induce poor performance or failure in a product.

Finally, any mistakes or omissions in design or manufacture will induce faults in an otherwise high-quality system.

The word failure clearly indicates a physical performance disorder related to the product. The number of failures per year for a given production lot results in the annual failure rate; the time at which the same failures occur epidemically is the lifetime. The absence of failures defines good reliability.

The terms quality defect and poor quality can refer to a variety of problems that appear before customer use, such as flaws in the external case or imperfections that indicate that some malfunction will occur during use, causing poor performance or the failure to work at all.

Similarly, the term good quality has diverse meanings. In the marketplace, the phrase connotes good aesthetic design as well as good performance.

Let’s consider that the term quality defect, in the narrower sense commonly used in manufacturing, means that some aspect of the product is out of tolerance. If something can be inspected off the line and found to be out of specification, it is considered to have a quality defect.

In the early 1990s, a company in Korea mobilized its employees to participate in a quality improvement campaign to reduce the defect rates of its components to 100 parts per million, or 0.01%; the program has now progressed to a limit of a single ppm in order to upgrade quality even further. The after-sales service rate was also lowered a little, contrary to the company’s expectations.

Many readers have noticed that the differences between quality defects and after-sales service rates never change, and that advanced companies have surprisingly low after-sales service rates in spite of high defect rates.

Why do the products of advanced companies need repair so seldom and have virtually no failures?

The reason is that the concepts of defect rate and failure are completely different.

Activities to decrease quality defects for the purpose of improving reliability are as useless as trying to compare apples and oranges.

Since failure occurs even when all specifications are satisfied and no defects are evident, obviously something can be improved, by either revising specifications or making them more thorough.

Because failures are principally removed by design changes or specification elaboration, reliability pertains to design technology and should be called reliability design technology. Why should designs or specifications be changed?

**The two elements of failure: Stress and material**

As mentioned earlier, product failure is a physical problem.

Let’s focus on where the trouble occurs, or the failure site. Because every product is an aggregate of many structures, the unit structure of the failure site may only be visible when the product is disassembled. The loads applied to this unit structure produce stresses.

Failure occurs when the stress is greater than the material strength, or when the material cannot endure the applied stress. Understanding this process is called failure mechanics, and its two elements are stress and materials.

A desirable or reliable structure includes both well-dispersed stresses and reliable materials. This description sounds like it applies only to mechanical structures, but it is equally true of electronic devices, which have analogous structures.
For example, semiconductors experience thin-gate rupture due to high voltages (electrical overstress [EOS]), and circuit wires have narrow-width openings due to high currents (electro migration).

In either case, failure occurs when the materials comprising the structures do not survive the environmental and operational stresses applied to them; after all, every object from a microdevice to a high building has some architecture, and it is only as strong as its component parts. Because reliability is the relationship between stresses and materials, the solution for avoiding failure is altering the structures to better disperse stresses or replacing materials. But strong materials are not always reliable. Bronze is weaker than steel in mechanical strength, but better than steel in corrosive environments. It is best to select a material with consistently high mechanical strength throughout the product’s usage lifetime. The defect in the middle of the material will contribute to failure. Such defects are usually produced in the manufacturing process, which leads us to regard defects solely as a manufacturing issue.

The specifications for flaws should be revised and expanded to include such data as location and size. Altering the structure or materials or making specifications more detailed assumes that achieving reliability pertains to design, not manufacturing. Since product design includes reliability, it is properly done jointly by the product designer and a reliability specialist.

Consequently, design specifications should be classified into two categories: performance specifications and reliability-related specifications. When an item is inspected for a certain specification and turns out to be out of tolerance, the item will not work well if the specification relates to performance; if it is a reliability specification, the item cannot reach its design target life. However, there is no need to describe specifications in detail in both the performance and reliability categories; it is often awkward or difficult to discuss each category separately in written materials about the manufacture of the item, and in many cases, it is impossible to differentiate between the two. That means that generally, the specifications of materials and structures are performance specifications as well as reliability specifications.

For example, the structures in the fuselage of an airplane should be assembled without welding because the fuselage receives repetitive stresses from the expansion and contraction of the surrounding air at high altitude and ground, which can lead to breakage along welding lines. In this case, it would be hard to distinguish whether reliability engineering was included in the design or not. Moreover, it is easy to overlook minor details that might be regarded as appropriate reliability specifications.

For instance, the corners of plastic components are mostly produced with right angles but giving them small round radii would make them stronger, helping to prevent failure in the short run.

Likewise, the lead wires of resistors, which are used plentifully in printed circuit boards, are made of copper or brass and clad with tin for good soldering. There is a nickel layer between the two, which hampers both corrosion or the development of a galvanic cell for copper and the diffusion of zinc oxidation for brass— both of which could lead to failure due to an open circuit in a few years.

Designers frequently miss minor details that are critical to reliability, like the corner radius of components and the nickel layer for soldering materials, because they seem irrelevant to performance.

Generally, failure does not occur when a product first comes out, unless the stresses are sudden and catastrophic, but only after months or years, due to gradual intrinsic changes in its materials. Thus, engineers cannot always tell ahead of production whether or not the reliability specifications are adequate. This is why product designers are often inattentive to reliability specifications. At the same time, corrective actions that require analysis of failures occurring in the market are usually considered a matter of inadequate design or manufacturing specifications. Therefore, thorough product design should include conforming product materials and their associated stress dispersal structures, and carefully assessing the operational environmental and operational stresses. This also means that the lifetime and failure rate must be predetermined, although engineers cannot estimate them before testing completed products.

If a product is designed carelessly, the product lifetime will be short and its failure rate will be high, even if there is a well-organized quality control system. Thus, reliability is one of the intrinsic characteristics of a designed product. Top companies follow this practice, which is why customers prefer name brands for durables. It is impossible to produce reliability indices of materials without considering the stresses applied to them. There is no such thing as material reliability. If the materials are incorporated into the structure and the applicable environmental /operational conditions are applied, the effects on the materials are evident and the resulting failure phenomena are observable.

Generally, the reliability of a material can be determined, if stresses are applied to a test piece in accordance with certain conditions, but the result actually reflects the reliability of the total structure—that is, the test piece material and the stress-producing engine. Therefore, the result
cannot be applied accurately to a structure that differs from that of the test piece, with its relevant stresses. Rather, the result expresses the relative reliability characteristics of materials under the anticipated stresses.

As the basic concept of failure mechanics and its two elements indicates, the shape of the failure rate curve over time does not vary for different kinds of components.

Electronic components consist of a combination of materials experiencing stresses, just like mechanical components. The unit structure of electronic components, including the materials and their attendant stresses, is not essentially different from that of mechanical parts, but the shape of the failure rate curve is usually assumed to be different.

It is frequently said that the failure of electronic items occurs accidentally, or that failure is random, but that, for mechanical items, there are few failures within the expected lifetime and many wear-out failures near the end. Is it true that only wear-out failures occur in mechanical items and random failures in electronic products? This hypothesis was formulated in the 1950s in an attempt to explain the differences between the failure phenomena of mechanical and electronic components when the latter were just coming on the market, but it is a careless conclusion to draw today; it would be impossible to make a generalization such as this that would apply to all items.

Again, let’s look at an example. In the capacitors frequently used in electric power circuits, rubber caps around the terminals block the leakage of electrolytes.

As the rubber degrades over time and loses elasticity, the electrolyte flows out and creates wear-out failure. Rubber degradation can be accelerated by the stress of ambient high temperature (the solution is to locate electrolyte capacitors as far as possible from heat-producing devices). There is nothing random about this phenomenon. In the early industrial age, many random failures occurred in mechanical items.

Nowadays mechanical items, like car engines, have little or no trouble during use and seldom fail over their lifetimes, due to over 200 years of development in mechanical engineering. If car engines are well maintained, they seldom fail. Likewise, in the early stages of the development of electronic devices, there were many initial and random failures.

Now electronic devices have been improved to the point where initial failures have mostly disappeared and random failures have decreased enormously, as the limitations applicable to circuits were observed and corrected.

Pecht’s law has established that the failure rates of microcircuit devices of digital systems have decreased by about 50% every 15 months. It was possible to hypothesize and experimentally verify this theory because the understanding of both microstructures and the materials of electronic constituents has broadened substantially after several decades of research and failure analysis.

Reliability engineering as a science

So far, some confusing concepts have been clarified. These concepts include that the mean time to failure (MTTF) is inadequate as a reliability index, that failure occurs because of the relationship among materials and stress, and that corrective action for potential failures must be taken at the design stage.

Given these conclusions, reliability technology as it is now practiced is virtually groundless.

The Japanese call this adopted engineering technology, † meaning that reliability engineering has not been developed independently and borrows and adopts theories from other fields of science and technology.

In order to establish an independent technological field, the many related topics must be addressed and integrated from the ground up.

Let us briefly review the historical development of reliability engineering and the reasons why discrete reliability theories have never been established. Research into reliability began as a result of the exceptionally high number of failures in vacuum tubes used in World War II military weapons. Over half (60%) to 75% of vacuum tubes in communications modules had failed.

In the Korean War 5 years later, a similar phenomenon occurred. The Pentagon initiated an investigation into this enormously wasteful situation. The resulting project was undertaken by the Advisory Group on Reliability of Electronic Equipment (AGREE) and lasted for 5 years, from 1952 to 1957, when its findings were published.

Subsequently, however, many of the research projects on reliability issues, such as a survey on the substitution periods of aircraft components, have been classified as confidential military documents and have not been made public.

In related fields, scientists and engineers continued to make substantial contributions on reliability issues.

In the early 1950s, the basic mathematics of failure distributions was established statistically. In the 1960s, the methods called failure modes and effects analysis (FMEA) and fault tree analysis (FTA) were devised — methods that are now applied to many hardware systems.

In the 1970s, there were great advances in failure analysis, mainly in materials engineering — a success that was much indebted to rapid progress in destructive testing/non-destructive testing (DT/NDT) equipment, such as the scanning electron microscope (SEM) and the scanning acoustic microscope (SAM).
Despite these achievements, since the 1970s it has been acknowledged that there is a major gap between reliability theory and its application to industrial fields—the real problem to be solved. Since the results estimated by current reliability prediction methods for electronic components differ widely from experimental data, specialists doubt the methods and their bases.

Recently, scientists have been working to clarify in detail the processes leading to failure (failure mechanisms), using the physics of failure (PoF) approach.

Why reliability concepts and methodologies have not been still established in a modern society, which has been producing so many technical advances daily?

The first reason is that reliability information has not been made readily available by industries to reliability specialists. Corporations fear that releasing survey results about reliability problems will negatively influence business, so they keep them confidential.

The second reason is that the solutions to reliability failures, the methodologies to determine solutions, and the foundational reliability concepts should be established through a multidisciplinary approach.

The current system relegates reliability issues to the lowest level of the engineering staff or to specialists in only one field; this hampers contact among experts and the necessary interconnections among researchers in various fields. Thus, the concepts of reliability technology are not systemized and each specialist understands them idiosyncratically.

What fields comprise reliability technology? As we noted, failure occurs from the relationship between stress and material, which must be understood first and foremost. We need to know how the stresses from both the environmental conditions in which the item is operated (indoors, in a field, under the ocean, in outer space, and so on) and operational conditions (mechanical, electrical, and electronic) are concentrated or dispersed in the item’s structure.

We also must know how the materials in the structure reach the point of breakage, degradation, or wear due to these stresses. In order to quantify the degradation process, a model of failure mechanics must be introduced and analysed, together with a time variable, which pertains to the field of physical chemistry.

Furthermore, the manufacturing process has to be understood from the original materials to the final product, because every step influences the inner structure of the materials.

Finally, knowledge of statistics is crucial for design of experiments and estimating their results. It is quite fortunate that observed results from the various interactions between stresses and materials have been investigated and publicized in the name of relevant failure mechanisms.

After establishing the basics in the next chapter, such as that reliability comes under the larger heading of product quality, the relevant terminology will be defined, with an explication of the hierarchy of those terms. Then the systematic sequence of reliability engineering will be explored.

The concepts of reliability technology are discussed and a degradation model, for which the time variable function is easy to memorize, will be connected with failure mechanisms, and the statistics for estimating lifetime will be integrated.

**Conclusion**

If a new product is designed with new materials or a new structure, approval specifications for it should also be addressed anew. This is a universally applicable approach. Just because a set of specifications is well established does not mean they remain valid regardless of changed circumstances.

Any change requires revising the specifications, sometimes quantitatively. This also pertains to design functions; otherwise, reliability accidents related to poor quality will occur at some time even in advanced corporations.

Reliability technology involves estimating the future behaviour of hardware products after 10 or 20 years. It is difficult but not impossible to do this. Accurately grasping reliability concepts can motivate an organization to adopt scientific principles, minimizing quality problems down the line.

**References**


Mr. Băjenescu is the author of many technical books - published in English, French, German and Romanian. His latest book entitled (in German) "Zuverlässige Bauelemente für elektronische Systeme" was published in 2020 by the prestigious publisher SPRINGER.

Since 1991, he won many Awards and Distinctions, presented by the Romanian Academy, Romanian Society for Quality, Romanian Engineers Association, etc. for his contribution to reliability science and technology.

Recently, he received the honorific titles of Doctor Honoris Causa from the Romanian Military Academy and from Technical University of the Republic of Moldavia.

In 2013, he obtained, together with prof. Marius Băzu (head of reliability laboratory of Romanian Research Institute for Micro and Nanotechnologies IMT) the Romanian Academy "Tudor Tănăsescu" prize for the book Failure Analysis, published by John Wiley & Sons.

e-mail-address: tmbajenesco@gmail.com