The Physics of Degradation in Engineered Materials and Devices

Fundamentals and Principles

Edited by Jonathan Swingler
with major contribution
by Alec Feinberg

THE PHYSICS OF DEGRADATION IN ENGINEERED MATERIALS AND DEVICES

THE PHYSICS OF DEGRADATION IN ENGINEERED MATERIALS AND DEVICES

FUNDAMENTALS AND PRINCIPLES

Edited by JONATHAN SWINGLER
with major contribution by ALEC FEINBERG

The Physics of Degradation in Engineered Materials and Devices:
Fundamentals and Principles
Copyright © Momentum Press®, LLC, 2015.

ABSTRACT

Degradation is apparent in all things and is fundamental to manufactured as well as natural objects. It is often described by the second law of thermodynamics where entropy, a measure of disorder, tends to increase with time in a closed system. Things age!

This concise reference work brings together experts and key players engaged in the physics of degradation to present the background science, current thinking and developments in understanding, and give a detailed account of emerging issues across a selection of engineering applications. The work has been put together to equip the upper level undergraduate student, postgraduate student as well as the professional engineer and scientist in the importance of physics of degradation. The aim of the work is to bridge the gap between published textbooks on the fundamental science of degradation phenomena and published research on the engineering science of actual fabricated materials and devices.

A history of the observation and understanding of physics of degradation is presented. The fundamentals and principles of thermodynamics and entropy are extensively discussed. This is the focus of this work with an extended chapter by Alec Feinberg on equilibrium thermodynamic damage and non-equilibrium thermodynamic damage. The work concludes with two particular technologies to give examples of areas of application.

KEY WORDS

ageing, degradation, electrical power systems, entropy, ferroelectric devices, thermodynamics

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Jonathan Swingler

1.1   INTRODUCTION

This concise reference work has been put together to equip upper-level undergraduate and postgraduate students as well as professional engineers and scientists with the importance of degradation in engineered materials and devices. It brings together experts and key players engaged in the physics of degradation to present the background science, current thinking, and developments in understanding and give a detailed account of the emerging issues across a selection of engineering applications.

   This introductory chapter presents the rationale and emphasis of the book. It explains the relevance of what is termed in the title and the subtitle of the work. The chapters on thermodynamics and entropy are very briefly surveyed, along with two other chapters on example technologies and applications to show the significance of the subject area.

1.2   RATIONALE AND EMPHASIS

The aim of the work is to bridge the gap between published textbooks on the fundamental science of degradation phenomena and published research on the engineering science of actual fabricated materials and devices.

1.2.1   PHYSICS OF DEGRADATION IN ENGINEERED MATERIALS AND DEVICES

Degradation is apparent in naturally occurring materials and structures as well as human-engineered materials and devices. In everyday experience, it is the ever-present phenomena of spontaneous loss of some quality, functionality and order. Work has to be done from outside the system of interest to maintain that functionality, or that order. The second law of thermodynamics formally captures this idea with the concept of entropy or disorder, which states:

> *During real processes,*
> *the entropy of an isolated system always increases.*
> *In the state of equilibrium the entropy attains its maximum value.*
> (Mandl 1989)

This loss of order or degradation has many terms or phrases to label the phenomena, such as ageing, deterioration, devolution, and wear-out. It is this degradation in engineered systems that this book explores.

The mechanisms of degradation for a variety of materials and structures cover a wide range of discipline categories such as thermal, mechanical, chemical, biological, and so on. These are shown in Figure 1.1, which is not an exhaustive listing, and there may be more than one process behind each mechanism. All mechanisms require the knowledge and understanding of natural processes within these disciplines and thus are grouped together as the *physics* of degradation. The fundamental underlying principle is entropy and the second law of thermodynamics.
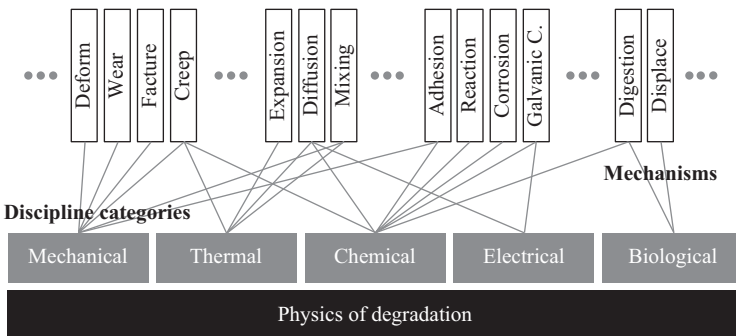
Figure 1.1. Mechanisms in the physics of degradation.

The title contains the phrase "engineered materials and devices," particularly to emphasize "engineered" as the designer or architect is interesting in developing or fabricating a material or device that has a specific purpose or function. The scope of the materials and devices themselves span a vast array of products from mechanical to electrical, to microelectronic, and so on.

1.2.2 FUNDAMENTALS AND PRINCIPLES

Chapter 2 lays out the history of the observation and understanding of physics of degradation. The fundamentals and principles, the key features, are identified, which have led to the current understanding. However, as Jones highlights in Chapter 2, these details are limited mainly to the past 100 years or so, since the literature in the discipline has become available.

The fundamentals and principles of thermodynamics and entropy are extensively discussed in Chapters 3 and 4. This is the focal point of the book.

1.2.3 THE READER AND THE AUTHORS

The target readers are those who are interested in taking thermodynamics of ageing beyond the theoretical ideas to specific applications. For the upper-level undergraduate student embarking on project work, it opens up opportunities to develop predictions from theory and conduct experiments to explore the limits of those predictions. For the postgraduate and postdoctoral researcher, it offers springboard ideas to push the boundaries of current thinking. For the professional engineer and scientist, it gives grounding in the current thinking of the underlying science to enable the development of innovative solutions.

The contributing authors are scientists and engineering working in the field from both academia and industry. There is a wealth of experience from initiating novel concepts for pushing forward the research effort, developing innovative solutions for commercial products, and communicating these ideas.

1.3   THERMODYNAMICS AND ENTROPY

Two chapters address the role of thermodynamics in the understanding of degradation or ageing in engineered materials and devices. The first

of these chapters focuses on an approach, involving irreversible processes, described as the degradation entropy generation (DEG) theorem. The second of these chapters focuses on an approach, again emphasizing irreversible processes, highlighting the interaction of the system with its environment and entropy damage. This second chapter is an extended chapter exploring equilibrium thermodynamic damage and non-equilibrium thermodynamic damage.

1.3.1 DEG THEOREM

The DEG theorem is presented by Bryant in Chapter 3. This links the rate of degradation or ageing to rates of production of irreversible entropy by underlying dissipative physical processes. The second law of thermodynamics plays a central role here. To explain this approach, a review of thermodynamics is given in the chapter discussing entropy and the production of irreversible entropy under particular conditions. Four conditions are identified:

1. The Stationary Condition
   The system is in its simplest state of constant internal energy and entropy.
2. The Mechanical Strain and Isothermal Condition
   The system undergoes mechanical change at constant temperature, which is calculated with Helmholtz free energy.
3. The Isentropic and Isobaric Condition
   The system undergoes volume change at constant pressure, which is calculated using enthalpies.
4. The Isothermal and Isobaric Condition
   The system undergoes volume change due to chemical activity at constant temperature and pressure, which is calculated with Gibbs free energy.

   Ageing and degradation mechanisms are then discussed in relation to irreversible entropy and dissipative processes. Bryant identifies nine dissipative processes, giving the mathematical relationship between the irreversible entropy and the process: (1) adhesion, (2) plastic deformation, viscous dissipation, and cutting, (3) fracture, (4) phase changes, (5) chemical reactions, (6) diffusion, (7) mixing, (8) heat transfer, and (9) thermal expansion.

The chapter gives three detailed examples of using the DEG theorem to describe degradation due to:

1. wear in a system under mechanical sliding,
2. fatigue in a system under mechanical load, and
3. multiple processes in an electrical battery system.

1.3.2 EQUILIBRIUM THERMODYNAMIC DAMAGE

Chapter 4 is an extended chapter by Feinberg covering two aspects concerning thermodynamic damage assessment of materials and devices in a state of equilibrium and non-equilibrium. Equilibrium thermodynamics provides techniques to describe the initial and final states of a system, whereas the non-equilibrium thermodynamics provides techniques to describe in detail the paths from the initial to the final state of a system.

The first half of the chapter focuses on equilibrium thermodynamics and is concerned with defining and discussing the main concepts.

The second law of thermodynamics plays a central role in the chapter and is rephrased in several places to emphasize particular ideas presented. The main idea is that it is the interaction between the system and its environment that drives the degradation of the system as the system comes into thermodynamic equilibrium with its environment. Other ideas identified are that, left to itself, the system–environment interaction will increase total entropy and the system will undergo a decrease in free energy towards a minimum.

Thermodynamic work and the first law of thermodynamics are introduced. The change in energy of a system is due to the work performed between the system and the environment, and the heat exchanged between the environment and the system. Generalized conjugate mechanical work variables of a generalized force and a generalized displacement are discussed.

The second law of thermodynamics, device damage, catastrophic failure, and parametric failure are all elaborated upon. Simple examples are given to illustrate the relevant points. Feinberg mathematically defines four thermodynamic ageing states in terms of free energy and entropy behavior:

1. Non-equilibrium ageing state
2. Equilibrium non-ageing state

This is associated with the catastrophic failure of a device (system) when that system has a free energy as small as possible and the entropy is as large as possible with the system in a final thermodynamic equilibrium state with the environment.

3. Parametric equilibrium ageing state

This is associated with parametric failure of a device (system) when its free energy and entropy have attained a critical threshold value.

4. Relative equilibrium ageing state

This is associated with activated failures.

Measuring damage entropy processes and measures that can be used to characterize system-level entropy damage are discussed in the first half. This is packed with useful information and analysis. A quasi-static measurement process is devised for measuring damage moving from an initial state to a final state.

1.3.3   NON-EQUILIBRIUM THERMODYNAMIC DAMAGE

The second section of Chapter 4 focuses on non-equilibrium thermodynamics and is concerned with how degradation or ageing takes place over time. A detailed analysis and description enable mathematical modeling of the ageing process. The quasi-static measurement process, first proposed in Section 1, is further used, but this time as many quasi-static thermodynamic measurement states are added as required to describe the process.

Cumulative damage to a system during a cyclic process is analyzed using the generalized conjugate work variables. A damage ratio for tracking degradation over time and consequently an acceleration factor is found. The "damage acceleration factor," as Feinberg calls it, is found to be closely related to that acceleration factor used in reliability testing. The damage acceleration factor is different, though, with an additional stress term arising from energy-related thermodynamic work. Examples are given in Chapter 4 to illustrate these points relating to electrochemistry in battery devices and microelectronics.

The approaches developed in Chapter 4 on thermodynamic damage are worked through in a number of mechanical systems. Examples on creep, wear, thermal cycling fatigue, and vibration fatigue are presented in detail.

The chapter concludes with an extensive discussion on the change in free energy during degradation: the thermal activation free energy

approach. The system degradation path to a minimum may not be the final minimum available to the system. An energy barrier or barriers may exist to further or the final minimum.

1.4   TECHNOLOGIES AND APPLICATIONS

The final two chapters in the book focus on particular technologies to give examples of areas of application. The first of these is in electrical power systems and methods of monitoring the types of degradation. The second of these is in ferroelectric devices and the observed degradation mechanisms.

1.4.1   MONITORING DEGRADATION IN ELECTRICAL POWER SYSTEMS

Chapter 5 by Ma reviews measurement and monitoring technologies for electrical power systems concentrating on characterization of metallic material, condition monitoring of insulation and condition monitoring of associated assets.

The chapter focuses on insulation degradation and its measurement and discusses the thermal, mechanical, and electrical stresses that lead to degradation. Defect formation is highlighted, which can lead to partial discharges in operation. It is the interaction of thermal, mechanical, and electrical stresses on defects that increases the probability of partial discharge activity and eventual failure of the system. Thus, Ma makes the case that it is imperative that partial discharge in such systems is detected at an early stage for a quick remedy.

Ma extensively discusses the different approaches to partial discharge detection and measurement. A particularly helpful section is on noise issues and denoising of measurements. Partial discharge measurements can often be immersed and lost by electrical noise and require noise reduction techniques that fall into two categories:

1.   Specialist hardware—purpose-built circuits, and
2.   Specialist software—purpose-built signal processing techniques.

A selection of these circuits and techniques is discussed in detail. An example using a wavelet-based technique is presented and proves to be impressive.

1.4.2 PHYSICS OF DEGRADATION IN FERROELECTRIC DEVICES

Ferroelectric properties of particular materials are discussed in Chapter 6 by Weaver. These materials can be fabricated into many passive and functional devices, which are listed by Weaver, and include, for example, capacitors and actuators. Piezoelectric materials are classified as a subset of pyroelectric materials, which are a sub-set of ferroelectric materials. Humidity and temperature effects are the main factors comprehensively discussed in this chapter that lead to degradation.

The influence of humidity on the performance and thus degradation of piezoelectric materials and devices is also comprehensively discussed. Under such conditions, material properties, geometric factors, and electric field stresses all have an impact on degradation. Conduction mechanisms, and their development within the dielectric (electrically insulting piezoelectric) material are examined. The importance of ceramic properties, electrode materials, and barrier coatings are presented.

Temperature effects, and particularly thermal expansion, of the ferroelectric material as well as the devices structure is presented as the next important factor concerning degradation. Actual devices are discussed with real data to illustrate the issues involved.

This chapter also focuses upon and highlights issues concerning measurement techniques because to develop a scientific understanding of degradation of ferroelectric materials in harsh environments, it requires the ability to measure complex coupled properties. Therefore, a description of measurement challenges, techniques, and recent developments in this area are presented.

1.4.3 ESTABLISHED AND EMERGING TECHNOLOGIES

Only two applications are presented in this volume to illustrate the significance of the physics of degradation to commercial products. The electrical power system application could be argued is well established in reliability issues and ageing, whereas many aspects of ferroelectric technologies are still an emerging area. Concerning reliability in electrical and electronic systems, the reader is directed to texts like *Reliability Characterisation of Electrical and Electronic Systems* edited by Swingler (2015). There are many more established and emerging technologies that are at various stages of maturity in having a coherent understanding of the ageing phenomena and reliability. These are shown in Figure 1.2, which is not an
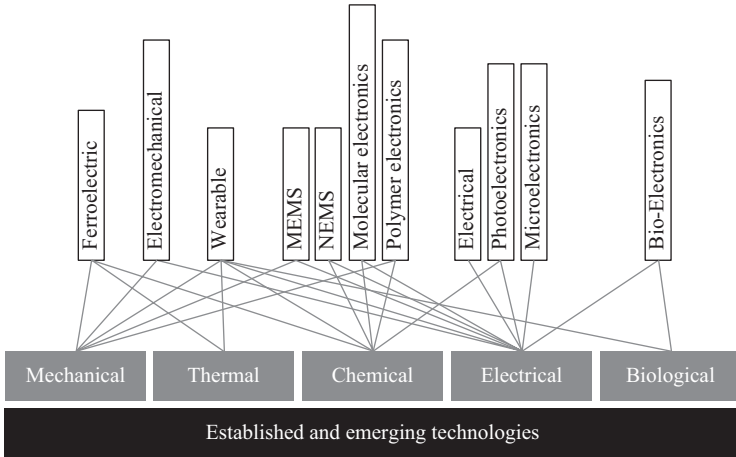
Figure 1.2. Established and emerging technologies.

exhaustive list. Some are, for example, Polymers Electronics, Bio-Electronics, Micro-Electro-Mechanical Systems, Wearable electronics, and Electromechanical Machines.

1.5   CONCLUSION

This introductory chapter has laid out the aim of the book. The book is designed to bridge the gap between published textbooks on the fundamental science of degradation phenomena and published research on the engineering science of actual fabricated materials and devices. The rationale and emphasis of the book are presented, explaining the relevance of the subject matter.

An introduction to the physics of degradation is given with an illustration of the types of mechanisms involved. These mechanisms cover a wide range of disciplines such as thermal, mechanical, chemical, biological, and so on and are underpinned by thermodynamics and entropy. The chapters on thermodynamics and entropy are very briefly surveyed along with the two chapters on example technologies and applications to show the significance of the subject area. Further established and emerging technologies are illustrated, showing the wide scope of the applicable area. These are at different levels of maturity in having a more complete understanding of the degradation phenomena leaving much room for research and innovation.

REFERENCES

Mandl, F. 1989. *Statistical Physics*. 2nd ed. Chichester, UK: John Wiley and Sons.
Swingler, J. 2015. *Reliability Characterisation of Electrical & Electronic Systems*. Cambridge, UK: Woodhead Publishing, an imprint of Elsevier.

CHAPTER 2

History of the Physics of Degradation

Jeff Jones

2.1 CONTEXT

Most failures of engineering systems result from a gradual and irreversible accumulation of damage that occurs during a system's life cycle. This process is known as degradation and is fundamental to all things in nature. This is a consequence of the Second Law of Thermodynamics (Darrow 1942), which states that the entropy, which is a measure of disorder, of closed systems will tend to increase with time. Degradation is self-evident: wood rots, iron oxidizes, metal cracks, material properties alter, and concrete crumbles.

Since degradation is a natural phenomenon, it is really not possible to identify when the effect was first noticed. It is likely that, ancient man observed that as things got older or were utilized more frequently, their performance would deteriorate. This includes everything, from people, to animals, to homes, to tools. For instance, it is very unlikely that the ancients didn't notice that a stone blade would become more blunt with use or age. It is also likely that the ancients identified that this degradation would proceed at a rate proportional to use or age and that they would put "maintenance" in place, by replacing it or sharpening it to improve the performance, before it became too dull to be useful. This process would demonstrate an understanding of degradation, an appreciation of how system performance is affected by a system parameter—in this case, sharpness—and

the ability to forecast how the parameter changes with time. This also implies the use of a model of some kind. It is not possible to think about degradation without considering a model, which means, the physical process of degradation and its modelling cannot be treated separately.

Therefore, in any history of system degradation, it is not possible to identify where the science of degradation began. Similarly, we cannot trace the history up to the present time, since, in this area (which now forms part of asset management) research is moving along at a pace, especially in the prognostics field, where anything that is written would be outdated in a very short time.

This leads to a dilemma about what exactly should be in a document about the history of degradation? The approach taken here has been to try to identify key features that have led to where we are today. To do this limits how far we can look back, since it is necessary to have evidence and knowledge, which can only be achieved through the published work in this area. Since we are limited to approximately the past 100 years. the real title of this chapter should be "Key Moments in the History of Degradation Over the Past 100 or So Years."

2.2   HISTORY

The concept of degradation of materials is intimately tied to the concept of fatigue since that is where much of the early work related to the physics of degradation was carried out. Albert in 1837 published one of the first fatigue-test results (Albert 1837). To generate his data, he constructed a test machine for conveyor chains that had failed in service in the Clausthal mines in Germany. He generated data on the actual component, not just the materials. It seems that in these early days, much of the focus was on transport, for obvious safety reasons, with work by Rankine (1842), who discussed the fatigue strength of railway axles. York (1842) also conducted experiments with railway axles. In 1853, Morin suggested that the axles of horse-drawn mail coaches be replaced after every 60,000 km—an early example of life time prediction (Morin 1853). While these authors are discussing fatigue, the word itself was not mentioned until 1854 by Braithwaite in a paper that describes fatigue failures of such things as brewery equipment, crankshafts, axles, and more. Braithwaite, however, attributes the coinage to someone else (Braithwaite 1854).

The earliest reference to a degradation process in a technical system that is more general than fatigue appears to be "On The Causes Of Failure

Of Deep-Sea Cables," which was published in 1884 but discussed work carried out around 1865 (Graves 1884). In this paper, the author states that the most feasible explanation of the observed failure is

> ....that from some geological cause or other at the bottom of the ocean the iron wires become weakened by oxidation, the hemp as a consequence becomes destroyed, and that the accumulated twists on both sides of a weak place concentrate their force upon it and wring the cable asunder, and that these failures are not necessarily due to abrasion on elevated ridges.

This paper suggests that by this time it is known that fatigue can cause failure, which is already a well-known concept, but that another process, in this case, oxidation, changes the material properties such that the fatigue life is changed. It is the start of this sort of understating that the system mechanical properties are changed by environmental, in this case, chemical, effects that leads to the understanding of the physics of the failure itself.

The first proper systematic study of fatigue life that is recorded is that of Wöhler, who around 1867 published work at the Paris exposition that examined the fatigue behavior of materials (Wöhler 1870a). Of course, earlier, he too had been examining railway axles, from 1858 to 1860, and had been developing deflection gauges (Wöhler 1858, 1860) to measure them under load.

His work at that time was driven by safety issues. There had been a number of accidents in the railway industry (Smith 1990), and he was attempting to work out the probability of failure for different designs and so develop a safety factor that could be used to make designs safer. In fact, his work touched on many areas, the measurement of service loads and the calculation of the corresponding stresses, the design for life and estimation of the probability of survival, as well as some observation on crack propagation (Wöhler 1870b).

Wöhler's work is the first exploration of what were to be known as S–N curves, or alternatively, Wöhler curves, which characterize the fatigue behavior of materials. Wöhler himself drew no curves, presenting his data as tables, and the curves themselves were first drawn by his successor Spangenberg (1874, 1875, 1879). The curves show stress applied on the vertical axis, with a number of application or cycles before failure on the horizontal axis, as in Figure 2.1. The lines represent the behavior of the material and are different for each material. The endurance limit, which is sometimes known as the fatigue limit or
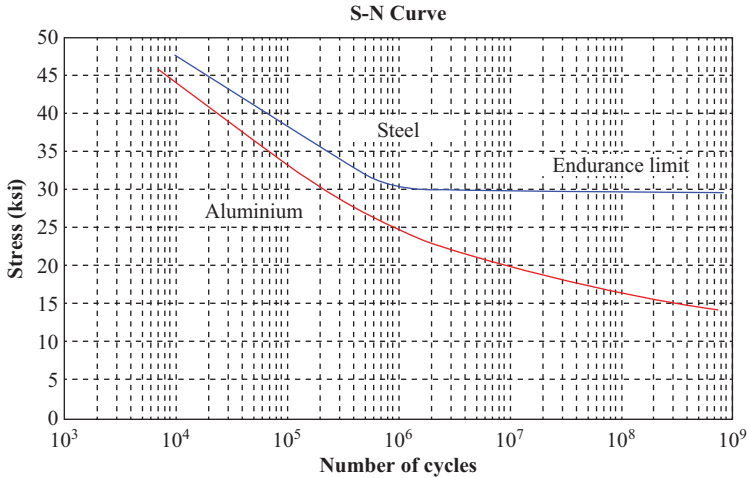
Figure 2.1. Example of an S–N or Wöhler curve (adapted from http://commons. wikimedia.org/wiki/File:S-N_curves).

fatigue strength (displayed by some materials) is the limit at which the material will survive without failure, no matter how many cycles of the load are applied.

Degradation is more than fatigue, and work prior to 1914 on insulators by Austin also noted a degradation effect, here called depreciation. This definition does not refer to overuse of the systems, but rather to environmental conditions (Austin 1914).

> *Investigation shows that insulator trouble increasing with time is not due to fatigue in the material under applied working loads, but rather to depreciation caused by the absorption of water by porous material or by the cracking of the dielectric from high* and *internal mechanical stress set up by uneven temperature in the dielectric, or by greater expansion of cement or metal, or stress from a combination of these….* (Austin 1914)
>
> *The study of depreciation shows that routine tests which will tend to eliminate future depreciation, or refinements in the mechanical features of the insulator, are of far more importance in producing reliability than the designing of insulators to withstand extremely severe dielectric design tests, for insulators which may have extremely high dielectric strength will cause trouble through cracking from internal stress.* (Austin 1914)

In this work, Austin identified the likely cause of the degradation and noted that developing the correct tests for degradation was important. This is probably the first reference to testing to understand the rate of degradation under different environmental conditions.

The data gathered by Wöhler (1870a) was analyzed by Basquin in 1910 using slightly more up-to-date mathematical analysis techniques to produce the initial log–log models for stress life (Basquin 1910). Further, in a similar vein Miner (1945) formalized the linear damage hypothesis that was first suggested by Palmgren (1924). In addition to the Palmgren–Miner rule, this paper contained a four-parameter equation relating the tensile strength to the fatigue limit for the S–N curve, a relationship that was later used by Weibull. Palmgren (1924) also specified B10-life as a design criterion, which is a survival probability statement and the first case was developed for ball bearings.

The Palmgren–Miner rule, more commonly known only as Miner's (1945) rule, is an early example of a damage accumulation model that deals with failure caused by fatigue. This is a progressive, localized, permanent damage that occurs when a material is subject to cyclic stresses below the tensile strength limit of the material. Failure actually begins with the initiation of a microscopic crack in the material, which then develops in a fairly predictable way. This predictability arises since the actual crack propagation rate is a property of the material and is related to the applied stress level and the amount of damage that already exists in the material. Eventually, the crack will become unstable under the applied stress and failure will occur. This will happen after a predictable duration for any particular material and stress level. It is this predictability that leads to degradation models such as the Wöhler or S–N curves (Wöhler 1870a), Miner's law (Miner 1945), Paris's relationship (Paris 1961, 1964), and the Coffin–Manson relation (Coffin and Manson 1954).

Paris (1961) had established that fatigue–crack propagation could be described by a simple equation that can be considered equivalent to an S–N curve under certain limiting conditions and with many assumptions. The model was later refined by Paris (1964) to remove some limitations and generalize it. For a while it was considered to be a large step forward to be able to calculate the propagation of cracks of arbitrary shape under arbitrary stress types even though the equation and some of its assumptions are now considered to have vastly oversimplified the physical reality.

In 1954, and based on the work by Bauschinger (1880, 1881, 1886), Coffin (1954a, 1954b) and Manson (1953) described the behavior of metallic materials under cyclic inelastic strain amplitudes using a four-parameter

relationship that bears their name and creates a new field, known as low-cycle fatigue analysis, which is applied to items loaded over relatively fewer cycles at elevated temperatures.

Fatigue is in fact one of the most persistent problems in engineering design, causing failure in everything, from rotating and reciprocating machinery and components, through failure in complex systems such as aircraft and ships, to large civil engineering structures such as bridges and buildings. Even at the electronic component level, fatigue is a major problem in solder joints, bond wires, and copper-plated vias (Dasgupta 1993).

In 1961, Professor Waloddi Weibull published a book on materials and fatigue testing while working as a consultant for the U.S. Air Force Materials Laboratory (Weibull 1961). Prior to this work, Weibull had originated one of the most useful techniques aiding people in understanding and analyzing degradation, which originated from a paper published in 1951 (Weibull 1951). Up until about 1963, Weibull, who was then a Swedish mechanical engineer, also wrote a number of papers related to fatigue and other failure mechanisms. Weibull's original paper (Weibull 1951) did little more than propose a multiparameter distribution, but it became widely appreciated and was shown by Abernethy, working with Pratt and Whitney in 1967, to have some application to the analysis of defect data (Pratt and Whitney 1967). The versatility of this technique allows it to "mimic" a number of other distributions, including the log-normal distribution, which is one of the main distributions used for the analysis of degradation. This can be theoretically derived under assumptions matching many degradation processes common to material failure mechanisms such as corrosion, diffusion, migration, crack growth, electromigration, and many other chemical reactions and processes (Epstein 1947). While the Weibull distribution has proven to be very powerful and extremely useful in this area, it should be noted that Lawless (1983) commented that an enormous amount of detailed work on the Weibull distribution should have been directed to the investigation of other statistical models.

The log-normal distribution is often considered for modelling the failure time distribution when the corresponding degradation process is based on rates that combine multiplicatively, and is convenient for modelling fatigue crack growth in metals and composites as proposed by Kao (1965). The log-normal distribution is an exponential transformation of the normal distribution and fits many diverse types of data well because it has a great variety of shapes. It is widely used for life data showing early increasing failure rates, and later, decreasing failure rates with increasing age (Lamberson and Kapur 1977; Nelson 1990; Yang 2007). Examples of its application include metal fatigue, electronic components, electrical

insulation, and other stress–strength phenomena. Aitchison and Brown (1957) described various properties and applications of the log-normal distribution. Further, there are several probabilistic models in the literature explaining the origin of the log-normal distribution. For instance, it has been shown in 1969 by Gertsbakh and Kordonsky, that this distribution may appear in a shock accumulation situation (Gertsbakh and Kordonsky 1969). Presently, it is commonly used as a distribution for repair time and analysis of fatigue failures (Hoyland and Rausand 1994) but as shown by Wolstenholme (1999) it can be difficult to analyze and has a particularly awkward hazard function. Sobczyk and Spencer (1992) feature many examples of its use.

Dodson and Howard (1961) published "High Stress Aging to Failure of Semiconductor Devices" that uses the Arrhenius equation as a basis for the assessment of temperature induced ageing of semiconductor devices and interest in this paper led to the annual conference on reliability physics, which started in September 1962. This conference has been fundamental in encouraging and presenting work on failure mechanisms, degradation, and basic reliability physics and testing ever since. For example, Crook (1979) presented a model for time-dependent dielectric breakdown of semiconductors as a function of operational and environmental conditions and the physical parameters of the device and Hieber and Pape (1984) presented a creep–rupture equation that calculates time to rupture as a function of applied mechanical load and temperature.

Davis (1962) developed mathematical models to describe mechanisms by which components fail as a function of time, environmental, and operating stresses. Beau followed this up in 1964 where he described three classical causes of failure (i) reliability limitation inherent in the design, (ii) reliability degradation caused by the factory process, and (iii) reliability degradation caused by the user (Beau 1964). This was addressed through the human factors approach of dealing with mistakes and errors in workmanship to which he attributed most failures. However, the paper addresses the two main degradation effects. Later, Shiomi (1965) introduced a generalized cumulative degradation model for estimation and prediction of component life under successive and different stress levels. This is based on the Arrhenius model and Miner's law, but is extended to more general applicability.

Considerations of stochastic processes are also used to infer lifetime distributions from damage models but they do not always lead to the classical distributional shapes. In 1969, to deal with this problem, Birnbaum and Saunders (1969) described an empirical life distribution model that could be derived from a physical fatigue process where crack growth

causes failure. This was an important step forward for new models that described degradation processes. Meeker and Escobar, much later in 1998, showed that most degradation models will not yield lifetime distributions with closed-form solutions (Meeker and Escobar 1998).

One of the main uses of degradation modeling has been to use it in order to provide an input for reliability prediction. Gertsbakh (1989) suggests special attention to lifetime investigation for failure caused by fatigue. Lu and Meeker (1993) combined regression model, simulation, and bootstrap methods and applied it to degradation data in order to estimate the time-to-failure distribution. Tang and Chang (1995) described a conceptual framework for reliability evaluation from non-destructive accelerated degradation data, where the data were treated as a collection of stochastic processes for which parameters depend on stress level. Yang and Yang (1998) developed a method for rapidly evaluating reliability using degradation data. Crk later in 2000 extended Lu and Meeker's work in developing a multiple multivariate regression model that considered the case when the parameters are random, dependent, and vary with applied stresses (Crk 2000).

Degradation models are getting increasingly important for modern high-cost and high-reliability products, such as aerospace components, where generating failure data through life testing is time consuming and costly. Here, the use of degradation models can be used to shorten testing since time to failure can be predicted when enough data has been collected (Jones 1999). It is also possible, if failure is defined in terms of a specified level of degradation, to use a degradation model to define a particular time-to-failure distribution (Amster and Hooper 1983; Lu and Meeker 1993; Meeker and Escobar 1998; Nelson 1990).

Understanding of degradation processes and degradation modelling has evolved since the early days in the 1830s and the modern engineer or scientist has a vast array of tools and techniques to choose from to help understand the degradation process in the numerous materials used today in technological systems.

2.3 CONCLUSION

The material that has been covered in this chapter is selective and is what the author feels are the most important publications and by implication, pieces of work, in the physics of degradation field over the past 100 years. An attempt has been made to generalize and to look at publications that address more general themes rather than the detailed history of the development of

a single degradation failure mechanism, for example, a mechanical part or an electronic device.

As time progresses and technology moves on, some of the failure mechanisms studied in the earliest days no longer apply because of changes in materials, manufacturing processes, and usage, but these are replaced by new degradation mechanisms introduced by the new materials, new manufacturing processes, and new usage types. The history of the physics of degradation will go on generating new physics, new ideas, new models, and new problems for the foreseeable future.

REFERENCES

Aitchison, J., and J.A.C. Brown. 1957. *The Log Normal Distribution*. New York and London: Cambridge university press.

Albert, W.A.J. 1837. liber Treibseile am Harz. Archiv fur Mineralogie, Georgnosie. Bergbau und Huttenkunde 10, pp. 215–34.

Amster, S.J., and J.H. Hooper. 1983. Accelerated life tests with measured degradation data and growth curve models. *American Statistical Association Annual Meeting*, Vol. 5, no. 9.

Austin, A.O. December 11, 1914. Insulator Depreciation and Effect on Operation 302d Meeting of the American Institute of Electrical Engineers, New York.

Basquin, O.H. 1910. The exponential law of endurance tests. *Proceedings of American Society of Testing Materials*, Vol. 10, pp. 625–30.

Bauschinger, J. 1880. "Uber das Kristallinischwerden und die Festigkeitsverminderung des Eisens durch den Gebrauch." Dinglers J. 235, pp. 169–73.

Bauschinger, J. 1881. "Uber die Veranderung der Elastizitatsgrenze und des Elastizitatsmoduls verschiedener Metalle." Civilingenieur 27, pp. 289–348, Leipzig: Felix.

Bauschinger, J. 1886. "Uber die Veranderung der Elastizitatsgrenze und der Festigkeit des Eisens und Stahls durch Strecken und Quetschen, durch Erwarmen und Abkiihlen und durch oftmals wiederholte Beanspruchung." Mitt. Mech. Tech. Lab. Munch. 13, pp. 1–115.

Beau, J.F. 1964. Management of the human element in the physics of failure. *Proceedings of 3rd Annual Symposium on the Physics of Failure in Electronics*. pp. 264–79.

Birnbaum, Z.W., and S.C. Saunders. 1969. "A new family of life distributions." *Journal of Applied Probability* 6, no. 2, pp. 319–27.

Braithwaite, F. 1854. "On the fatigue and consequent fracture of metals." *Institution of Civil Engineers, Minutes of Proceedings*, Vol. 8, pp. 463–74. London.

Coffin, L.F. 1954a. "A study on the effect of cyclic thermal stresses on a ductile metal." *Transaction of American Society of Mechanical Engineers (ASME)* 76, pp. 931–50.

Coffin, L.F. 1954b. The problem of thermal stress fatigue in austenitic steels at elevated temperatures. ASTM STP No. 165, American Society for Testing and Materials, *Symposium on Effect of Cyclic Heating and Stessing on Metals at Elevated Temperatures*.

Crk, V. 2000. Reliability assessment from degradation data. *Proceedings Annual Reliability and Maintainability Symposium (RAMS)*, pp. 155–61.

Crook, D.L. 1979. Method of determining reliability screens for time dependent dielectric breakdown. *Proceedings of 17th Annual Reliability Physics Symposium*, pp. 1–7.

Darrow, K.K. June, 1942. "Entropy." *Bell System Technical Journal* 21, no. 1, pp. 51, 74, doi: 10.1002/j.1538-7305.1942.tb01301.x

Dasgupta, A. 1993. "Failure mechanism models for cyclic fatigue." *IEEE Transactions on Reliability* 42, no. 4, pp. 548–55.

Davis, H. September 26–27, 1962. Introduction. *Proceedings of the 1st Annual Symposium on Physics of Failure in Electronics*, pp. 1–3.

Dodson, G.A., and B.T. Howard. January, 1961. High stress aging to failure of semiconductor devices. *Proceedings of 7th National Symposium on Reliability and Quality Control*. Philadelphia, PA.

Epstein, B. December, 1947. "The Mathematical Description of Certain Breakage Mechanisms Leading to the Logarithmic-Normal Distribution." *Journal of the Franklin Institute* 244, no. 6, pp. 471–77.

Gertsbakh, I.B. 1989. *Statistical Reliability Theory*. New York: Marcel Dekker Inc, pp. 42–7.

Gertsbakh, I.B., and K.B. Kordonsky. 1969. *Models of Failure*. Berlin: Springer.

Graves, J. 1884. "On the Causes of Failure of Deep-Sea Cables." *Journal of the Society of Telegraph-Engineers and Electricians* 8, no. 51, pp. 119–26.

Hieber, H., and K. Pape. 1984. Lifetime of bonded contacts on thin film metallization. *Proceedings of 22nd Annual Reliability Physics Symposium*, pp. 128–33.

Hoyland, A., and M. Rausand. 1994. *System Reliability Theory, Models and Statistical Methods*. New York: John Wiley & Sons.

Jones, J.A. 1999. "A Toolkit for Parametric Drift Modelling of Electronic Components." *Reliability Engineering and System Safety* 63, pp. 99–106.

Kao, J.H.K. 1965. Statistical Models in Mechanical Reliability. *Proceedings of the 11th National Symposium on Reliability and Quality Control*. pp. 240–7.

Kapur, K.C., and L.R. Lamberson. 1977. *Reliability in Engineering Design*. New York: John Wiley & Sons.

Lawless, J.F. 1983. "Statistical Methods in Reliability." *Technometrics* 25, no. 4, pp. 305–16.

Lu, C.J., and W.Q. Meeker. 1993. "Using Degradation Measures to Estimate a Time-to-failure Distribution." *Technometrics* 35, no. 2, pp. 161–73.

Manson, S.S. 1953. Behaviour of materials under conditions of thermal stress. *Heat Transfer Symposium,* University of Michigan Engineering Research Institute, pp. 9–75.

Meeker, W.Q., and L.A. Escobar. 1998. *Statistical Methods for Reliability Data*. John Wiley & Sons.

Miner, M.A. September, 1945. "Cumulative Damage in Fatigue," *Journal of Applied Mechanics* 12, no. 3, pp. A l59–64.

Morin, A. 1853. Lecons de mecanique practique—resistance des materiaux. Paris, Librairie de L. Hachette et Cie, p. 456.

Nelson, W. 1990. *Accelerated Testing: Statistical Models, Test Plans and Data Analysis*. New York: John Wiley & Sons.

Palmgren, A. 1924. "Durability of Ball Bearings." *ZDVDI* 68, no. 14, pp. 339, (published in German).

Paris, P.C. 1964. "The Fracture Mechanics Approach to Fatigue," In *FATIGUE— An Interdisciplinary Approach*, eds. J.H. Burke, N.L. Reed, and V. Weiss. New York: Syracuse University Press.

Paris, P.C., M.P. Gomez, and W.E. Anderson. 1961. "A Rational Analytic Theory of Fatigue." *The Trend in Engineering*, 13, pp. 9–14.

Place, C.S., J.E. Strutt, K. Allsopp, P.E. Irving, and C. Trille. 1999. "Reliability Prediction of Helicopter Transmission System Using Stress-strength Interference with Underlying Damage Accumulation" *Quality and Reliability Engineering International*, 15, pp. 69–78.

Pratt and Whitney Aircraft. 1967. *Introduction to Weibull Analysis*. PWA 3001: East Hartford, CT.

Rankine, W.J.M. 1842. "On the Causes of the Unexpected Breakage of the Journals of Railway Axles, and on the Means of Preventing Such Accidents by Observing the Law of Continuity in Their Construction." Institution of Civil Engineers, *Minutes of Proceedings*, Vol. 2, pp. 105–8. London.

Shiomi, H. 1965. Cumulative degradation model and its application to component life estimation. *Proceedings of Fourth Annual Symposium on the Physics of Failure in Electronics*. pp. 74–94.

Smith, R.A. 1990. "The Versailles railway accident of 1842 and the first research into metal fatigue." In vol. 4 of *Fatigue '90*, Birmingham: EMAS.

Sobczyk, K., and B.F. Spencer. 1992. *Random Fatigue from Data to Theory*. Boston: Academic Press.

Spangenberg, L. 1874. "Uber das Verhalten der Metalle bei wiederholten Anstrengungen." Z. Bauw. 24, pp. 473–95 and 25, pp. 78–98.

Spangenberg, L. 1875. Uber das Verhalten der Metalle bei wiederholten Anstrengungen. Berlin, Ernst and Korn.

Spangenberg, L. 1879. "Uber Festigkeits-Versuche mit Eisen und Stahl." Glasers Ann. Gew. 5, pp. 6–15.

Tang, L.C., and D.S. Chang. 1995. "Reliability Prediction Using Non-Destructive Accelerated Degradation Data: Case Study on Power Supplies." *IEEE Transaction on Reliability* 44, no. 4, pp. 562–6.

Weibull, W. 1961. *Fatigue Testing and Analysis of Results*. London: Pergamon Press.

Weibull, W. November, 1951. "A Statistical Distribution Function of Wide Application." *Journal of Applied Mechanics* 18, pp. 293–7.

Wöhler, A. 1858. "Bericht iiber die Versuche, welche auf der konigl. Nieder-schlesisch- markischen eisenbahn mit Apparaten zum Messen der Biegung und Verdehung von Eisenbahnwagenachsen wahrend der Fahrt angestellt wurden." *Zeitschrift fur Bauwesen* 8, pp. 641–52.

Wöhler, A. 1860. Versuche zur Ermittlung der auf die Eisenbahnwagenachsen einwirkenden Krafte und die Widerstandsfahigkeit der Wagen-Achsen. Zeitschrift fur Bauwesen 10, pp. 583–616.

Wöhler, A. 1870a. Uber die Festigkeitsversuche mit Eisen und Stahl. Auf Anord-nung des Ministers fur Handel, Gewerbe u. offentl. Arbeiten, Grafen Itzen-blitz, angestellt. Berlin: Ernst und Korn.

Wöhler, A. 1870b. Über die Festigkeitsversuche mit Eisen und Stahl. Zeitschrift für Bauwesen 20, pp. 73–106.

Wolstenholme, L.C. 1999. *Reliability Modelling: A Statistical Approach*. New York: Chapman and Hall/CRC.

Yang, G. 2007. *Life Cycle Reliability Engineering*. New York: John Wiley & Sons.

Yang, K., and G. Yang. 1998. "Degradation Reliability Assessment Using Severe Critical Values." *International Journal of Reliability, Quality and Safety Engineering* 5, no. 3, pp. 85–95.

York, J.O. 1842. "An Account of a Series of Experiments on the Comparative Strength of Solid and Hollow Axles." *Institution of Civil Engineers, Minutes of Proceedings*, Vol. 2, pp. 105–8, London.

THERMODYNAMICS OF AGEING AND DEGRADATION IN ENGINEERING DEVICES AND MACHINES

Michael D. Bryant

3.1 INTRODUCTION TO DEGRADATION AND AGEING

Degradation and ageing of engineering devices and components compromise performance, and lead to device failure. Historically, modeling ageing and degradation were often heuristic: For the degradation under study, collect data on a variable that measures the degradation and on related phenomenological variables, and from correlations and curve fits, obtain a function. The modeling usually did not invoke fundamental laws or principles. Ageing and degradation processes are generally slow, with characteristic times much longer than operation times of the device. Degradation involves irreversible changes of material structures of components that result from dissipative processes. In the absence of shock or other catastrophes, the main drivers of component ageing are effects arising from the machine's environment, and dissipation and power losses due to machine operation. For example, pad–disc rubbing during brake use dissipates kinetic energy to slow the vehicle, and also causes brake wear. Fatigue of a machine's load-bearing member involves plasticity and fracture induced by repeated application of the machine's loads, usually

near stress concentrators. Thermal deterioration of magnet wire insulation on coils in motors and transformers arise principally from coil resistance losses. Irreversible loss of energy storage capacity in a lead acid battery arises from precipitants coming out of solution to form compounds that diminish electrochemical action in solution.

The above examples of ageing and degradation all involve irreversible dissipative physical processes, very often driven by machine power and energy losses. In accordance with the second law of thermodynamics, these dissipative processes must generate irreversible entropy. This chapter will review the thermodynamic degradation paradigm, which posits that ageing and degradation can be posed in terms of the entropy produced; review the laws of thermodynamics and thermodynamic processes; review the degradation entropy generation (DEG) theorem, which links rate of degradation to rates of production of irreversible entropy by underlying dissipative physical processes; overview many of the dissipative processes that age and degrade engineering devices and components; and apply the above to situations of sliding wear, material fatigue, and battery ageing.

3.2   THERMODYNAMIC DEGRADATION PARADIGM

Ling et al. (2002) hypothesized degradation of any type to be "a consequence of irreversible thermodynamic processes that 'disorder' components." This began a thermodynamic characterization of degradation dynamics that employed entropy, thermodynamic disorder, as an implicit but fundamental measure of degradation. Stating that entropy "offers a natural measure of component degradation, which should be a direct consequence of the second law of thermodynamics," Ling et al. (2002) proposed the thermodynamic degradation paradigm, wherein degradation of any kind can be related to the concomitant irreversible entropy generated by the dissipative processes associated with the degradation. This paradigm was formally defined and quantified into the DEG theorem, which mathematically relates degradation to entropy.

3.3   REVIEW OF THE DEG THEOREM

Pivotal to constructing equations of degradation is the DEG theorem (Bryant, Khonsari, and Ling 2008), which links rate of degradation to rates of production of irreversible entropy by the underlying dissipative

physical processes. A degradation mechanism consists of one or more dissipative processes (Bryant 2010) that together irreversibly damage or change material structures. For example, common dissipative processes that damage tribology interfaces include adhesion, surface plastic deformation, fracture, chemical reaction, material phase changes, viscous dissipation, heat dissipation, and material mixing, among others (Bryant 2010). In the DEG theorem, Bryant, Khonsari, and Ling (2008) related rate of degradation to a weighted sum of the rates of production of irreversible entropy produced by the dissipative processes underlying the degradation or ageing. Suppose irreversible degradation can be measured by a function

$$w = w\left\{ p_i\left(\zeta_i^j \right) \right\}, \tag{3.1}$$

which depends monotonically on the energies $p_i\left(\zeta_i^j \right)$ associated with the $i = 1, 2, \dots n$ underlying irreversible dissipative processes, where index j refers to the time-dependent phenomenological variables ζ_i^j that characterize each process energy p_i. To be in accordance with the second law of thermodynamics (Kondepudi and Prigogine 1998), the dissipative irreversible processes *must* produce irreversible entropy:

$$S' = S'\left\{ p_i\left(\zeta_i^j \right) \right\}, \tag{3.2}$$

that must depend on the same $p_i\left(\zeta_i^j \right)$. Since the second law demands $dS' \geq 0$, the function in Equation 3.2 must be monotonic and positive. The apostrophe notation indicates irreversible entropy. Through the second law of thermodynamics, the structure of dependency of S' on the phenomenological variables ζ_i^j is similar to that of w. The DEG theorem exploits this dependency. By applying the chain rule to Equations 3.1 and 3.2, rates of degradation and entropy are

$$\frac{dw}{dt} = \sum_{i,j}\left(\frac{\partial w}{\partial p_i}\frac{\partial p_i}{\partial \zeta_i^j} \right)\frac{\partial \zeta_i^j}{\partial t}, \quad \frac{dS'}{dt} = \sum_{i,j}\left(\frac{\partial S'}{\partial p_i}\frac{\partial p_i}{\partial \zeta_i^j} \right)\frac{\partial \zeta_i^j}{\partial t}. \tag{3.3}$$

Multiplying terms within the sum of the first of Equation 3.3 by unity in the form $(\partial S'/\partial p_i)^{-1}$ $(\partial S'/\partial p_i)$ gives

$$\frac{dw}{dt} = \sum_{i,j}\left\{ \frac{\partial w / \partial p_i}{\partial S'/ \partial p_i} \right\}\left(\frac{\partial S'}{\partial p_i}\frac{\partial p_i}{\partial \zeta_i^j} \right)\frac{\partial \zeta_i^j}{\partial t} = \sum_{i,j} B_i\left(\frac{\partial S'}{\partial p_i}\frac{\partial p_i}{\partial \zeta_i^j} \right)\frac{\partial \zeta_i^j}{\partial t}. \tag{3.4}$$

Equation 3.4 expresses rate of degradation as a sum of terms identical to those in the second of Equation 3.3, but weighted by a degradation coefficient

$$B_i = \frac{\partial w / \partial p_i}{\partial S' / \partial p_i} = \begin{cases} \dfrac{\partial w}{\partial S'}\bigg|_{p_i} & (3.5a) \\[2ex] T_i \dfrac{\partial w}{\partial p_i} & (3.5b) \end{cases}$$

that arises from the term within the curly brackets in Equation 3.4. In Equation 3.5a, the subscript p_i denotes $\partial w / \partial S'$ with dissipative process p_i active, or geometrically, $\partial w / \partial S'$ along the p_i direction in relevant multi-dimensional spaces that graph Equations 3.1 and 3.2 as surfaces versus coordinates p_i. The B_i, material properties that can be measured or related to other material properties, may be constant or depend on the thermodynamic states such as temperature. Definition of entropy (and the units involved) suggests $\partial S' / \partial p_i = 1/T_i$, where T_i is a temperature associated with process energy p_i. With this, Equation 3.5b gives another way to obtain the B_i. The terms in Equation 3.4 that multiply $\partial S' / \partial p_i$ involve dissipated power components

$$\frac{dp_i}{dt} = \sum_j \frac{\partial p_i}{\partial \zeta_i^j} \frac{\partial \zeta_i^j}{\partial t} = \sum_j e_{ij} f_{ij} \qquad (3.6)$$

associated with the ith dissipative process. The products within the sums are power conjugates involving a rate of a phenomenological variable $\partial \zeta_i^j / \partial t$ analogous to a generalized velocity f_{ij}, and its power conjugate $\partial p_i / \partial \zeta_i^j$ analogous to a generalized force e_{ij}. Equation 3.4 becomes

$$\frac{dw}{dt} = \sum_{i,j} B_i \left(\frac{\partial S'}{\partial p_i} \frac{\partial p_i}{\partial \zeta_i^j} \right) \frac{\partial \zeta_i^j}{\partial t} = \sum_{i,j} B_i \frac{e_{ij} f_{ij}}{T_i}$$

$$= \sum_i B_i \frac{dp_i / dt}{T_i} = \sum_{i,j} B_i \frac{dS'_i}{dt} \qquad (3.7)$$

The last two terms in Equation 3.7 explicitly show degradation rate to be a linear combination of the rates of entropy production (or power dissipation when relevant), where S'_i and dp_i / dt are entropy generated and power dissipated by the ith dissipative process. If entropy or power dissipated can be formulated, a quantitative degradation model can be formulated that is underpinned by the laws of thermodynamics. Thermodynamics

will now be reviewed with a goal of formulating the production of entropy of dissipative processes.

3.4 REVIEW OF THERMODYNAMICS

All systems must conserve energy, stated by the first law of thermodynamics

$$dE = dQ - dW + \sum_k \eta_k dN_k. \tag{3.8}$$

Here E is the internal energy stored within a control volume that surrounds the relevant part of the system, Q and W are heat flow and work performed across the boundary of the control volume, N_k is the number of moles of species k within the control volume, and

$$\eta_k = \frac{\partial E}{\partial N_k} \tag{3.9}$$

is the chemical potential associated with species k. Here, passage of time dt is implicit in the differentials. The second law of thermodynamics states that a change in the entropy S within the control volume

$$dS = dS' + dS_e \tag{3.10}$$

consists of a reversible change dS_e from flow of entropy into and out of the control volume, and an irreversible change that arises from generation of entropy dS' inside the control volume. Entropy flow dS_e arises from heat transfer

$$TdS_e = dQ + \sum_k \eta_k d_e N_k. \tag{3.11}$$

due to flow of heat dQ (via conduction and radiation) and flow of matter $d_e N_k$ across the control volume boundary. Here, chemical potential η_k gauges the energy contained per mole of matter. Change in number of moles

$$dN_k = d'N_k + d_e N_k \tag{3.12}$$

consists of a reaction term $d'N_k$ (chemical or phase change) and the matter transport term $d_e N_k$. Note the terms $d_e N_k$ extend the foregoing equations to open systems.

3.5   ENTROPY AND PRODUCTION OF IRREVERSIBLE ENTROPY

The traditional inequality form $dS \geq 0$ of the second law, Equation 3.10, arises from the nonnegative production $dS' \geq 0$ of entropy. Since entropy measures randomness and disorder, $dS \geq 0$ suggests a tendency for a system to become disordered. All natural processes, including degradation, proceed so as to increase the total entropy of a system and its surroundings. A system produces entropy, $dS' > 0$, until equilibrium. At equilibrium, a system's entropy is maximum and entropy production ceases, $dS' = 0$ (Kondepudi and Prigogine 1998). Once created, entropy cannot be destroyed, but only transported away via entropy flow dS_e. Entropy flow allows reversible processes.

To assess degradation, Equation 3.7 demands assessment of entropy production S', which will be determined using Equations 3.8 to 3.12. Substitution of Equation 3.12 into Equation 3.8 with grouping of terms according to Equation 3.11 shows the presence of term TdS_e on the right side of Equation 3.8. Substitution of Equation 3.10 for dS_e into Equation 3.8 for the "terms" that comprise TdS_e, and solving for entropy production gives

$$dS' = dS - \frac{dE}{T} - \frac{dW}{T} + \frac{1}{T}\sum_k \eta_k d'N_k .$$

(3.13)

Besides the work dissipated (Klamecki 1984; Frederick and Chang 1965; Bejan 1988) and material reaction terms (Kondepudi and Prigogine 1998), Equation 3.13 asserts that entropy production also arises from internal energy and entropy changes dE and dS within the control volume. To assess entropy production and thus degradation, conditions must be placed on dE and dS within the control volume. For other thermodynamic conditions, the thermodynamic potentials

$$A = E - TS, \qquad H = E + pV, \qquad G = E - TS + pV$$

(3.14)

of Helmholtz free energy, enthalpy, and Gibbs free energy, respectively, can be applied. Under these three definitions, Equation 3.13 becomes, respectively,

$$dS' = -\frac{SdT}{T} - \frac{dA}{T} - \frac{dW}{T} + \frac{1}{T}\sum_k \eta_k d'N_k,$$

(3.15)

which needs conditions on dA and dT;

$$dS' = dS + \frac{Vdp}{T} - \frac{dH}{T} - \frac{dW'}{T} + \frac{1}{T}\sum_k \eta_k d'N_k, \qquad (3.16)$$

which needs conditions on dH and dp; and

$$dS' = -\frac{SdT}{T} - \frac{dG}{T} + \frac{Vdp}{T} - \frac{dW'}{T} + \frac{1}{T}\sum_k \eta_k d'N_k, \qquad (3.17)$$

which needs conditions on dG, dp, and dT. In Equations 3.16 and 3.17, $dW' = dW - pdV$ denotes work other than pressure volume.

3.5.1 THERMODYNAMIC SYSTEM CONDITIONS

Stationary systems: The simplest condition for Equation 3.13 is the stationary system with $dE = dS = 0$ (constant internal energy and entropy). Here, combinations of flows across, and production inside the control volume of energy and entropy balance, rendering constant levels of internal energy and entropy within the control volume. Stationary systems are at a steady state but not in equilibrium.

Systems under strain and isothermal: Conditions for Equation 3.15 involve dA and dT. Helmholtz free energy A, energy untainted by entropy or exergy, can include energy associated with conservative forces (such as elastic strain energy of deformed material dependent on displacements), and the work of isothermal separation of surfaces, important for quantifying fracture and fatigue phenomena, among others. For isothermal systems, $dT = 0$.

Isentropic and isobaric systems: Conditions for Equation 3.16 involve dH, dS, and dp. Enthalpy H, with pressure volume expansion work subtracted from internal energy, pertains to gases and phase changes of materials. For isothermal systems, $dp = 0$.

Isothermal and isobaric systems: Conditions for Equation 3.17 involve dH, dT, and dp. Gibbs free energy G, with pressure volume expansion work subtracted from internal energy, pertains to $dT = dp = 0$ (constant temperature and pressure) conditions under which chemical reactions are standardized. For isothermal systems, $dp = 0$. This dissipated work must eventually diffuse through heat flow dQ or mass flow $\sum \eta_k dN_k$ or both, as entropy flow dS_e. Open systems demand balancing flows of entropy, heat, work, energy, and mass over a control volume about the degrading body or system.

3.6   DISSIPATIVE MECHANISMS AND AGEING

Ageing and degradation mechanisms consist of one or more dissipative processes operating simultaneously or in sequence. Dissipative processes drive degradation and ageing. Bryant (2010) identified several common dissipative processes operative at tribological interfaces and estimated the irreversible entropies produced. This section will present the results of Bryant (2010).

   Adhesion of surfaces and films onto other bodies generate entropy

$$\Delta S' = \frac{\Delta \gamma}{T_m} \Delta A_s,$$ (3.18)

where the interface surface energy $\Delta \gamma$ is the work per unit area needed to separate surfaces in a solid to create new surface area ΔA_s, and T_m is the local temperature of the media (Maugis 1999). Surface energies γ are 1 to 3 Jm^{-2} for clean metals, 0.1 to 0.5 Jm^{-2} for ionic crystal ceramics, and less than 0.1 Jm^{-2} for molecular crystals and polymers (Maugis 1999).

   Plastic Deformation, Viscous Dissipation, and Cutting of solids induce an entropy change

$$\Delta S' = \frac{U_c}{T_m} \Delta V,$$ (3.19a)

where $U_c = dW_p/dV$ is the work of plastic deformation W_p or cutting expended per unit volume ΔV of the affected zone, and T_m is the temperature of the affected material media. Plastic flow in ductile metals occurs when the metal is overloaded in shear. The dissipation function of continuum mechanics for large deformations gives

$$U_c = \int \sigma d\varepsilon.$$ (3.19b)

   Here σ is the stress tensor and ε is the strain tensor. The form of Equation 3.19b is valid for solids and fluids. For plastic dissipation in solids, σ is the deviatoric stress tensor.

   Fracture separates continuous materials and forms new free surfaces within the cracked body. Overstressed bodies fracture, especially brittle materials. Fracture is associated with rupture of bodies, fatigue of bodies, and surface damage. For a crack of length a, in a linearly elastic brittle material, under plane strain, at uniform temperature T_{cr}, with no material

loss, and no chemical reactions or phase changes, Rice (1978) derived the irreversible entropy produced

$$dS' = \frac{G - 2\gamma_o}{T_{cr}} da,$$

(3.20)

where energy release rate $G = -\partial U_s/\partial a$ is the energy released from the elastic strain energy U_s per increment of crack extension and γ_o is the surface energy needed to establish free surfaces in the solid under isothermal conditions.

Phase changes, including melting, solidification, crystallization, and evaporation, alter material structure. Temperature ageing can induce a phase change. Material properties such as strength or elastic modulus of components can degrade. Phase changes, associated with surface melting and recrystallization of metals, induce entropy change

$$\Delta S' = \frac{\Delta H}{T_{phase}}$$

(3.21)

Here, enthalpy change *DH* is the latent heat absorbed or shed during the phase change, and T_{phase} is the temperature associated with the phase change. For solid to liquid and vice versa, T_{phase} is the melting temperature, ΔH is the molar enthalpy of fusion, and $\Delta S'$ is the molar entropy of fusion. From liquid to gas, T_{phase} is the boiling temperature, ΔH is the molar enthalpy of vaporization, and $\Delta S'$ is the molar entropy of vaporization. Phase change occurs quickly enough such that the phase change process can be considered instantaneous, rendering rates immaterial.

Chemical reactions alter the composition of materials, which changes material properties and behavior of components. For a system at constant temperature and pressure $dT = dp = 0$, and no other work crossing the system boundary $dW' = 0$, the chemical reaction

$$\sum_{reactants} \kappa_i r_i \rightarrow \sum_{products} \kappa_i p_i.$$

(3.22a)
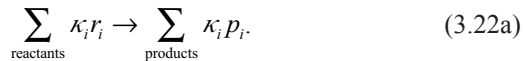
involving reactant species r_i and product species p_i at temperature T_m generates entropy

$$dS' = \frac{A}{T_m} d\xi, \quad A = \sum_{reactants} \kappa_i \eta_i - \sum_{products} \kappa_i \eta_i.$$

(3.22b)

Chemical affinity A depends on the chemical potentials η_i and the coefficients κ_i in the stoichiometric Equation 3.22a that specifies the overall chemical reaction. Since the stoichiometry of the reaction relates to changes of molar masses dN'_i of reactants and products involved in the reaction, the extent of the reaction ξ has differential change

$$d\xi = \pm \frac{dN'_i}{\kappa_i} \qquad (3.22c)$$

for each reactant or product involved in the reaction, where minus pertains to reactants, and plus pertains to products. The number of moles $dN'_i = \pm\kappa_i d\xi$ of reactants and products in the reaction depend on the stoichiometric coefficients κ_i.

To illustrate, oxidation of metallic copper to cuprous oxide involves intermediate reactions that form copper ions, transfer electrons to oxygen, and form the final products. The reaction has three steps, but the entropy generated depends only on the overall reaction stoichiometry $4Cu + O_2 \rightarrow 2Cu_2O$. The chemical affinity $A = 2\eta Cu2O - 4\eta Cu - \eta O_2$ and the extent of reaction $d\xi = -dN'Cu/4 = -dN'O_2 = dN'Cu_2O/2$. The numbers in the denominators are the stoichiometric coefficients from the overall chemical reaction equation. Thus, $dS' = (2\eta Cu_2O - 4\eta Cu - \eta O_2)\, dN'Cu_2O/2T_m$. Chemical potentials η_i for most elements and compounds are well tabulated in *CRC Handbook of Chemistry and Physics* (Lide 2006).

Diffusion migrates material (molecules, ions, or particles) from regions of higher concentration to regions of lower concentration. Concentration differences or gradients can alter thermodynamic states, thermodynamic energies, and chemical potentials, see Equations 3.8, 3.9, and 3.14. The mass gained by a region of lower concentration during diffusion from a region of higher concentration must equal the mass lost from the region of higher concentration, that is, $dN_{low} = -dN_{high}$. For entropy generation, diffusion is viewed as a "reaction" (Kondepudi and Prigogine 1998), with higher concentration material being "reactant," and material transported to regions of lower concentration being "product." Rate of entropy production is then governed by Equations 3.22b and 3.22c, but with chemical potentials η_{low} and η_{high} principally determined by concentrations of species, and Equation 3.22c defined by mass conservation. For diffusion, T in Equation 3.22b is the temperature of the diffusion media. For the simple example cited, Equation 3.22c becomes $d\xi = dN_{low} = -dN_{high}$, affinity of Equation 3.22b becomes $A = \eta_{high} - \eta_{low}$, and the rate of entropy production

$$\frac{dS'}{dt} = \frac{\eta_{high} - \eta_{low}}{T} \frac{d\xi}{dt}$$ (3.23)

Mixing of n different material species generate entropy

$$\Delta S' = -R \sum_{i}^{n} \frac{N_i}{N} \ln \frac{N_i}{N}, \quad N = \sum_{i}^{n} N_i.$$ (3.24)

Here R is the universal gas constant, N_i denote molar masses of the $i = 1, 2, \ldots, n$ different species, and ratios N_i/N are molar fractions. Before mixing, the pure species were in separate volumes. Equation 3.24 was derived as the entropy of the final mixed state minus the sum of the entropies of the initial pure states.

Heat transfer, associated with diffusion of heat dQ from a region of higher temperature T_h to a region of lower temperature T_l, generates entropy.

$$dS' = \left(\frac{1}{T_l} - \frac{1}{T_h} \right) dQ.$$ (3.25)

Transferring heat dQ from a hotter body to a cooler body reduces the hotter body entropy by dQ/T_h, but increases the cooler body entropy by dQ/T_l. The difference is the entropy generated as shown in Equation 3.25.

Thermal expansion is the tendency for material bodies to expand or dilate under increased temperature. For unrestrained free expansions, the deformations are usually elastic and almost reversible. Here, the entropy generation by the underlying dissipative processes is via heat transfer, Equation 3.25. If thermal expansions are restrained and this results in irreversible material changes such as plastic deformation or fracture, in addition to heat transfer, entropy would be generated by plastic deformation, Equations 3.19, and fracture, Equation 3.20.

For each of these dissipative processes, the entropy produced has the form of a product of a weighting term and a differential or difference of a phenomenological variable.

Generalizing, each of the entropies can be written in one of the forms:

$$dS' = Xd\zeta, \quad \Delta S' = X\Delta\zeta$$ (3.26)

where X is the weighting function and ζ is the phenomenological variable.

3.7   EXAMPLE APPLICATIONS OF THE DEG THEOREM

Presented in this section will be examples of component degradation involving wear, fatigue, battery degradation, and grease degradation. For each of these examples, the dissipative process will be identified and the DEG theorem will be applied. For wear and fatigue, measured data for degradation and entropy production will be presented to support the analysis.

3.7.1   DEGRADATION BY WEAR

Rubbing one body against another induces friction forces between bodies and simultaneously generates sliding wear. Wear, defined here as loss of volume from material during rubbing, irreversibly reorders the rubbed material, and as such, irreversible entropy must be produced. Friction force F, and wear w, are manifestations of the same dissipative processes that occur at the interface between rubbing bodies. Historically, friction and wear were often treated as separate phenomena, but thermodynamics and the DEG theorem admit a unified approach (Bryant, Khonsari and Ling 2008). Many sliding wear systems operate at a steady state, wherein $dE = dS = 0$, which suggests use of Equation 3.13. With no reactions or phase change of material ($dN'k \approx 0$) within the control volume, Equation 3.13 asserts dissipated work to be the prevalent entropy producer. Since the direction of the friction force vector always opposes the direction of the sliding velocity vector during rubbing, with $v = dx/dt$ in Figure 3.1(a), friction dissipates power $-Fv$ and via Equation 3.13 $dS'/dt = Fv/T$.

Through the DEG theorem of Equation 3.8, the rate of degradation by wear is

$$\frac{dw}{dt} = B\frac{dS'}{dt} = B\frac{Fv}{T} = B\frac{\mu Nv}{T},\qquad (3.27a)$$

where $F = \mu N$ describes coulomb friction with friction coefficient μ, and temperature T of the contact region between the rubbing bodies is affected by the friction heating and heat transfer. Equation 3.27a can be compared to the classic wear law of Archard (1953) $w = kNx/H_m$ expressed in the rate form $dw/dt = kNv/H_m$. Here, H_m is the hardness pressure (Rabinowicz 1965, 1980) of the softer sliding pair and k is the nondimensional wear coefficient. Dividing Equation 3.27a by the rate of Archard's law gives

(a)

(b)
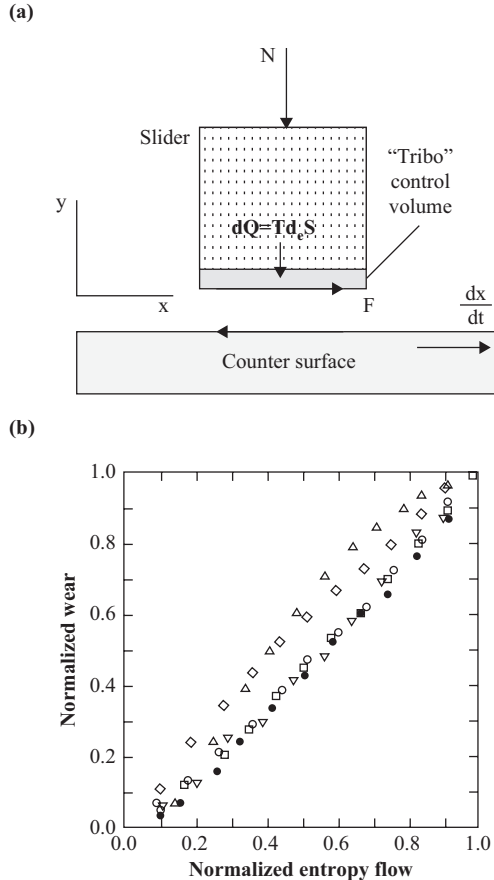
Figure 3.1. (a) Sliding wear with friction, and (b) wear versus entropy.

$$k = B \frac{\mu H_m}{T}.$$
(3.27b)

Alternatively, this k to B relation can be obtained via Equation 3.5b with $T_i = T$ and Archard's law rewritten as $w = k p_i / \mu H_m$, wherein the process energy $p_i = Fx = \mu N x$. Since, for steady sliding, T, k, H_m, and μ are constants, wear B should be constant.

Figure 3.1(b) depicts a wear test machine from Ling et al. (2000), wherein a copper rider (E) loaded by a 9.1 kg dead weight (A) slid over the circumference of boundary lubricated steel rotor (D) rotating at 1,000 rpm, rendering $v = 3.3$ ms^{-1} surface speed. Three thermocouples in (F) measured

the temperature of the sliding interface (E–D) to estimate conduction up through the rider of heat generated by sliding friction losses. The cylindrically shaped rider was carefully insulated about its outer radius to encourage vertical conduction heat flow. Rider wear was measured as a displacement loss by photonic gap sensor (H). Since sliding of the rider was kept in a stationary state, $dS = 0$, leading to $dS' \approx -dS_e$. Entropy flow $\sum_n \Delta Q^{(n)} / T^{(n)}$ was estimated from the temperature measurements for heat flow $\Delta Q^{(n)}$ and the interface temperatures $T^{(n)}$ during the nth time interval. Wear w versus entropy flow S_e measurements are plotted in Figure 3.1(b). The coordinate axes were normalized by maximum values of wear w_{max} and entropy flow S_{emax}. Since the slope of this curve in a nondimensional space is approximately 1, Equation 3.5a suggests that $B = w_{max}/S_{max} = 4.0 \times 10^{-10}$ m³/(JK⁻¹). This B value, along with handbook values of μ and H_m, rendered $k = B_\mu H_m/T \approx 1.01 \times 10^{-4}$. This k value, derived from the measurements of wear, force, and *temperatures*, was within a percent of Rabinowicz's (1980) value, derived from measurements of wear, force, and *distance slid*. Here, phenomenologically different measurements rendered nearly identical k values.

Other sliding wear tests (Bryant and Khonsari 2008) with different materials, different geometry, and under different test conditions gave similar results, further supporting the thermodynamic degradation paradigm and the DEG theorem. Fretting wear, which occurs via a different mechanism and dissipative processes, also follows the thermodynamic degradation paradigm. Following Huq and Celis (2002) and Fouvry et al. (1997, 2001, 2003), the dominant dissipative process p is the work of plastic deformation, driven by friction force F. Through the DEG theorem (Bryant, Khonsari, and Ling 2008) as posed by Equation 3.7,

$$\frac{dw}{dt} = B\frac{dS'}{dt} = B\left[2(\delta e)(\delta f)\right] = 2B\frac{F}{T}\frac{dx}{dt} = \left(\frac{2B}{T}\right)\frac{dE_\mu}{dt}. \quad (3.27c)$$

The delta notations on e and f indicate perturbations from the current equilibrium state. The 2 arises from the second variation of entropy (Dai, Yang, and Xue 2000), and T is the near constant temperature at equilibrium. Here, the volume of material lost w is proportional to the friction energy $E\mu$ dissipated, which is observed in fretting wear (Fouvry et al. 1997, 2001, 2003).

3.7.2   DEGRADATION BY FATIGUE

Damage from fatigue weakens many metals, ceramics, composites, and polymers, rendering components susceptible to rupture under mechanical

loads normally too small to cause rupture. Under a repetitive subcritical cyclic load such as reversed bending or torsion, the many small flaws (stress concentrators or elevators) present in a material structure elevates stresses, resulting in plastic deformations and eventual creation of small cracks in the material structure (Juvinall and Marshek 1991). Continued cyclic loading accumulates crack damage, weakening the material structure. This affects material properties such as elastic modulus; fatigue strength S_M, the load or stress needed to cause rupture after logging M cycles; or material toughness U_M (Duyi and Zhenlin 2001), the amount of dissipated strain energy that the part can absorb before rupture. After sufficient weakening, strength diminishes to load, and the component fails. The number of cycles M_f to failure often serves as a fatigue life or failure criteria. Observations distilled into the measured S–N curve of Figure 3.2, taken from Juvinall and Marshek (1991), show declining fatigue strength

$$S_M = S_M(M) \tag{3.28}$$

versus number of cycles M. Under different fatigue conditions (load, geometry, kinematics, environment, etc., indexed by k), the diminished strength S_M equals the load after M_f^k cycles, and the component fails. For a same member subjected to different fatigue conditions during portions of its life, cycles $M_k < M_f^k$ represent the fraction $M_k/M_f^k < 1$ of the member's life under the kth load conditions. For a member sequentially loaded under K different load characteristics, Miner's (1945) rule

$$\sum_k^K \frac{M^k}{M_f^k} \leq 1 \tag{3.29}$$

$S = 0.9S_u$ (in ksi, $S \approx 0.45 \cdot$ Bhn; in MPa, $S \approx 3.10 \cdot$ Bhn)

$S'_n = 0.5S_u$ (in ksi, $S'_n \approx 0.25 \cdot$ Bhn; in MPa, $S'_n \approx 1.73 \cdot$ Bhn)
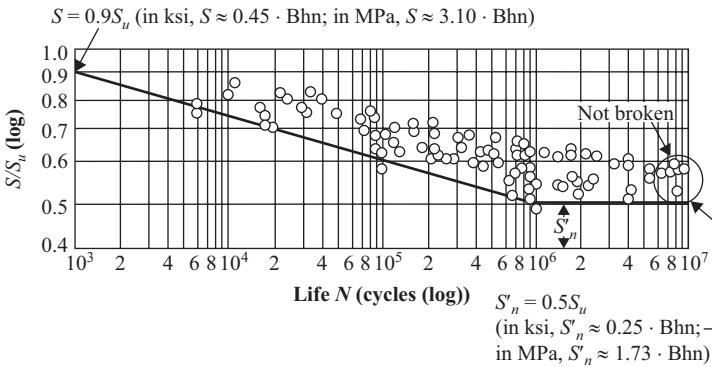
Figure 3.2. S–N diagram showing reduction of strength versus number of cycles.

states that the sum of the fractions M_k/M_f^k of the life spent under each load condition must not exceed unity. From this, the remaining number of cycles under different conditions can be estimated.

Amiri and Khonsari (2012), Naderi and Khonsari (2010), and Amiri, Naderi, and Khonsari (2011) have shown that fatigue of some metals, composites (Naderi and Khonsari 2012), and polymers obey the thermodynamic degradation paradigm and the DEG theorem. Collectively, these references experimentally correlate the cumulative effects of fatigue damage with the entropy produced in a fatiguing member. Amiri, Naderi, and Khonsari (2011) formulated an expression for the irreversible entropy S' generated during progression of fatigue, and measured entropy flow S_e under an approximately stationary process $dS = 0$. Through Equation 3.10, $dS_e = - dS'$, which allows entropy flow to estimate entropy generation. Naderi, Amiri, and Khonsari (2010) determined entropy flow via infrared temperature measurements on a cantilevered member undergoing reverse bending, and used finite element calculations to estimate stress, the dissipated work of plastic deformations, the heat released from these dissipations, and the entropy flow associated with the heat flow over the fatigued member. The temperatures estimated by finite element were consistent with the infrared measured temperatures.

During the first hundred fatigue cycles or so, temperatures increased due to heat of plastic work dissipation (Naderi, Amiri, and Khonsari 2010; Amiri, Naderi, and Khonsari 2011), see Figure 3.2. During the second phase of thousands of cycles, temperatures stabilized at near constant levels set by equilibrium between heat transfer and the dissipated work. Finally, just before the end of the component's fatigue life, temperatures abruptly increased, and Naderi, Amiri, and Khonsari (2010) proposed temperature to predict onset of catastrophic failure. Similar to Figure 3.1, Naderi, Amiri, and Khonsari (2010), Amiri, Naderi, and Khonsari (2011), and Naderi and Khonsari (2011) graphed normalized cycles M/M_f versus normalized entropy S_e/S_f, see Figure 3.3, where M_f and S_f are cycles and entropy at specimen rupture. Collectively, the articles mentioned varied fatigue conditions—geometry, displacement amplitude, frequency, and kinematics——and load type——reversed bending and reversed torsion. All cases rendered the linear relation

$$\frac{S_e}{S_f} \approx \frac{M}{M_f},\qquad(3.30)$$

suggesting failure after generating a critical cumulative entropy S_f, or logging total number of cycles M_f Naderi, Amiri, and Khonsari (2010)

measured S_f of 4 MJ m^{-3} K^{-1} for Al 6061-T6 and 60 MJ m^{-3} K^{-1} for SS 304. These measured values were found to be independent of fatigue geometry and load characteristics. Through independent measurements and a probabilistic analysis of elastic and plastic strains and stresses for fatigue of aluminum rod and sheet specimens, Whaley (2012) also discovered a critical entropy S_f to fatigue failure. This is consistent with the philosophy and experience of Sosnovskiy and Sherbakov (2009), who, after studying many types of man-made and biological systems, have concluded that failure occurs whenever an item of any kind has experienced a critical amount of entropy.

Substituting Equation 3.30 into Miner's Equation 3.29 gives $\sum_{k}^{K} \dfrac{S^k}{S_f^k} \leq 1$, showing that Miner's rule is equivalent to an entropy statement. This was also found by Amiri and Khonsari (2012) via a different analysis. As discussed in the previous paragraph, since Naderi, Amiri, and Khonsari (2010) and Whaley (2012) found that critical entropy S_f depended only on material type and not fatigue geometry and load conditions, all $S_f^k = S_f$, rendering

$$\sum_{k}^{K} S^k \leq S_f. \tag{3.31a}$$

Equation 3.31a accumulates entropy, which can be stated in integral form:

$$\int dS' \leq S_f. \tag{3.31b}$$

Equation 3.31 implies that whatever conditions (load, geometry, load kinematics, load type, etc.) exist, a fatigued member fails after generation of total entropy S_f, suggesting total entropy accumulation tantamount to number of cycles for fatigue life.

Fatigue is a slow process with almost steady temperatures (Amiri, Naderi, and Khonsari 2011) and negligible material loss or change, giving $dT \approx d'N_k = 0$. Systems under these conditions can be described by Equation 3.15. Here, the entropy generated per cycle dS' arises from the applied work dW and Helmholtz free energy dA—which includes stored elastic strain energy and surface energy for cracking—via dissipative processes of plastic deformation and cracking; in comparison, Amiri, Naderi, and Khonsari (2011) showed that the entropy generated by heat flow per cycle is negligible. Consideration of Figure 3.2 and Equation 3.28 in light of Equations 3.30 and 3.31 suggest that fatigue strength

$$S_M = S_M(S'). \tag{3.32a}$$

Integration over time of Equation 3.7 with degradation measure w taken to be S_M gives

$$S_M - S_0 = \int B \frac{dS'}{dt} dt = \int B dS'. \tag{3.32b}$$

Comparison of Equations 3.32 suggests validity of the DEG theorem for application to fatigue. Figure 3.2 suggests the degradation coefficient $B < 0$ to denote diminishing strength. Through Equation 3.5, $B = \partial S_M / \partial S'$, but with entropy generation proportional to number of cycles for a given test on a specimen, $B = (\partial S_M / \partial M)/(\partial S'/\partial M) = (\partial S_M / \partial M)(M_f / S_f)$ and the strength equation can be written in terms of entropy generation $S_M = S_u (M_f/S_f S')^{-\alpha}$. For this case, B is not constant but is a function of entropy.

For Figure 3.2, the curve through the data points, a straight line with slope $-\alpha$, $\alpha > 0$, in the log–log space, if normalized by initial strength S_0 and cycles M_e to the endurance limit, can be approximated as[1]

$$\frac{S_M}{S_0} = \left(\frac{1 + M / M_e}{M} \right)^{\alpha} \tag{3.33a}$$

$$U = U_0 + \frac{U_0 - U_{M_f - 1}}{\ln(M_f)} \ln(1 - M/M_f) \tag{3.33b}$$

where subscript 0 refers to the initial rupture strength. When sufficient entropy accumulates, fatigue strength S_M equals the applied load, and the specimen ruptures.

3.7.3 BATTERY DEGRADATION

Batteries store energy electrochemically and release energy as electric power. Popular battery types include lead acid and lithium ion batteries. Batteries consist of anode and cathode electrodes, electrolyte, separator, and terminals. Batteries have finite lifetimes, which are usually limited by manufacturing defects and ageing effects. This section will focus on the ageing effects on battery life. Battery health is usually measured in terms of capacity C (Ah), the amount of charge in ampere-hours a battery can deliver when discharged at a rated current, and/or the growth of

---

[1]In a log-log space for $M \ll M_e$, Equation (3.28) will plot roughly as a straight line of slope $-a$, similar to an integrator in a Bode magnitude plot. For $M \gg M_e$, the slope will vanish due to the "zero" like term in the numerator, forming an S-N "knee".
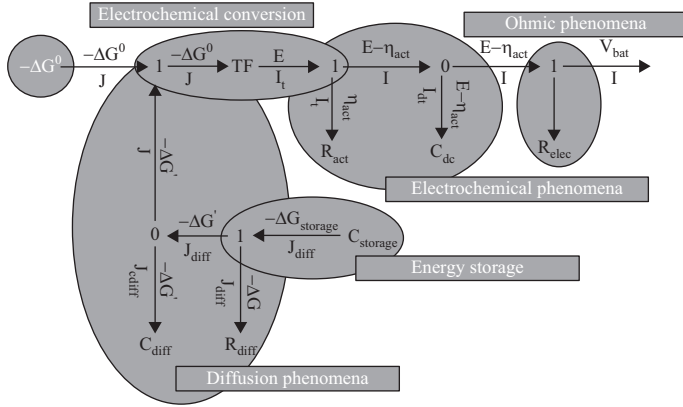
Figure 3.3. Bond graph systems dynamic model of a Li ion battery.
Source: Adapted from Ménard, Fontès, and Astier (2010).

internal cell impedance Z (ohms) (see Broussely et al. 2005), where the internal cell impedance Z pertains to a Thevenin equivalent circuit of the battery. Ageing reduces capacity C and increases impedance Z. Cycles of charging and discharging age a battery. Battery cycle life is rated as the number of complete charge–discharge cycles a battery can undergo before (1) capacity C falls below 80 percent of initial rated capacity, and/or (2) the internal resistance Z increases 1.3 to 2 times its initial value.

Battery life, typically 500 to 1,200 charge discharge cycles, depends on many factors (Broussely et al. 2005 and Vetter et al. 2005). Most prevalent are:

a. Number of charge–discharge cycles experienced by the battery: More cycles diminish remaining life.
b. Depth of discharge (DOD), the percent of battery capacity discharged during a charge–discharge cycle: A larger DOD reduces cycle life and increases the increment of energy dissipated per cycle.
c. Electrolyte decomposition enhanced by high temperature and Li plating.
d. Electrode plating by Li, which increases resistance and fades capacity, is exacerbated by lower temperature.

The effects on battery life stated in the list are summarized in Table 3.1.

A battery operational model in bond graph form from Ménard, Fontès, and Astier (2010) models the dynamic electrochemical phenomena in a Li-ion battery. The bond graph of Ménard, Fontès, and Astier (2010) was copied and is presented as Figure 3.3. A bond graph maps

Table 3.1. Observations of various physical effects on battery life

Observation	Source	Incorporated into model
"Most…effects can be considered as additive"	Vetter et al. (2005)	Sum over terms from dissipative effects
More charge/discharge cycles diminish capacity	Broussely et al. (2005)	More cycles accumulate more entropy
Larger DOD (percent of battery capacity discharged per cycle) reduces life	Broussely et al. (2005) Electropaedia (2005)	Increases energy dissipated per cycle. Entropy produced accumulates faster.
Electrolyte decomposition enhanced by high temperature, high state of charge	Vetter et al. (2005)	Expression terms $$B_{diff} \frac{R}{\alpha FI_{\lim}} I_t + B_{act} \frac{R}{\alpha FI_o} I_t^2$$
Electrode plating by Li, which increases resistance and fades capacity, exacerbated by higher charging rate and lower temperature	Broussely et al. (2005) Vetter et al. (2005)	Expression term $$\frac{B_{elec}}{T_{elec}} R_{elec} I^2$$

where and how power flows in a physical system, and where energy is stored and dissipated. The half arrows in Figure 3.3 indicate the direction of positive power flow in a Li-ion battery. In a bond graph, potential energy is stored in capacitance elements C, kinetic energy is stored in inertance elements I, and power is dissipated in resistance elements R. From a completed bond graph, the differential equations that govern the physics and dynamics of a system can be extracted (Brown 2006; Karnopp, Margolis, and Rosenberg 2000). The bond graph of Figure 3.3 labels where in the battery system are the "energy storage," "diffusion phenomena," "electrochemical conversion," "electrochemical phenomena," and "ohmic phenomena." On the far right of Figure 3.3, battery terminal voltage V_{bat} and current I appear across the terminals of a physical battery. Chemical capacitance $C_{storage}$ in "energy storage" stores the battery energy via electrochemical charge separation involving Li^+ ions and electrons. Gibbs free energy $-\Delta G_{storage}$ (J) and molar flow of lithium ions J (mol sec^{-1}) appear as effort and flow on multiple bonds, indicating the chemical thermodynamics embedded in this bond graph. The

minus sign on $-\Delta G_{storage}$ refers to energy leaving the main storage $C_{storage}$. In "diffusion phenomena," capacitance C_{diff} and resistance R_{diff} together set a time constant $C_{diff} R_{diff}$, which controls the slow dynamics of Li$^+$ ion diffusion in the electrolyte, which transports charge through the electrolyte. Transformer TF:nF in "electrochemical conversion" has the modulus of the Faraday constant F (9.649×10^4 C mol^{-1}) and the number of moles of electrons n exchanged for each mole of lithium ions involved in the electrochemical reaction at the electrodes. This transformer element converts electrochemical power to electrical power. Effort source $-\Delta G^0$ establishes a reference thermodynamic potential within the bond graph. The power needed to activate electrochemical reactions at the electrode–electrolyte interfaces imposes resistance R_{act}. This power is dissipated. Capacitance C_{dc} is due to a layer of charge (electrons and lithium ions Li$^+$) that forms about the electrode–electrolyte interface. Low mobility of Li$^+$ ions through the electrolyte relative to electron flows in the battery causes ohmic resistance R_{elec}.

In electric circuits, electric resistances dissipate power $V I$, where V and I are the voltage drop across and current through the physical resistor. The electric resistance R_{elec} in Figure 3.3 has voltage drop η_{elec} and current I. Similar to an electrical resistance, the power dissipated by other resistances is the product of the equivalent voltages and currents, shown as labels on the half arrow bonds. In the bond graph of Figure 3.3, resistances R_{elec}, R_{diff}, and R_{act} dissipate powers $\eta_{elec} I$, $-\Delta G_{diff} J_{diff}$, and $\eta_{act} I_t$, respectively. The entropy generated dS' is the dissipated power divided by a temperature associated with the dissipative process. Thus from the resistance elements and considerations associated with Equation 3.7,

$$\frac{dS'}{dt} = \frac{-\Delta G_{diff} J_{diff}}{T_{diff}} + \frac{\eta_{act} I_t}{T_{act}} + \frac{\eta_{elec} I}{T_{elec}}. \tag{3.34a}$$

By way of Equation 3.7, the battery degradation is

$$\frac{dw}{dt} = B_{diff} \frac{-\Delta G_{diff} J_{diff}}{T_{diff}} + B_{act} \frac{\eta_{act} I_t}{T_{act}} + B_{elec} \frac{\eta_{elec} I}{T_{elec}}, \tag{3.34b}$$

where degradation measure w can be battery capacity C and/or internal impedance Z. Temperatures T_{diff}, T_{act}, and T_{elec} are associated with the diffusion, activation, and electric domains of the battery. Equation 3.34b relates rate of capacity or impedance change (Erdinc, Vural, and Uzunoglu 2009) to power dissipated, and sums dissipative effects from Li-ion diffusion into/out of electrodes, the energy of activation of Li/Li-ions at electrodes,

and ohmic effects associated with mobility of Li ions in electrolyte. As the first row of Table 3.1 states, the sum over effects in Equation 3.34b is consistent with Vetter et al. (2005), who reviewed ageing mechanisms in Li-ion batteries and stated that diverse "effects can be considered as additive." Equation 3.34b is consistent with item (b) of the list, since a larger DOD yields larger power dissipations from all effects, with greater per cycle changes in w, and with Broussely et al. (2005) who found capacity faded and impedance increased with more charge–discharge cycles.[2] In Equation 3.34b, coefficients B_{diff}, B_{act}, and B_{elec} should be adjusted to reflect the relative importance of each entropy term on the degradation. For w being C or Z, these coefficients must be negative or positive, respectively, to model capacity fade or impedance increase. Equation 3.34b can be posed in terms of phenomenological variables via constitutive relations of Equations 14 and 17 of Ménard, Fontès, and Astier (2010), wherein

$$-\Delta G_{diff} J_{diff} = \frac{RT_{diff}}{\alpha FI_{lim}} I_t, \qquad \eta_{act} I_t = \frac{RT_{act}}{\alpha FI_o} I_t^2, \qquad \eta_{elec} I = R_{elec} I^2 \quad (3.35a)$$

giving

$$\frac{dw}{dt} = B_{diff} \frac{R}{\alpha FI_{lim}} I_t + B_{act} \frac{R}{\alpha FI_o} I_t^2 + B_{elec} \frac{R_{elec} I^2}{T_{elec}}. \qquad (3.35b)$$

Here R is the molar gas constant, $\alpha = 1/2$, and I_{lim} and I_o are diffusion currents dependent on the number of lithium ions. Cancellation of T_{diff} and T_{act} between numerators of Equation 3.35a and denominators of Equation 3.34b rendered the first two terms of Equation 3.35b explicitly independent of temperature. These terms, which pertain to electrolyte diffusion and electrode–electrolyte interface activation, should increase with temperature, for the model to be consistent with row five of Table 3.1 and item (c) of the list. This suggests that the coefficients B_{diff} and B_{act} should monotonically increase with temperature, in a manner consistent with the observations. Note that this is consistent with remarks that followed Equation 3.5. The last term of Equation 8b has T_{elec} in the denominator, suggesting more influence from this term at lower temperature, consistent with item list (d) and the last row of Table 3.1.

---

[2]With each increment of energy dissipated during each charge–discharge cycle an increment of entropy must be produced, via the second law of thermodynamics. As cycles accumulate, the entropy produced accumulates, and list item (a) suggests battery life diminishes with increased entropy accumulation.

Finally, since I_t increases with state of charge, see Ménard, Fontès, and Astier (2010), a higher state of charge increases the first two terms of Equation 3.35b, which elevates the degradation rate, consistent with row four of Table 3.1.

3.8   CONCLUSION

A method to construct models of degradation and ageing of any kind was presented. This method was based on the thermodynamic degradation paradigm and the DEG theorem, which in turn are based on the fundamentals of thermodynamics. The method selects a suitable degradation measure, and then relates the rate of the degradation measure to the rate of irreversible entropy generated by the dissipative processes that underlie the degradation. This is encapsulated in Equation 3.4, wherein the rate of degradation is expressed as a linear combination of the rate of generation of irreversible entropy. The method was demonstrated on three very different examples: sliding wear, fatigue, and battery ageing. The model constructed from the DEG theorem was quantitatively consistent with past degradation models for wear and fatigue, and the expression derived for battery ageing was qualitatively consistent with physical observations. This method permits a structured approach toward degradation modeling, inasmuch as the rate of entropy generation can be related to the mechanics of the dissipative processes that underlie the entropy generation. Often this is the power dissipated divided by temperature. Using this method, equations that govern degradation and ageing can be derived and related to the operational physics of the system.

REFERENCES

Amiri, M., and M.M. Khonsari. 2012. "On the Role of Entropy Generation in Processes Involving Fatigue." *Entropy* 14, no. 12, pp. 24–31. doi: 10.3390/e14010024

Amiri, M., M. Naderi, and M.M. Khonsari. 2011. "An Experimental Approach to Evaluate the Critical Damage." *International Journal of Damage Mechanics* 20, no. 1, pp. 89–112, doi: http://dx.doi.org/10.1177/1056789509343082

Archard, J.F. 1953. "Contact and Rubbing of Flat Surfaces." *Journal of Applied Physics* 24, no. 8, pp. 981–88. doi: http://dx.doi.org/10.1063/1.1721448

Bejan, A. 1988. *Advanced Engineering Thermodynamics*. New York: Wiley.

Broussely, M., P. Biensan, F. Bonhomme, P. Blanchard, S. Herreyre, K. Nechev, and K. Staniewicz. 2005. "Main Aging Mechanisms in Li ion

Batteries." *Journal of Power Sources* 146, no. 1–2, pp. 90–96. doi: http://dx.doi.org/10.1016/j.jpowsour.2005.03.172

Brown, F.T. August 2006. *Engineering System Dynamics, A Unified Graph-Centered Approach*. 2nd ed. CRC Press.

Bryant, M.D. 2010. "Unification of Friction and Wear." In *Recent Developments in Wear Prevention, Friction and Lubrication*, ed. G.K. Nikas, pp. 159–96. Austin, TX: Transworld Research Network (Research Signpost). ISBN 978-81-308-0377-7.

Bryant, M.D. and M.M. Khonsari. 2008. "Application of Degradation-Entropy Generation Theorem to Dry Sliding Friction and Wear." *Proceedings of STLE/ASME International Joint Tribology Conference*. Miami, FL. IJTC2008-71079.

Bryant, M.D., M.M. Khonsari, and F.F. Ling. 2008. "On the Thermodynamics of Degradation." *Proceedings of the Royal Society London Series A* 464, no. 2096, pp. 2001–14. doi: http://dx.doi.org/10.1098/rspa.2007.0371

Dai, Z., S. Yang, and Q. Xue. 2000. "Thermodynamic Model of Fretting Wear." *Journal of Nanjing University of Aeronautics & Astronautics (China)* 32, no. 2, pp. 125–31.

Duyi, Y. and W. Zhenlin. 2001. "A New Approach to Low Cycle Fatigue Damage Based on Exhaustion of Static Toughness and Dissipation of Cyclic Plastic Strain Energy during Fatigue." *International Journal of Fatigue* 23, no. 8, pp. 679–87. doi: http://dx.doi.org/10.1016/s0142-1123(01)00027-5

Electropaedia. 2005. *Battery and Energy Technologies: Battery Life (and Death)*. http://www.mpoweruk.com/life.htm

Erdinc, O., B. Vural, and M. Uzunoglu. June 9–11, 2009. "A Dynamic Lithium-Ion Battery Model Considering the Effects of Temperature and Capacity Fading." *2009 International Conference on Clean Electrical Power* pp. 383–6. doi: http://dx.doi.org/10.1109/ICCEP.2009.5212025

Fouvry, S. and P. Kapsa. 2001. "An Energy Description of Hard Coating Wear Mechanisms." *Surface and Coating Technology* 138, no. 2–3, pp. 141–8. doi: http://dx.doi.org/10.1016/s0257-8972(00)01161-0

Fouvry, S., P. Kapsa, H. Zahouani, and L. Vincent. 1997. "Wear analysis in fretting of hard coatings through a dissipated energy concept." *Wear* 203–204, pp. 393–403. doi: http://dx.doi.org/10.1016/s0043-1648(96)07436-4

Fouvry, S., T. Liskiewicz, P. Kapsa, S. Hannel, and E. Sauger. 2003. "An Energy Description of Wear Mechanisms and Its Applications to Oscillating Sliding Contacts." *Wear* 255, no. 1–6, pp. 287–98. doi: http://dx.doi.org/10.1016/s0043-1648(03)00117-0

Frederick, D. and T.S. Chang. 1965. *Continuum Mechanics*. Boston, MA: Allyn and Bacon.

Huq, M.Z. and J.-P. Celis. 2002. "Expressing Wear Rate in Sliding Contacts Based on Dissipated Energy." *Wear* 252, no. 5–6, pp. 375–83. doi: http://dx.doi.org/10.1016/s0043-1648(01)00867-5

Juvinall, R.C. and K.M. Marshek. 1991. *Fundamentals of Machine Component Design*. 2nd ed. New York, NY: Wiley.

Karnopp, D.C., D.L. Margolis, and R.C. Rosenberg. 2000. *System Dynamics, a Unified Approach*. 3rd ed. New York: Wiley.

Klamecki, B.E. 1984. "Wear— An Entropy Based Model of Plastic Deformation Energy Dissipation in Sliding." *Journal of Wear* 96, no. 3, pp. 319–29.

Kondepudi, D., and I. Prigogine. 1998. *Modern Thermodynamics from Heat Engines to Dissipative Structures*. New York, NY: Wiley.

Lide, D.R, ed. 2006–2007. *CRC Handbook of Chemistry and Physics*. 87th ed. USA: CRC Press.

Ling, F.F., Bryant, M.D., and Doelling, K.L. December 2002 December 2002. "On Irreversible Thermodynamics for Wear Prediction". *Wear* 253, no. 11–12, pp. 1165–72.

Maugis, D. 1999. *Contact, Adhesion and Rupture of Elastic Solids*. Berlin, Germany: Springer.

Ménard, L., G. Fontès, and S. Astier. 2010. "Dynamic Energy Model of a Lithium-Ion Battery." *Mathematics and Computers in Simulation* 81, no. 2, pp. 327–39. doi: http://dx.doi.org/10.1016/j.matcom.2010.07.026

Miner, M.A. 1945. "Cumulative Damage in Fatigue." *Transactions of ASME* 12, 159–65.

Naderi, M. and M.M. Khonsari. 2010. "A Thermodynamic Approach to Fatigue Damage Accumulation under Variable Loading." *Materials Science and Engineering: A* 527, no. 23, pp. 6133–39. doi: http://dx.doi.org/10.1016/j.msea.2010.05.018

Naderi, M. and M.M. Khonsari. 2011. "On the Thermodynamic Entropy of Fatigue Fracture." *Structural Health Monitoring* 10, no. 2, pp. 189–97. doi: http://dx.doi.org/10.1177/1475921710373295

Naderi, M. and M.M. Khonsari. 2012. "A Comprehensive Fatigue Failure Criterion Based on Thermodynamic Approach." *Journal of Composite Materials* 46, no. 4, pp. 437–47. doi: http://dx.doi.org/10.1177/0021998311419540

Naderi, M., M. Amiri, and M.M. Khonsari. 2010. "On the Thermodynamic Entropy of Fatigue Fracture." *Proceedings of Royal Society of London Series A* 466, no. 2114, pp. 423–38. doi: http://dx.doi.org/10.1098/rspa.2009.0348

Rabinowicz, E. 1965. *Friction and Wear of Materials*. New York, NY: Wiley & Sons Inc.

Rabinowicz, E. 1980. *Wear Control Handbook*. 486. New York, NY: ASME press.

Rice, J.R. 1978. "Thermodynamics of the Quasi-Static Growth of Griffith Cracks." *Journal of Mechanics and Physics of Solids* 26, no. 2, pp. 61–78. doi: http://dx.doi.org/10.1016/0022-5096(78)90014-5

Sosnovskiy, L.A. and S.S. Sherbakov. 2009. *Surprises of Tribo-Fatigue*. Minsk Belarus: Magic Book. ISBN 978-985-6863-24-3.

Vetter, J., P. Novak, M.R. Wagner, C. Veit, K.C. Moller, J.O. Besenhard, M. Winter, M. Wohlfahrt-Mehrens, C. Vogler, and A. Hammouche. 2005. "Ageing Mechanisms in Lithium-ion Batteries." *Journal of Power Sources* 147, no. 1, pp. 269–81. doi: http://dx.doi.org/10.1016/j.jpowsour.2005.01.006

Whaley, P.W. November 9–15, 2012. "Critical Entropy Threshold: An Irreversible Thermodynamic Theory of Fatigue." *Proceedings of the ASME 2012 International Mechanical Engineering Congress and Exposition IMECE2012.* Houston, Texas.

CHAPTER 4

THERMODYNAMIC DAMAGE WITHIN PHYSICS OF DEGRADATION

Alec Feinberg

SECTION 1: EQUILIBRIUM THERMODYNAMIC DAMAGE ASSESSMENT

4.1 INTRODUCTION

There are a lot of interesting approaches for understanding the physics of device degradation. Reviewing the literature, one might note that thermodynamics seems underutilized for this area. You may wonder why we need another approach. The answer is in many cases you do not. However, sometimes, systems are complex and made up of many components. How do we describe the ageing of a complex system? Here is possibly where thermodynamics can be invaluable, since it is mainly an energy approach. In addition, we will see that assessing thermodynamic damage can be very helpful in quantifying the life of different devices and to understanding their failure mechanisms.

More importantly, thermodynamics is a natural candidate to use for understanding device ageing. For example, although most people who study thermodynamics are familiar with its second law, not many think of it as a good explanation of why a device degrades over time. We can manipulate a phrasing of the second law to clarify our point, that is:

Second law in terms of device thermodynamic damage: *The sponta-neous irreversible damage processes that take place in a device interacting with its environment, do so in order to go toward thermodynamic equilib-rium with its environment.*

There are many phrasings of the second law. This phrasing describes ageing, and we use it in this chapter as the second law of thermodynamics damage occurring in devices as they age. We provide some examples of this statement in regard to ageing to help clarify this.

When we state that degradation is irreversible, we mean either non-repairable damage or that we cannot reverse the degradation with-out at the same time employing some new energetic process to do so. *We see there is a strong parallel consequence of the second law associated with spontaneous degradation processes. The science presents us with a gift, for its second law actually explains the ageing processes.* Thus, we are compelled to look toward this science to help us in our study of device degradation. Currently, the field of reliability physics includes a lot of thermodynamic-type explanations. Yet, realistically, the application of thermodynamics to the field of device degradation is in many ways not fully mature. Its first and second laws can be difficult to apply to complex ageing problems. Yet, we anticipate that a thermodynamic approach to complex ageing may be a very useful tool.

When we talk about device damage, we should not lose sight of the fact that that we are using it as an applicable science for device reliability. To this end, we would like to keep our sights on this goal. Thermodynamic reliability is a term that can apply to degradation physics of a device (Fein-berg and Crow 2001; Feinberg and Widom 2000) after it is taken out of the box and subjected to its use stress environmental conditions.

In this chapter, we will start by introducing some of the parallels of thermodynamics that can help in our understanding of physics-of-degradation problems. Here, fundamental concepts will be introduced to build a basic framework in the hope that the science of thermodynamic damage in reliability physics will eventually mature.

When building a semiconductor component, manufacturing a steel beam, or simply inflating a bicycle tire, a system is created that interacts with its environment. Left to itself, the interaction between the system and environment degrade the system of interest in accordance with our second law phrasing of device degradation. That is, the degradation is driven by this tendency of the system/device to come into thermodynamic equilibrium with its environment. The total order of the system plus its environment tends to decrease. The air in the bicycle tire will start to dif-fuse through the rubber wall; impurities from the environment will diffuse into otherwise more pure semiconductors; internal manufacturing stresses

will cause dislocations to move into the semiconductor material; iron alloy steel beams will start to corrode as oxygen atoms from the atmospheric environment diffuse into the steel. In all of these cases, the spontaneous processes creating disorder are irreversible. For example, the air is not expected to go back into the bicycle tire; the semiconductor will not spontaneously purify; and the steel beam will only build up more and more rust. The original order created in a manufactured product diminishes in a random manner, and becomes measurable in our macroscopic world.

Associated with the increase in total disorder or entropy is a loss of ability to do useful work. The total energy has not been lost but degraded. The total energy of the system plus the environment is conserved during the process when total thermodynamic equilibrium is approached. The entropy of the ageing process is associated with that portion of matter that has become disorganized and is affecting our device's ability to do useful work. For the bicycle tire example, prior to ageing, the system energy was in a highly organized state. After ageing, the energy of the gas molecules (which were inside the bicycle tire) is now randomly distributed in the environment. These molecules cannot easily perform organized work; the steel beam, when corroded into rust, has lost its strength. These typical second-law examples describe the irreversible processes that cause ageing.

More precisely, if the entropy has not increased, then as stated earlier, the device/system has not aged. Sometimes, it will be helpful to separately talk about entropy in two categories: (a) "entropy damage" as compared with (b) "general entropy change" which does not cause device damage. For example, the bicycle tire that has degraded due to energy loss did not experience damage and can be reused. However, the corrosion of the steel beam is permanent damage. In some cases it will be obvious, in other cases we may need to keep tabs on entropy damage. In most cases, we will mainly be looking at entropy change due to device ageing as compared with absolute values of entropy, since entropy change is easier to measure. Entropy in general is not an easy term to understand, it is like energy, the more we learn how to measure it, the easier it becomes to understand.

4.2   THE SYSTEM (DEVICE) AND ITS ENVIRONMENT

In thermodynamics, we see that it is important to define both the device and its neighboring environment. Traditionally, this is done quite a bit in thermodynamics. Note that we use the term system. Here this term applies to some sort of device, complex subsystem, or even a full system made up of many devices. The actual term system or controlled mass is often used

in many thermodynamic text books. In terms of the ageing framework, we will mean:

- *The **system** is some sort of device set apart for study. From an engineering point of view, of concern is the possible ageing that can occur to it.*
- *The **environment** is the neighboring matter, that interacts with the system in such a way that it drives it toward its thermodynamic equilibrium ageing state.*

   This interaction between a device and its environment drives the device toward a thermodynamic equilibrium ageing state. It is important to realize that there is no set rule on how the system or the environment is selected. The key is that the final results be consistent.

4.2.1   THERMODYNAMIC PROCESSES ARE IRREVERSIBLE AND CAUSE DAMAGE

Thermodynamic processes are commonly referred to as reversible or irreversible. Sanding a piece of wood is an irreversible process that causes damage. We create heat from friction, which raises the internal energy to the surface; some of the wood is removed, creating higher-disordered wood particles so that the entropy increases. The disorder wood particles can be thought of as entropy damage; the wood block undergoes an increase in its internal energy from heating, which also increases its entropy as well; some of the wood at the surface is loose but does not break free. Thus, not all the entropy production goes into damage (removal of wood). Since we cannot do a reversible cycle of sanding that removes the wood particles and then puts it back to its original state, the process is irreversible and damage has occurred. Although this is a gross example, in a sense there are no reversible real processes. This is because work is always associated with an energy loss. The degree of this loss can be minimized in many cases for a quasi-static process (slow varying in time). Then we are closer to a reversible or a less irreversible process. For example, current flowing through a transistor will cause the component to heat up and emit electromagnetic radiation, which cannot be recovered. Moreover, commonly associated with the energy loss is degradation to the transistor; this is a consequence of the environment doing work on the transistor. In some cases, we could have a device doing work on the environment (like a battery). There are a number of ways to improve the irreversibility of

the ageing transistor. Improve the design's reliability so that less heat is generated, or we can lower the environmental stress such as decrease the power applied to the transistor. In the limit of reducing the stress to zero, we approach a reversible process. Therefore, reversible processes must be quasi-static. However, this does not imply that all quasi-static processes are reversible. In addition, the system may be repairable to its original state from a reliability point of view.

A **quasi-static process** *ensures that the system will go through a gentle sequence of states such that a number of important thermodynamic parameters are well-defined functions of time, if infinitesimally close to equilibrium (so the system remains in quasi-static equilibrium), the process is typically reversible.*

A repairable system is in a sense "repairable-reversible" or less irreversible from an ageing point of view. However, we cannot change the fact that the entropy of the universe has permanently increased from the original failure and that a new manufactured part had to be made for the replaceable part. Such entropy increase has likely caused damage to the environment that we live in.

4.3 THERMODYNAMIC WORK AND THE FIRST LAW

As a system ages, work is performed by the system on the environment or vice versa. The non-equilibrium process involves an energy exchange between these two. Measuring the work isothermally (constant temperature) performed by the system on the environment, and if the effect on the system could be quantified, then a measure of the change in the system's free energy can be obtained. If the process is quasi-static, then generally the energy in the system ΔU can be decomposed into the work ΔW done by the environment on the system and the heat ΔQ flow.

The bending of a paper clip back and forth illustrates cyclic work done by the environment on the system that often causes dislocations to form in the material. The dislocations cause metal fatigue, and thereby the eventual fracture in the paper clip; the diffusion of contaminants from the environment into the system may represent chemical work done by the environment on the system. We can quantify such changes using the first and second laws of thermodynamics. The first law is a statement that energy is conserved if one regards heat as a form of energy.

The first law of thermodynamics: *The energy change of the system ΔU is partly due to the work ΔW performed on the system by the*

environment and partly due to the heat ΔQ that flows from the environment to the system

$$\Delta U = \Delta Q + \Delta W \qquad (4.1)$$

In the case where heat and work are added to the system, then either one or both can cause damage. If we could track this, we would find that the portion of related damage is causing loss in the system's free energy (which is discussed in Section 4.7).

If heat flows from the system to the environment, then our sign convention is that $\Delta Q < 0$. Similarly, if the work is done by the system on the environment then our sign convention is that $\Delta W < 0$. That is, adding ΔQ or ΔW to the system is positive, increasing the internal energy. (Note that the first law does not prohibit a degraded system from spontaneous repair, which is a consequence of the second law.)

During the quasi-static process, the work done on the system by the environment has the form

$$\delta W = \sum_a Y_a dX_a \qquad (4.2)$$

Each generalized displacement dX_a is accompanied by a generalized conjugate force Y_a. Note that because work is a function of how it is performed (often termed path dependent in thermodynamics), we used the symbol δW instead of dW, indicating this for an infinitesimal increment of work. For a simple system, there is but one displacement X accompanied by one conjugate force Y. Some examples of basic conjugate work variables are given in Table 4.1 (Feinberg and Crow 2001; Feinberg and Widom 2000).

A system's state is often defined by macroscopic state variables. Thermodynamic state variables are needed to define the equilibrium state of the system. Common examples of state variables are temperature, volume, pressure, energy, entropy, number of particles, mass, and chemical composition. We see that some pairs of state variables are directly related to mechanical work in the table above. These macroscopic parameters depend on the particular system under study and can include voltage, current, electric field, vibration displacements, and so forth. Thermodynamic parameters can be categorized as intensive or extensive. Intensive variables have uniform values throughout the system such as pressure or temperature. Extensive variables are additive such as volume or mass. For example, if the system is sectioned into two subsystems, the total volume V is equal to the sum of the volumes of the two subsystems. The pressure is intensive. The intensive pressures of the subsystems are equal and the

Table 4.1. Generalized conjugate mechanical work variables

Common systems δW	Generalized force Y	Generalized displacement X	Mechanical work $\delta W = YdX$
Gas	Pressure (−P)	Volume (V)	−P dV
Chemical	Chemical potential (μ)	Molar number of atoms or molecules (N)	μ dN
Spring	Force (f)	Distance (x)	f dx
Mechanical wire/bar	Tension (J)	Length (L)	J dL
Mechanical strain	Stress (σ)	Strain (e)	σ de
Electric polarization	Polarization (−p)	Electric field (E)	−p dE
Capacitance	Voltage (V)	Charge (q)	V dq
Induction	Current (I)	Magnetic flux (Φ)	IdΦ
Magnetic polarizability	Magnetic intensity (H)	Magnetization (M)	H dM
Linear system	Velocity (v)	Momentum (m)	v dm
Rotating fluids	Angular velocity (ω)	Angular momentum (L)	ωdL
Resistor	Voltage (V)	Current (I)	$\int VI\, dt$

same as before the division. Intensive parameters can be defined in the small neighborhood of a point.

4.4 THERMODYNAMIC SECOND LAW IN TERMS OF DEVICE ENTROPY DAMAGE

We have stated that as a device ages, measurable disorder (degradation) occurs. We mentioned that the quantity of entropy defines the property of matter that measures the degree of microscopic disorder that appears at the macroscopic level. Therefore, we can also restate our phrasing for device ageing of the second law in terms of entropy when the system–environment interaction is isolated:

  The second law of thermodynamics in entropy terms of device degradation: *The spontaneous irreversible damage process that takes place*

in the system–environment interaction when left to itself increases the total entropy that results from a tendency *of the system to go toward thermodynamic equilibrium with its environment.*

This is more evident (measurable) if the system is immersed in an isothermal (constant temperature) environment. There is no way to reverse this process without creating more entropy. This alternate phrasing of the second law is another way of saying that the total order in the device plus the environment changes toward disorder. Degradation is a natural process that starts in one equilibrium state and ends in another; it will go in the direction that causes the entropy of the system plus the environment to increase for an irreversible process and to remain constant for a reversible process.

An ageing process can always be associated with entropy increase where:

The entropy generated associated with device damage, we can term, **"entropy damage."**

Entropy is an extensive property, so that the total entropy between the environment and the device is the sum of the entropies of each. Therefore, the device and its local environment can be isolated to help explain the entropy change. We can write that the entropy generated S_{gen} in an ageing process as

$$S_{gen} = \Delta S_{total} = \Delta S_{device} + \Delta S_{env} \geq 0 \tag{4.3}$$

Now in a degradation process, the device and the environment can both have the entropies changed. For example, matter that has become disorganized, such as a phase change, that affects device performance. In theory, "damage entropy" is separable in the ageing process related to the device (system) such that

$$\Delta S_{device} = \Delta S_{damage} + \Delta S_{non\text{-}damage} \geq 0 \tag{4.4}$$

So that, by this definition, *damage entropy* change, ΔS_{damage}, must be greater than zero or ageing in the device is not measurable. And nondamage entropy increase, $\Delta S_{non\text{-}damage}$, is when more disorganization occurring in the device that is not currently affecting the device performance; for example, vibration in the lattice of a crystal might increases when heat is added but there is no permanent disorder to that area of the device that is affecting its performance. However, we might still be able to measure that portion of the heat added that did not cause device damage.

Entropy damage in a system: While non damage entropy can be added or removed from a system without causing measurable degradation, such as by adding or removing heat in the system, damage system entropy can only be increased in a measurable way to its maximum value where the system approaches failure.

4.4.1 ENTROPY AND FREE ENERGY

Prior to ageing, our device has a certain portion of its energy that is "available" to do useful work. This is called the thermodynamic *free energy*. The thermodynamic free energy is the internal energy input of a system minus the energy that cannot be used to create work. This unusable energy is the entropy (S) multiplied by the temperature (T) of the system. There are two thermodynamic free energies widely used. They are Helmoltz free energy (F), which is the capacity to do mainly mechanical (useful) work, while the Gibbs' energy is primarily for non-mechanical work in Chemistry. The system's free (or available) energy is in practice less than the system's energy U; that is, if T denotes the temperature of the environment and if S denotes the system entropy, then Helmholtz free energy is $F = U - TS$, which obeys $F < U$. The free energy is then the internal energy of a system minus the amount of unusable (or useless) energy that cannot be used to perform work. The unusable energy is given by the entropy of a system multiplied by the temperature of the system. If the system's initial free energy is denoted by F_i (before ageing) and the final free energy is denoted by F_f (after ageing), then $F_f < F_i$. The system is in thermal equilibrium with the environment when the free energy is minimized. We can also phrase the second law for ageing in terms of the free energy

The second law of thermodynamics in free energy in terms of device degradation: The spontaneous irreversible damage process that takes place over time in a system decreases the free energy of the system toward a minimum value. This spontaneous process reduces the ability of the system to perform useful work on the environment, which results in ageing of the system.

For an environment at a fixed temperature, the isothermal change in the system's free energy is equal to the work done by the system on the environment;

$$Work = (F_f - F_i) \tag{4.5}$$

(This is for a closed system, indicating no exchange of matter).

4.4.2 WORK, RESISTANCE, AND GENERATED ENTROPY

Work is done against some sort of resistance. Resistance is a type of friction that increases temperature, creates heat and entropy. In the absence of friction, the change in entropy is zero. We think of mechanical friction as the most common example, but we can include electrical resistance, air resistance, resistance to a chemical reaction, resistance of heat transfer, and so on. Real work processes always have friction generating heat. If heat flows into a system, then the system's entropy increases. Heat flowing out decreases the system's entropy. Some of this heat that flows into a system can cause entropy damage—permanent degradation to the system.

Typically, instead of talking about some form of "absolute entropy," measurements are generally made on the change in entropy that takes place in a specific thermodynamic process. For example, in an isothermal process, the change in entropy (ΔS) is the change in heat (δQ) divided by the absolute temperature (T):

$$dS \geq \frac{\delta Q}{T} \tag{4.6}$$

Here, the equality holds for reversible processes, and the inequality for irreversible ones (as defined in the next section). Note that because the entropy measurement depends on our observation of heat flow, which is a function of the heat dissipation process (often termed path dependent in thermodynamics), we again use the symbol δQ instead of dQ. The inequality in the above equation reminds us that the entropy change of a closed system during an irreversible process is always greater than the entropy transfer. That is, some entropy is generated during an irreversible process (S_{gen}), and this generation is due to the presence of the irreversibilities. So, instead of the inequality, we can write this as

$$\Delta S_{Sys} = S_2 - S_1 = \int_1^2 \frac{\delta Q}{T} + S_{Gen} \tag{4.7}$$

4.4.3 ENTROPY MAXIMIZE PRINCIPLE—COMBINED FIRST AND SECOND LAWS

Typically, entropy change is measured, not absolute values of entropy. We may combine the general first law statement

$$dU = \delta Q + \delta W = \delta Q + \sum_a Y_a dX_a \tag{4.8}$$

with the statement of the thermodynamic second law $\delta Q = TdS$ (for reversible processes). The full statement of the two thermodynamic laws involves *only exact differential forms*. For a general system, the combined statement reads

$$U = U(S, X_1, \cdots, X_n), \tag{4.9}$$

$$dU = TdS + \sum_a Y_a dX_a \tag{4.10}$$

The internal energy U is a function of the entropy and the generalized displacements. During a quasi-static process, the energy change may be decomposed into a term representing the heat flow TdS from the environment into the system plus the work $\sum_a Y_a dX_a$ done by the environment on the system.

In terms of the entropy function, we have

$$S = S(U, X_1, \cdots, X_n), \tag{4.11}$$

$$dS = \left(\frac{1}{T}\right) dU - \sum_a \left(\frac{Y_a}{T}\right) dX_a \tag{4.12}$$

The entropy maximum principle in terms of ageing: *Given a set of physical constraints, a system will be in equilibrium with the environment if and only if the total entropy of the system and environment is at a maximum value.*

$$S_{tot} = S_{Sys} + S_{en} = S_{Max} \tag{4.13}$$

At this point, the system has lost all its useful work, maximum damage has occurred, and further ageing no longer will occur.

Example 4.1 Thermal Equilibrium

Consider a system that can exchange energy with its environment. The system can exchange energy with the environment subject to the constraint of overall energy conservation. The total conserved energy is written

$$U_{tot} = U + U_{en} \tag{4.14}$$

where we have dropped the "Sys" subscript. Under a quasi-static exchange of energy we have

$$dU_{tot} = dU + dU_{en} = 0, \tag{4.15}$$

$$dS_{tot} = \frac{dU}{T} + \frac{dU_{en}}{T_{en}} \tag{4.16}$$

Since $dU = -dU_{en}$

$$dS_{tot} = \left(\frac{1}{T} - \frac{1}{T_{en}} \right) dU, \tag{4.17}$$

$$dS_{tot} \geq 0 \tag{4.18}$$

In order to insure that the total entropy goes to a maximum, we have positive increments under the exchange of energy in which the system energy change is dU. If the environmental temperature is more than the system temperature $T_{en} > T$, then this dictates that $dU > 0$. If the environmental temperature is less than the system temperature $T_{en} < T$, then this dictates that $dU < 0$. The energy then flows from the higher temperature region to the lower temperature region. If the entropy is at the maximum value, then the first-order differential vanishes:

$$dS_{tot} = \left(\frac{1}{T} - \frac{1}{T_{en}} \right) dU = 0. \tag{4.19}$$

At equilibrium, the environmental temperature is equal to the system temperature $T = T_{en}$, and the energy exchange flow comes to a halt. This is in accordance with the entropy maximum principle, which occurs when the system and the environment are in equilibrium. In terms of ageing, we might have a transistor that is being cooled by a cold reservoir. If the transistor fails, it comes to equilibrium with the environment so that the temperature of the system becomes equal with the environment.

Example 4.2   Equilibrium and Charge Exchange

Shown in Figure 4.1 is a system (capacitor) in contact with an environment (battery).

A (possibly) nonlinear capacitor is connected to a battery. The capacitor and the battery can exchange charge and energy subject to conservation of total energy and total charge. Then, with these constraints, we wish to maximize
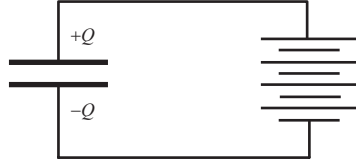
Figure 4.1. System (capacitor) and environment (battery) circuit.

$$S_{tot} = S(U,q) + S_{en}(U_{wn},q_{en}) \tag{4.20}$$

with the total energy and charge

$$U_{tot} = U + U_{en}, \tag{4.21}$$

$$q_{tot} = q + q_{en} \tag{4.22}$$

obeying

$$dU_{tot} = dU + dU_{en} = 0, \tag{4.23}$$

$$dq_{tot} = dq + dq_{en} = 0. \tag{4.24}$$

The entropy for the capacitor and battery requires

$$dS_{tot} = dS + dS_{en}, \tag{4.25}$$

Then we can write from the combined first and second law equations

$$dS_{tot} = \left(\frac{dU}{T} - \frac{Vdq}{T}\right) + \left(\frac{dU_{en}}{T_{en}} - \frac{Edq_{en}}{T_{en}}\right), \tag{4.26}$$

or equivalently

$$dS_{tot} = \left(\frac{1}{T} - \frac{1}{T_{en}}\right)dU - \left(\frac{V}{T} - \frac{E}{T_{en}}\right)dq, \tag{4.27}$$

$$dS_{tot} \geq 0. \tag{4.28}$$

Therefore, for equilibrium, when the entropy is maximum $T = T_{en}$ and $V = E$, so that $dS_{tot} = 0$ the battery has degraded so that no more useful work

can be done. Of course, we can have a quasi-static equilibrium condition when the capacitor is fully charged and negligible current is flowing.

Example 4.3   Available Work

Consider a system in a state with initial energy E and entropy S that is not in equilibrium with its neighboring environment. We wish to find the maximum amount of useful work that can be obtained from the system if it can react in any possible way with the environment. Let's say the system wants to expand to come to equilibrium with the environment and we have a shaft on it in one direction that is doing useful work. From the first law we have

$$\delta U = - \delta W - P_o dV - \delta Q \qquad (4.29)$$

The entropy production is

$$\delta S_{Gen} = dS - \frac{\delta Q}{T} \qquad (4.30)$$

Here, dS is the increase in entropy storage and $\delta Q/T$ is the entropy outflow. Combining these two equations, then the useful shaft work is

$$\delta W = - \delta U - P_o dV - T_o (\delta S_{Gen} - dS) \qquad (4.31)$$

Integrating from the initial to final state we have for the actual work

$$W_{if} = (U + P_o V - T_o S)_I - (U + P_o V - T_o S)_f - T_o S_{Gen} \qquad (4.32)$$

The irreversible work is

$$W_{irr} = T_o S_{Gen} \qquad (4.33)$$

The reversible work is

$$W_r = (U + P_o V - T_o S)_I - (U + P_o V - T_o S)_f \qquad (4.34)$$

The S_{Gen} is positive from Equation 4.32 or zero so that

$$W_r = (\Delta U + P_o \Delta V - T_o \Delta S) \qquad (4.35)$$

is the available work (free energy). Thermodynamics sometimes refers to a term called *exergy* (or availability) as the maximum available work when a state of the system is reduced to full equilibrium with the environment (also called the dead state). In reliability terms, this could represents a catastrophic failed state. To see this we can find the minimum at fixed T_o and P_o

$$\left(\frac{\partial W_r}{\partial U} \right)_V = 0, \quad \left(\frac{\partial W_r}{\partial V} \right)_U = 0 \quad (4.36)$$

which gives

$$1 - T \left(\frac{\partial S}{\partial U} \right)_V = 0, \quad P_o - T_o \left(\frac{\partial S}{\partial V} \right)_U = 0 \quad (4.37)$$

These are the thermodynamic definitions of temperature and pressure. We see that these conditions are satisfied when the system has the same temperature and pressure as the environment at the final state. Thus, the largest maximum work is possible when the system ends up in equilibrium with the environment. Then, it is impossible to extract further energy as useful work.

The irreversible work is

$$W_{irr} = W_{rev} - W_{actual} \quad (4.38)$$

The efficiency and inefficiency for work-producing components are

$$\eta = \frac{W_{actual}}{W_{rev}} = \frac{W_{actual}}{W_{actual} + W_{irr}} \text{ and } 1 - \eta = \frac{W_{irr}}{W_{actual} + W_{irr}} \quad (4.39)$$

4.5 THERMODYNAMIC CATASTROPHIC AND PARAMETRIC FAILURES

We have defined the maximum entropy and minimum free energy as equilibrium states. We now use these definitions to formally define failure:

Catastrophic failure occurs when the system's free energy is as small as possible and its entropy is as large as possible, and the system is in its true final thermodynamic equilibrium state with the neighboring environment such that

$$\frac{d\varphi}{dt} = 0, \text{ and } \frac{dS_{Total}}{dt} = 0 \quad (4.40)$$

Typically, catastrophic failure occurs due to permanent device degradation caused by maximum entropy damage. We note that there are non-catastrophic intermediate equilibrium states that a system can be in with its environment. Some examples might include: when an electrical system is turned off or is in a standby state, if we have a secondary battery that has cycled in numerous charge–discharge cycles, each fatigue cycle represents an intermediate degraded state where the battery has not fully degraded to the point where it cannot perform useful work (see Example 4.7), and so on. Quasi-static measurements and quasi-intermediate states are discussed more in Section 4.7. In such cases, the degraded system has not failed as it is not in its "true final equilibrium state" relative to its stress environment. So we have to define what we mean by the equilibrium state relative to looking for a maximum in entropy.

We can also envision a situation in which a device such as a transistor or an engine degrades to a point where it can no longer perform at the intended design level. The transistor's power output may have degraded 20 percent; the engine's efficiency may have degraded 70 percent. When a parametric threshold is involved, we are likely not at a true final equilibrium state. Therefore, we note that:

Parametric failure occurs when the system's free energy and entropy have reached a critical parametric threshold value such that

$$\varphi \le \varphi_{Threshold}, \text{ and } S_{Total} \ge S_{Threshold} \tag{4.41}$$

In such cases, parametric failure is also the result of entropy damage.

4.5.1 EQUILIBRIUM AND NON-EQUILIBRIUM AGEING STATES

In this section, we have actually been describing equilibrium and non-equilibrium thermodynamics. More formally, equilibrium thermodynamics provides methods for describing the initial and final equilibrium system states without describing the details of how the system evolves to final state. Whereas non-equilibrium thermodynamics describes in more detail what happens during the evolution to the final equilibrium state, for example, the precise rate of entropy increase or free energy decrease. Those parts of the energy exchange broken up into heat and work by the first law are also tracked during the evolution to an equilibrium final state. At the point where the irreversible process virtually slows to a halt, it approaches reversibility.

Table 4.2. Thermodynamic ageing states

Entropy definitions	Measurement definition	Free energy definitions
Non-equilibrium ageing		
$\dfrac{dS_{Total}}{dt} > 0$	$\dfrac{d\Delta S_{Total}}{dt} > 0$	$\dfrac{d\varphi}{dt} < 0$
Equilibrium non-ageing state (associated with catastrophic failure)		
$\dfrac{dS_{Total}}{dt} = 0$	$\dfrac{d\Delta S_{Total}}{dt} = 0$	$\dfrac{d\varphi}{dt} = 0$
Parametric equilibrium ageing state (associated with parametric failure)		
$S_{Total} \geq S_{Threshold}$	$\Delta S_{Total} \geq \Delta S_{Threshold}$	$\varphi \leq \varphi_{Threshold}$
Relative equilibrium ageing state (associated with activated failure)		
$S_{B\,Total} \leq S_{Total} \leq$ $S_{B\,Total} + \Delta s$	$\Delta S_{B\,Total} \leq \Delta S_{Total} \leq$ $\Delta S_{B\,Total} + \Delta s$	$\varphi_B \leq \varphi \leq \varphi_B + \Delta\varphi$

For example, as work is performed by a chemical cell (a battery with an electromotive force), the cell ages and the free energy decreases. Non-equilibrium thermodynamics describes the evolution that takes place as current passes through the battery, and the final equilibrium state is achieved when the current stops and the battery is dead. "Recharging" can revive a secondary battery. However, this is cyclic work that also degrades the battery's capacity after each cycle. (See Example 4.7.)

Table 4.2 summarizes the key ageing states described in Sections 4.4 and 4.5 (Feinberg and Crow 2001; Feinberg and Widom 2000).

4.6   ENTROPY OF A COMPLEX SYSTEM

Entropy is an extensive property, thus the total entropy of a system is equal to the sum of the entropies of the parts of the system. The parts may also be sub-systems. If we isolate an area enclosing the system and its environment such that no heat, mass flows, or work flows in or out, then we can keep tabs on the total entropy. In this case the entropy generated from the isolated area is

$$S_{Gen} = \Delta S_{Total} = \sum_{i=1}^{N} \Delta S_i = \Delta S_{Sys} + \Delta S_{Surroundings} \geq 0 \qquad (4.42)$$

where the equality holds for reversible processes and the inequality for irreversible ones. ***This is an important result for thermodynamic damage***. If we can keep tabs on ΔS_{Total} over time, we can determine if ageing is occurring even in a complex system. To do this, we need a repeatable method or process to make ageing measurements at different times. If we find that the entropy has changed over time from a repeatable quasi-static measurement process, then we are able to measure and track the ageing that occurs between the systems initial, intermediate, and final states. We can call this the entropy of an ageing process. (Note that, during system ageing, we do not have to isolate the system. We only need to do this during our measurement process.)

4.7 MEASURING DAMAGE ENTROPY PROCESSES

We can theorize that any irreversible process that increases or decreases the change in entropy in a system under investigation causes some degradation to the system. However, if we cannot measure this degradation, then in our macroscopic world, the system has not actually aged. In terms of entropy generated from an initial and final state, we have

$$S_{Gen} = S_{initial} - S_{final} \geq 0 \qquad (4.43)$$

where the equal sign is for reversible process and the inequality is for irreversible one. However, what portion of the entropy generated causes degradation to the system and what portion does not? To clarify similar to Equation 4.4,

$$S_{gen} = S_{damage} + S_{non\text{-}damage} \qquad (4.44)$$

There is really no easy way to tell unless we can associate the degradation through a measurable quantity. Therefore, in thermodynamic damage, we are forced to define S_{damage} in some measurable way.

As well in thermodynamics, we typically do not measure absolute values of entropy, only entropy change. Let us devise a nearly reversible quasi-static measurement process f, and take an entropy measurement of interest at time t_1

$$\Delta S_f(t_1) = S(t_1 + \Delta t) - S(t_1) \qquad (4.45)$$

The measurement process f must be consistent to a point that it is repeatable at a much later ageing time t_2, we can observe if some measurable degradation has occurred to our device where we record the entropy change

$$\Delta S_f(t_2) = S(t_2 + \Delta t) - S(t_2), \text{ where } t_2 >> t_1 \qquad (4.46)$$

Then we can determine if damage has occurred. If our measurement process f at time t_1 and t_2 is consistent, we should find the entropy damage that has occurred between these measurement times as

$$\Delta S_{f\text{-}Damage}(t_2, t_1) = \Delta S_f(t_2) - \Delta S_f(t_1) \geq 0 \qquad (4.47)$$

where the equality occurs if no device degradation is measurable. If we do generate some damage entropy during our measurement process ($t_i + \Delta t_i$), it must be minimal compared to what is generated during the actual ageing process between time t_1 and t_2. Then, our entropy measurement difference should be a good indication of the device ageing/damage that is occurring between time t_1 and t_2. The actual ageing process to the system between time t_1 and t_2 might be a high level of stress applied to the system. Such stress need not be quasi-static. However, the stress must be limited to within reason so that we can repeat our measurement in a consistent manner at time t_2. That is, the stress should not be so harsh that it will affect the consistency of the measurement process f.

One might, for example, have a device ageing in an oven in a reliability test, then remove it, and make a quasi-static entropy measurement f at time t_1 and then put the device back in the oven and later do another measurement at time t_2. Any resulting measurement difference is damage entropy.

(Note that Equation 4.47 is a different statement from entropy flow that many books describe. Here, we are concerned with entropy damage over a repeatable measurement process.)

Example 4.4   Resistor Ageing

Resistor ageing is a fundamental example, since resistance generates entropy. A resistor with value R ohms is subjected to environmental stress over time at temperature T_1 while a current I_1 passes through it for one month. Determine a measurement process to find the entropy at two thermodynamic non-ageing states before the stress is applied and after it has

been applied at times t_{initial} and t_{final}, respectively. Determine if the resistor has aged from the measurement process and what is the final value for the resistor.

Before and after ageing at temperature T_1, we establish a quasi-static measurement process f to determine the entropy change of the resistor at an initial time t_i and final time t_f. A simple method would be to thermally insulate the resistor and pass a current through it at room temperature T_2, for a small time t and monitor the heat rise to temperature T_3. The internal energy of the resistor with work done on it is

$$U = W + Q = mCpT_3 = I^2Rt + mCpT_2 \qquad (4.48)$$

This yields

$$I^2Rt = mCp(T_3 - T_2) \qquad (4.49)$$

and

$$R(t_i) = mCp(T_3 - T_2)/I^2t \qquad (4.50)$$

The entropy for this quasi-static measurement process at initial time t_1 over the time period Δt is considered reversible so that the entropy can be written over the integral,

$$S_f(t_1 + \Delta t) - S_f(t_1) = \int \frac{\delta Q}{T} \qquad (4.51)$$

It is important to note that the entropy change is totally a function of the measurement process f, often termed path dependent in thermodynamics, which is why the symbol δQ is used in the integral. The integral for this process (d(Volume) = 0) is,

$$\Delta S_f(t_1) = \int_{T_2}^{T_3} \frac{mC_p(T)dT}{T} = mC_{p-avg}Ln\frac{T_3}{T_2} \qquad (4.52)$$

Here T_3 is the temperature rise of the resistor observed after time period Δt.

A month later, we repeat this exact quasi-static measurement process at time t_f and find that ageing occurred as the temperature observed is now T_4 where $T_4 > T_3$ and the entropy damage change observed over the final measurement time is

$$\Delta S_f(t_2) = mC_{p-avg} Ln\frac{T_4}{T_2} \tag{4.53}$$

and

$$R(t_f) = mCp(T_4 - T_2)/I^2 t \tag{4.54}$$

Therefore, the damage entropy change related to resistor degradation is

$$\Delta S_{Damage}(t_2, t_1) = mC_{p-avg} Ln\frac{T_4}{T_3} \tag{4.55}$$

or in terms of an ageing ratio

$$A_{Ageing-ratio} = Ln(T_3/T_2)/Ln(T_4/T_2) \tag{4.56}$$

The resistance change is

$$R(t_f) = \left[\frac{T_4 - T_2}{T_3 - T_2}\right] R(t_i) \tag{4.57}$$

We note

$$R = \rho \, L/A \tag{4.58}$$

Thus, an analysis of the problem indicates that the material properties and the dimensional aspects need to be optimized to reduce the observed ageing/damage that occurred.

One might now ask, so what! We can easily measure the resistance change of the ageing process with an ohm meter. On the contrary, we note that our ageing measurement was independent of the resistor itself. If we are only looking at ageing ratios, we did not even have to know any of its properties. In some cases, we are unable to make a direct measurement; we can detect ageing using thermodynamic principles. As well, the next example may be helpful to understand some advantages to this approach for complex systems.

Example 4.5   Complex Resistor Bank

In the above example we dealt with a simple resistor, but what if we had a complex resistor bank like a resistance bridge or some complex

arrangement? The system's ageing can still be detected even though we may not easily be able to make direct component measurements to see which resistor or resistors have aged over environmental stress conditions. In fact, components are often sealed, so they are not even accessible for direct measurement. In this case, according to our thermodynamic damage theory, if we isolate the system and its environment the total entropy change is from Equation 4.42

$$\Delta S_{Total} = \sum_{i=1}^{N} \Delta S_i$$

(4.59)

For a complex bank, if we are able to use the exact same measurement process f as we did for Example 4.4 then the ageing ratio is still

$$A_{Ageing-ratio} = Ln(T_3/T_2)/Ln(T_4/T_2)$$

(4.60)

4.7.1 INTERMEDIATE THERMODYNAMIC AGEING STATES AND SAMPLING

In the above examples, our ageing measurements were taken at an initial and final measurement times. These were quasi-static intermediate measurements states where little, if any, ageing occurred. We are sampling the ageing process. If it had failed, it would have been in equilibrium with its environment. We did not track how the entropy of the ageing process occurred over time. We were only concerned with whether or not degradation occurred. And our detection of the ageing process is only limited by our ability to come up with a quasi-static measurement process. In theory, we can even detect if a complex rocket ship is degrading. Each measurement process takes place at a non-ageing state at a key tracking point such as an initial, intermediate, or final measurement time. The measurement itself is taken in a small enough window of delta time to observe the state of the system but not cause any significant ageing during our observation times.

Given a system, what quasi-static measurement process f will best detect degradation at an intermediate ageing state? Part of the problem will always be resolution of our measurement process. Although we now have a tool that can possibly detect ageing of a large system, can we make a measurement with enough resolution to observe its degradation? Can we

take a partial sample in some way? For example, do we need to completely isolate the system and its environment in the entirety to make a measurement? We are now in a position that challenges our imaginations.

Thus, the principles will always be valid, but we may be limited by the practicality of the measurement. In the above case, the ageing was associated with heat. Not all ageing occurs in a manner that allows us to make degradation measurements in this way.

4.8 MEASURES FOR SYSTEM-LEVEL ENTROPY DAMAGE

We next ask what state variables can be measured as an indicator at the system level for the entropy of ageing. In this section, we will explore the state' system variables of temperature, system noise, and system failure rate.

4.8.1 MEASURING SYSTEM ENTROPY DAMAGE USING TEMPERATURE

We have noted that certainly temperature should be a good candidate for a system-level state variable. In terms of thermodynamics, if a part is incompressible (constant volume), and if a part is heating up over time t_2 compared to its initial time at t_1 due to some degrading process, then its entropy damage change is given by (e.g., see Equation 4.55)

$$\Delta S(t_2, t_1) = mC_{Avg} Ln \frac{T_2}{T_1} \tag{4.61}$$

where C is the specific heat ($Cp = Cv = C$) and for simplicity we are using an average value for C.

Then for a system made up of similar type of parts in accordance with Equation 4.61 the total damage entropy is

$$\Delta S_{Total} = \sum_{i=1}^{N} \Delta S_i = Ln \frac{T_2}{T_1} \sum_{i=1}^{N} m_i C_{Avg\,i} \tag{4.62}$$

This is an interesting result. It basically says that one measure of an ageing system is temperature. Although we might intuitively expect this, it is helpful to be specific. Note that we do not need to know the temperature of

each part; temperature is an intensive thermodynamic variable. It is a roughly uniform indicator for the system. Thus, monitoring temperature change of a large system is one key variable that is a major degradation concern for any thermal type system. Even for an ideal gas the expression is similar.

4.8.2 MEASURING SYSTEM ENTROPY DAMAGE USING NOISE

Another continuous intensive variable at the system level is noise. This may not be initially as obvious as the temperature state variable. In fact, one does not find this in most books as a thermodynamic state variable. Nevertheless, we can characterize one key state of a system using noise. In a mechanical or an electrical system, we suspect that system noise increase is a sign of disorder and increasing entropy. Simply put, if entropy damage increases, so should the system noise. For example, an electrical fan blade may become wobbly over time. The increase in its wobbliness can be thought of as noise. Not in the acoustic sense but its degree of how wobbly it is provides a measure of its increasing "noise level." Noise is a continuous random variable of some sort. The entropy of a continuous variable is treated in thermodynamics using the concept of differential entropy. For example, the statistical definitions of entropy for discrete and continuous variable X are well defined in thermodynamics as (Hazewinkel 2001; http://en.wikipedia.org/wiki/Differential_entropy).

$$Discrete\ X, p(x): \quad S(X) = -\sum p(x)\log_2 P(x) \qquad (4.63)$$

$$Continuous\ X, f(x): \quad S(X) = -\int f(x)\log_2 (f(x))dx = -E\big[\log f(x)\big] \qquad (4.64)$$

Note, in differential entropy, the variables are usually dimensionless. So if $X = voltage$, the solutions would be in terms of $X = V/Vref$, a dimensionless variable.

Here we are concerned with the continuous variable x having probability density function $f(x)$. Noise is often considered Gaussian. For example, Gaussian white noise is one common example and often reflects many real-world situations (note, not all white noise is Gaussian). When we find that a system has Gaussian white noise, the function $f(x)$ pdf is

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) \qquad (4.65)$$

Then when this function is inserted into the differential entropy equation, the results is given by (http://en.wikipedia.org/wiki/Differential_entropy)

$$S(X) = \frac{1}{2}\log(2\pi e\, \sigma(x)^2) \tag{4.66}$$

This is an important finding for system noise. We see that entropy for a Gaussian noise system, the differential entropy is only a function of its variance σ^2 (it is independent from its mean μ). Note, the results are also very similar if the noise is log normally distributed. For a system that is becoming noisier over time, the damage entropy can be measured in a number of ways where the change in the entropy at two different times t_2 and t_1 is

$$\Delta S_{Damage} = S_{t2}(X) - S_{t1}(X) = \Delta S(t2, t1) = \frac{1}{2}\log\left(\frac{\sigma_{t2}^2}{\sigma_{t1}^2}\right) \tag{4.67}$$

Interestingly enough, noise engineers are quite used to measuring noise with the variance statistic. One of the most common measurements of noise is called the Allan variance. This is a very popular way to measure noise and is in fact very similar to the Gaussian variance, given by

$$\text{Allan variance: } \sigma_y^2(\tau) = \tfrac{1}{2} <(Y_{avg2} - Y_{avg1})^2> \tag{4.68}$$

By comparison, the true variance is

$$\text{True variance: } \sigma_y^2(\tau) = \tfrac{1}{2} <(Y_{avg\,k} - <Y_{avg\,k}>)^2> \tag{4.69}$$

We note the Allan variance, commonly used to measure noise, is a continuous pair measurement of the population of noise values, whereas the true variance is nonpair measurement over the entire population. The Allan variance is used often because it is a general measure of noise and is not necessarily restricted to Gaussian-type noise.

The key results here are that entropy of ageing for system noise goes as the variance, which is also a historical way for measuring noise and is likely a good indicator of the entropy of ageing of a complex system. There are a number of historical options on how noise can be best measured. Recently, Feinberg 2015 has proposed an emerging technology for reliability detection using noise analysis for an operating system.

4.8.3 NOISE TEMPERATURE

It is not surprising that noise and temperature are related. For example, Johnson–Nyquist noise power goes directly with temperature. Electri-

cally, voltage and current are also functions of temperature. Simply put, as temperature increases, so too will device noise. Therefore, these two state variables are often related. However, each offers a unique way of characterizing ageing at the system level.

4.8.4 MEASURING SYSTEM DAMAGE ENTROPY USING FAILURE RATE

So far we have two states variable for measuring ageing. Another very obvious one that we should look at for consistency is the failure rate λ. We do not think of the failure rate as a thermodynamic state variable. In fact, to any reliability engineer, it is second nature that the failure rate is indeed a key degradation unit of measure. For consistency, we would like to put it in the context of damage entropy to see what we find. For a complex system, the most common distribution used comes from the exponential probability density function

$$f(x) = \lambda \exp(-\lambda x) \qquad (4.70)$$

The differential entropy can be found easily from the negative of the expectation of the Ln of $f(x)$ as (http://en.wikipedia.org/wiki/Differential_entropy)

$$S = -E[\text{Ln}(\lambda \exp - \lambda x)] = -E[\text{Ln}(\lambda) + \text{Ln}(\exp - \lambda x)] = -\text{Ln}(\lambda) + \lambda E[x])]$$
$$= \text{Ln}(1/\lambda) + 1 \qquad (4.71)$$

where, for the exponential distribution $E[x] = 1/\lambda$. For a system whose failure rate is constant over time (i.e., a property of the exponential distribution), the entropy is also constant. Thus, there is no change in the entropy in such a case over time. However, if the environmental stress changes, we can measure the entropy change. In this case, the entropy damage change would be for this system under two different environments

$$\Delta S = Ln(\lambda_{E1}/\lambda_{E2}) \qquad (4.72)$$

where the subscripts E1 and E2 are for environments 1 and 2.

4.8.5 ENVIRONMENTAL NOISE DUE TO POLLUTION

We have asserted that temperature, noise, and failure rate are some good key system thermodynamic state variables. However, depending on how

you define the system and the environment, they can also be state variables for the environment. An IC engine is a combustion system that interacts with the environment causing entropy damage to the environment. This pollution damage and that of other systems, on a global scale, are unfortunately measurable as meteorologist track global climate change. The temperature (global warming) effect is the most widely tracked parameter. However, it might also make sense from what we determined above to look at environmental noise degradation. As our environment ages, we might expect from our above finding that its variance will increase. Some indications of environmental variance change are larger swings in wind, rain, and temperature, causing more frequent and intense violent storms. It might prove prudent to focus on such environmental noise issues and not just track global temperature rise.

SECTION 2: NON-EQUILIBRIUM THERMODYNAMIC DAMAGE ASSESSMENT

4.9 EQUILIBRIUM VERSUS NON-EQUILIBRIUM AGEING APPROACH

We briefly discussed equilibrium versus non-equilibrium thermodynamics damage in Section 4.5.1. Per that discussion, in non-equilibrium thermodynamic degradation, we are concerned about how the ageing process takes place over time so that it can be modeled. In a sense, we already have one approach to this. In Section 1 of Chapter 4, we devised a consistent measurement process f to help measure the damage that may have occurred between times t_1 and t_2 finding the entropy damage as (see Equation 4.47)

$$\Delta S_{f\text{-}Damage}(t_2, t_1) =)\Delta S_f(t_2) - \Delta S_f(t_1) > 0 \qquad (4.73)$$

where $\Delta S_f(t_i)$ is a quasi-static measurement. Here we can add as many intermediate quasi-static thermodynamics measurements as needed to trace out the ageing path. Each measurement taken when the system is either under very little stress or in a short enough time period that we are able to sample the system's state variables over time. We then trace out how the damage is evolving over an extended time period. In this way, we will be able to model the ageing that is occurring by fitting the data. However, using this method, we need to keep track and measure a number of states throughout the ageing process. Often, we are able to model the degradation (damage) so we do not have to make many measurements. In

this section, we devise some methods helpful in non-equilibrium thermo-dynamic damage assessment.

4.9.1 CONJUGATE WORK AND FREE ENERGY APPROACH TO UNDERSTANDING NON-EQUILIBRIUM THERMODYNAMIC DEGRADATION

An approach to understand the types of thermodynamic stresses that are occurring to the ageing system is to use the conjugate work approach that we established in Section 4.3. During the quasi-static process, the work done on the system by the environment has the form

$$\delta W = \sum_a Y_a dX_a \text{ or } W = \int_{X_1}^{X_2} Y \, dX \tag{4.74}$$

where we need to sum the work. If the path is well defined, we can integrate over the path from X_1 to X_2. Each generalized displacement dX_a is accompanied by a generalized conjugate force Y_a. Recall that because work is a function of how it is performed (often termed path dependent in thermodynamics), we used the symbol δW instead of dW indicating this. For a simple system, there is but one displacement X accompanied by one conjugate force Y. Examples of basic conjugate work variables are given in Table 4.1.

Associated with work causing damage by the environment on the device is a loss of free energy, which we presented in Equation 4.5 and repeated here for convenience

$$Work = (F_f - F_i) \tag{4.75}$$

This is for an environment at a fixed temperature where the isothermal change in the system's free energy is equal to the work done by the system on the environment (and a closed system, indicating no exchange of matter).

In this section we will explore these two important equations as a basic approach to assessing non-equilibrium thermodynamic degradation.

4.10 APPLICATION TO CYCLIC WORK AND CUMULATIVE DAMAGE

In cyclic reversible work, the system undergoes a process in which the initial and final system states are identical. Earlier, a simple example was described of non-reversible cyclic work performed by bending a paper

Crystal lattice   Dislocation

1 Cycle   10 Cycles   100 Cycles
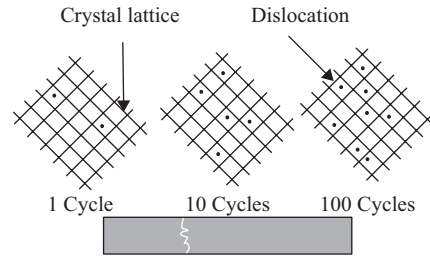
Breakage occurs when multiple dislocations line up

Figure 4.2. Conceptual view of cyclic cumulative damage.

Y-Stress   One cycle

$\Sigma = $ Area

X-Strain
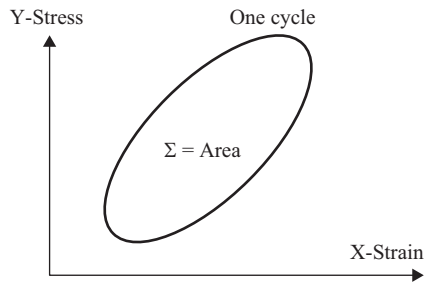
Figure 4.3. Cyclic work plane.

clip back and forth. The cyclic thermodynamic work is converted into heat and damage entropy in the system as dislocations are added at every cycle, causing plastic deformation. This produces metal fatigue and irreversible damage. This is illustrated in Figure 4.2. Ageing from such fatigue is due to external forces, which eventually result in fracture of the paper clip.

If we were trying to estimate the amount of damage done at each cycle, it might be more accurate to write the damage in terms of the number of dislocations produced during each cycle. Then we could also estimate the cumulative damage. However, this number is usually unknown. Therefore, we take a thermodynamic approach.

For a simple system, the path in the (X, Y) plane corresponding to a cycle as shown in Figure 4.3.

The work term is $\delta W_s = YdX$. Recall that the plane can represent any one of the conjugate work variable that can undergo cyclic work given in Table 4.1. In the (X, Y) plane, the cycle is represented by a closed curve C enclosing an area Σ. The closed curve is parameterized (during the time interval $t_i < t < t_f$ of the cycle as a moving point $(X(t), Y(t))$ in the plane.

For a cyclic process C in a simple system, the work done on the system by the environment is given by the integral around a closed curve

$$\text{Work(cycle)} = \oint_C Y dX = \int_{t_i}^{t_f} Y(t) \frac{dX(t)}{dt} dt \qquad (4.76)$$

The curve C representing the cycle is the boundary of an area Σ, which is written here as $C = \partial\Sigma$. Employing Stokes' integral theorem, one proves that the work done during a cycle may be related to the enclosed area via

$$\text{Work(cycle)} = \oint_{\partial\Sigma} Y dX = \iint_\Sigma dY dX \qquad (4.77)$$

In Figure 4.3, the system does work on the environment if the point in the plane transverses the curve in a clockwise fashion. The environment does work on the system if the point in the plane transverses the curve in a counterclockwise fashion.

4.10.1 THE THERMODYNAMIC DAMAGE RATIO METHOD FOR TRACKING DEGRADATION

In Example 4.3, we indicated an expression for inefficiency in Equation 4.39. We found the reversible work is the maximum useful work that can be obtained as a system undergoes a process between two specified states. The irreversibility is lost work, which is wasted work potential during a process as a result of irreversibilities. In terms of thermodynamic degradation in cyclic work, if we are creating damage at each cycle in the system, irreversible work damage cumulates until failure occurs. Some of the irreversibility in the total process is due to inefficiencies unrelated to system work damage. In theory, we should be able to track the true system work damage as it occurs, and we should distinguish the types of damage that can be measured and/or modeled

$$Damage = \frac{\sum_{i=1}^{m} w_{di}}{W_d} \quad and \quad Effective\ Damage = \frac{\sum_{i=1}^{m} w_{Ti}}{W_T} \qquad (4.78)$$

The work damage ratio: This consists of the work performed to the work needed to cause system failure. In system failure, we exhaust the maximum amount of useful system work. To consistently find this damage ratio, all work found must be taken over the same work path.

The path is important as there are many ways to walk up a hill; therefore, we will always require the numerator and denominator to travel the same path. For the damage expression, w_{di} is the cyclic work damage performed over i cycles, and W_d is the total work damage performed to cause system failure. This damage is created in the system by the environment doing work on the system or vice versa. When we cannot identify in modeling and/or in measuring the system damage part of the work performed, then we can use the work performed. Then the ratio is not the true damage, but it represents an effective amount of damage. In this case we write W_T as the total work performed to failure and w_{Ti} as the total cyclic work performed over i cycles. When we are trying to define damage, we should try and make the distinction. As a mechanical work example, damage might be related to plastic strain; we usually can measure and/or model this. At this point we will drop the subscript distinction, and just work with the concept (but one should try and keep the distinction in mind when working a specific problem). Then the cyclic damage can be written over n cycles as

$$Damage = \frac{\sum_n \oint Y_n dX_n}{W_{Failure}} \qquad (4.79)$$

When the Damage ratio value is 1, failure results for which we require the total work summed along the path traveled in the numerator to equal the full work to failure that occurred along the same path traveled in the denominator.

In many cases, it is actually possible to sum the work for each cycle in a consistent measurement process. For example, if Y is the stress and X the plastic strain, we can use strain gages and monitor the stress level and integrate out the area after each cycle using computer software. Thus, an accurate estimate can be made of the damage occurring. In other cases, where such measurements are not possible, often a technique called Miner's rule (Miner 1945) is used. This often serves as a good approximation of the effective damage and is discussed in Example 4.6.

Similarly, if the work is noncyclic with only one kind of work involved, and we are cautious to make sure the path traveled is identical in the numerator as in the denominator than the thermodynamic damage is

$$Damage = \frac{\int Y dX}{W_{Failure}} = \frac{Partial\ work}{Total\ work\ needed\ for\ failure} \qquad (4.80)$$

If there are other types of work stress causing damage, then these can be summed as long as the work path is the same. That is, we travel up the hill on the same path but we may at times carry a different amount of weight along the way. If we travel the path with one weight, we may breakdown just as we reach the top of the hill, while the next time we go up the path with more weight, we will breakdown sooner. The key is the path traveled, and then we can consider different stresses as we travel this path. Then we should be able to assess the second time up the hill if and when we will breakdown relative to the breakdown value of 1. In the case of other stresses we write

$$\text{Work} = \sum_i \int Y_i dX_i \tag{4.81}$$

This is our approach for assessing non-equilibrium thermodynamic damage. Many examples will follow to help clarify the concept.

4.10.2 DETERMINING DAMAGE ACCELERATION FACTORS USING THE DAMAGE RATIO PRINCIPLE

In nonequilibrium thermodynamics, as we seek to trace the degradation process over time, Equation 4.76 will often have the separable form

$$w = \int_{t_i}^{t_f} Y(t) \frac{dX(t)}{dt} dt = f(Y, k, E_a) t \tag{4.82}$$

Here t is the time and $f(Y, k, E_a)$ is some function of the environmental stress, k and E_a are the specific constants related to the degradation mechanism, such as a specific power exponent and activation energy. Then, according to Equation 4.79 or 4.80, if we have damage between two different environmental stresses Y_1 and Y_2, and failure occurs for each at times τ_1 and τ_2, respectively, then the damage value of 1 requires that

$$Damage = \frac{f(Y_2, k, E_a) \, \tau_2}{f(Y_1, k, E_a) \, \tau_1} = 1 \quad or \quad AF_D(1,2) = \frac{\tau_2}{\tau_1} = \frac{f(Y_1, k, E_a)}{f(Y_2, k, E_a)} \tag{4.83}$$

Here $AF_D(1, 2)$ is a damage acceleration factor that is closely related to the acceleration factor used in reliability testing as the time compression

acceleration factor between the two different stress environments 1 and 2. The difference is that the damage acceleration factor has an extra stress term as it comes from energy-related thermodynamic work. This will be clarified below in Examples 4.9 through 4.12. However, it is important to note the difference as it is helpful when distinguishing acceleration factor types and dealing with cyclic fatigue. The literature does not actually make a distinction as is done here for clarification (see, for instance, Example 4.12). Then, for any ith stress condition, the time to failure τ_i when $AF_D(1,i)$ is known is

$$\tau_i = \tau_1 AF(1,i) \text{ and } AF_D(1,i) = \frac{f(Y_1,k,E_a)}{f(Y_i,k,E_a)} \tag{4.84}$$

When AF_D is known, it allows us to write the damage at any time t between two stress environments along the same work path as

$$Damage = \frac{f(Y_1,k,E_a)\,t_1}{f(Y_2,k,E_a)\,\tau_2} = \frac{f(Y_1,k,E_a)\,t_1}{f(Y_2,k,E_a)\,AF_D(1,2)\,\tau_1} \tag{4.85}$$

The damage acceleration factor is an important tool in assessing the time evolution of thermodynamic damage occurring between different environments. Many examples are illustrated below using this concept.

Example 4.6   Miner's Rule

Equation 4.79 for cyclic damage may be used to derive Miner's empirical rule (Feinberg and Crow 2001; Feinberg and Widom 2000) that is commonly used for accumulated fatigue damage, and a number of other useful expressions in damage assessment of devices and machines.

Consider a system undergoing fatigue as in our paper clip example, which was bent back and forth a certain distance for three cycles. To find the actual work we need to sum up the cyclic work area of each bend cycle since both the stress σ and strain e will change slightly as the paper clip fatigues. If we use Stokes' theorem, it demonstrates the work is related to the cyclic area of each

$$w_1(\text{cycle}) = \sum_{i=1}^{3} \oint_{Area_i} \sigma de = \iint_{Area\,1} d\sigma de + \iint_{Area\,2} d\sigma de + \iint_{Area\,3} d\sigma de \tag{4.86}$$

In Miner's rule an approximation is actually made. Miner empirically figured that stress σ and cycles n were the main factors of damage

In our framework, this means $\text{Work}_n = \text{Work}(\sigma, n)$

Miner also empirically assumed that the work for n cycles of the same cyclic size is all that is needed—in our framework, this means

$\text{Work}_n = n\,\text{Work}(\sigma)$ (Miner's assumption versus reality as work is reduced in each σ cycle)

Using this assumption, then we average the three cyclic areas of the same stress multiplied by three in our case

$$w_i = 3\oint \sigma_i de \tag{4.87}$$

Following this approximation, for any stress level we just count the number of cycles so that the total thermodynamic work is

$$w(\text{for } k \text{ types of stress}) \approx n_1 \oint \sigma_1 de + n_2 \oint \sigma_2 de + n_3 \oint \sigma_3 de + ...$$

$$= n_1 w_1 + n_2 w_2 + n_3 w_3 + ... = \sum_{i=1}^{K} n_i w_i \tag{4.88}$$

Using this approximation, we obtain an effective damage (i.e., not the true damage) where the cumulative effective damage is

$$Effective\,Damage \approx \frac{\sum_{i=1}^{K} n_i w_i}{W_{Failure}} \tag{4.89}$$

or

$$Effective\,Damage \approx \frac{n_1 \oint \sigma_1 de + n_2 \oint \sigma_2 de + n_3 \oint \sigma_3 de + \cdots}{W_{Failure}} \tag{4.90}$$

$$= \frac{n_1 \oint \sigma_1 de}{W_{Failure}} + \frac{n_2 \oint \sigma_2 de}{W_{Failure}} + \frac{n_3 \oint \sigma_3 de}{W_{Failure}} + \cdots$$

Using the approximation above, the total work to failure is the same for each cyclic size type, along the same work path so that

$$W(for\ failire) = N_1 W_1 = N_2 W_2 = N_3 W_3 = \cdots \qquad (4.91)$$

So this yields

$$Effective\ Damage \approx \frac{n_1 \oint \sigma_1 de}{W_{Failure}} + \frac{n_2 \oint \sigma_2 de}{W_{Failure}} + \frac{n_3 \oint \sigma_3 de}{W_{Failure}} + \cdots \qquad (4.92)$$

$$= \frac{n_1 \oint \sigma_1 de}{N_1 W_1} + \frac{n_2 \oint \sigma_2 de}{N_2 W_2} + \frac{n_3 \oint \sigma_3 de}{N_3 W_3} + \cdots$$

Giving

$$Effective\ Damage \approx \frac{n_1 \oint \sigma_1 de}{N_1 \oint \sigma_1 de} + \frac{n_2 \oint \sigma_2 de}{N_2 \oint \sigma_2 de} + \frac{n_3 \oint \sigma_3 de}{N_3 \oint \sigma_3 de} + \cdots \qquad (4.93)$$

$$= \frac{n_1}{N_1} + \frac{n_2}{N_2} + \frac{n_3}{N_3} + \cdots$$

Therefore, Miner's rule (Miner 1945) is an approximation of the cumulative fatigue damage commonly written

$$Effective\ Damage \approx \frac{n_1}{N_1} + \frac{n_2}{N_2} + \frac{n_3}{N_3} + \cdots = \sum_{i=1}^{k} \frac{n_i}{N_i} \qquad (4.94)$$

4.10.3   MODIFICATION OF MINER'S RULE USING DAMAGE ACCELERATION FACTORS

It can be difficult to always know the cycles to failure at any ith stress level in Miner's application. However, since we are on the same work path, often, we can establish a time cycle acceleration factor between the stress levels. If AF_D and N_1 are known, then the Miner's rule can be simplified as per Equation 4.84 for the ith stress using

$$Ni = AF_D(1,\ i)\ N_1 \qquad (4.95)$$

Then the Miner's rule can be modified as

$$Effective\ Damage \approx \frac{n_1}{N_1} + \frac{n_2}{AF_D(1,2)N_1} + \frac{n_3}{AF_D(1,3)N_1} + \cdots$$

$$= \sum_{i=1}^{k} \frac{n_i}{AF_D(1,i)N_1} \tag{4.96}$$

4.10.4 DAMAGE IN ELECTROCHEMISTRY

In this section we provide some examples of how to use the concept of work damage in electrochemistry.

Example 4.7 Miner's Rule for Secondary Batteries

Miner's rule is not limited to mechanical stress; it can be applied to many types of cyclic fatigue situations. For example, one interesting application is to chemical cells of secondary batteries (Feinberg and Widom 2000). Here the cyclic work from Table 4.1 for k charge–discharge cycles is

$$w(\text{cycle}) = \sum_{i=1}^{k} \oint_{Area_i} V dQ \tag{4.97}$$

In an analogous manner, battery manufacturers plot something similar to S–N curves found in mechanical stress strain application. This is depth of discharge percent (DoD%) versus charging–discharging cycles to failure N. However, by comparisons, in this case, the DoD strain variable, charge is plotted using the DoD% (which is a percent of the charging capacity), instead of the stress cycles to failure by battery manufacturers for an apparent stress failure threshold of voltage as shown in the figure.

So in an application to the Miner's rule using such available data, the sum is over the DoD% ith level for battery life pertaining to a certain failure (permanent) voltage drop (such as 10 percent) of the initially rated battery voltage and then the effective damage done in n_i DoD% can be assessed when N_i is known for the ith DoD level. Then, just as before for k types of DoD% for various i-levels

$$Effective\ Damage \approx \sum_{i=1}^{k} \frac{n_i}{N_i} \tag{4.98}$$
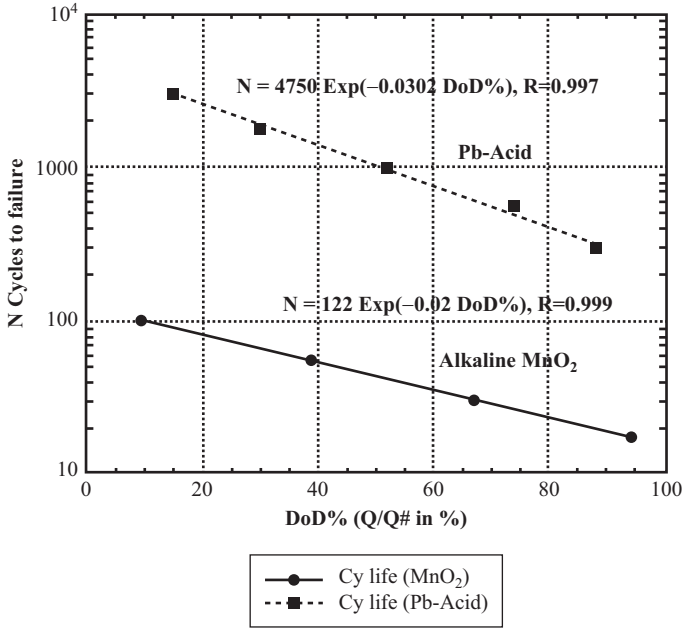
Figure 4.4. Lead acid and alkaline MnO_2 batteries fitted data (Feinberg and Widom 2000).

The S–N type of DoD%–N curve, we have modeled (Feinberg and Widom 2000) as chemical rate where activation plays a role to battery N cycle life

$$N_j = N_o e^{-\frac{\Phi_j}{K_B T}} = N_{jo} e^{-f_j DoD\%}$$ (4.99)

as is fitted in Figure 4.4. Here we assume the activation free energy is a linear function of discharge (Feinberg and Widom 2000)

$$\Phi_j = \varphi_j + \nu Q_j$$ (4.100)

where ν is the activation damage voltage and Q_j is that value for which the battery manufacturer assumes the chemical cell to be fully discharged for the jth battery type, ϕ is a damage activation energy, and then we identify

$$N_{jo} = N_o e^{-\frac{\varphi_j}{K_B T}}, \text{ and } f_j DoD\% = \frac{\nu Q_j}{K_B T}$$ (4.101)

Example 4.8 Chemical Corrosion Processes

In a corrosion process we may have N_i particles (or mole) of species i reacting. If the total reaction is consumed after N_i particles, and the reaction proceeds to the point where only $M_i < N_i$ particles have reacted, then we can immediately determine the damage from the thermodynamic work. The resulting work done on the chemically reacting system may be described in terms of the *chemical potential* μ_i of the species k. That is, under constant temperature and pressure, the work is directly related to the Gibbs free energy change so the damage from Table 4.1 and Equation 4.80 is

$$Damage = \frac{\Delta G_{M-particles}}{\Delta G_{N-particles}} = \frac{\sum_i \mu_i M_i}{\sum_i \mu_i N_i} \qquad (4.102)$$

However, chemical potentials are not easily assessed. Moreover, this says very little about the rate of the reaction. So the expression is not easily quantified. We need to work with measurable parameters that are available. For example, we may know the corrosion rate or corrosion current under a certain stress condition. Then, instead of using the chemical potential approach to assess damage, we can default to the thermodynamic work rule. For example, using the corrosion current, the work is $dW = VIdt$. Although we do not have cyclic work, we can think of it in terms of corrosion time for a single cycle. Then considering the corrosion time we write the corrosion damage

$$Damage\,(corrosion) = \sum_{i=1}^{n} \frac{(VI)_i t_i}{(VI)_i \tau_i} = \sum_{i=1}^{n} \frac{r_i t_i}{r_i \tau_i} = \sum_{i=1}^{n} \frac{t_i}{\tau_i} \qquad (4.103)$$

where for the ith stress condition (for example, temperature stress i), t_i is the corrosion time, τ_i time to fail. Note we have introduced r_i, the corrosion rate that goes as the corrosion current (see Equation 4.104). We find the expression simplifies and we do not actually need to know the corrosion current or rate at any stress level, if we know the time to failure τ_i at each ith level similar to the Miner's rule. We can view this for general, galvanic, or specific corrosion processes such as primary batteries.

Consider the case when the damage value is 1, failure occurs and we can derive a useful expression for the acceleration factor in terms of the stress. However, in this case, we will first need an expression for the corrosion rate.

A common expression for the corrosion rate in terms of mass transferred in the reaction is (Linden and Reddy 2002)

$$r = I_{Corr} \, B \, E_q / \rho \, A \tag{4.104}$$

where r is mm/year (mmpy), $I =$ corrosion current in amps, B is a units constant (3272 mm/(amp-cm-year)) of the corrosion rate, E_q is the equivalent weight (grams/equivalent), $\rho =$ density (grams/cm^3), A is the sample area (cm^3).

The corrosion current is dependent on the actual situation, but in general is thermally activated across a free energy barrier

$$I_{corr}(T) = K \, Exp(-\Delta\varepsilon/RT) \tag{4.105}$$

(R is the gas constant, more commonly used in electrochemistry than the Boltzmann's constant.) K is rate constant for generalized corrosion current (see other expression in Section 4.10.6 for the corrosion current). We can now consider two different environments at different temperatures. If we have obtained a damage value of 1 between these two stress environments, from Equations 4.103 and 4.105

$$Damage\,(corrosion) = 1 = \frac{I_1\tau_1}{I_2\tau_2}, \text{ then } AF = \frac{\tau_2}{\tau_1} = \frac{I_1}{I_2}$$
$$= \exp\left\{-\frac{\Delta\varepsilon}{R}(\frac{1}{T_1} - \frac{1}{T_2})\right\} \tag{4.106}$$

Here we canceled out the voltage potential as in different stress temperatures, only the current will likely change or dominate in the corrosion process, so

$$Damage\,(corrosion) = \frac{I_1 t_1}{I_2 \tau_2} = AF(1,2)\frac{t_1}{\tau_2} \tag{4.107}$$

The reader should note that in general the rate constant K cancels out, but depending on the way Equation 4.105 is written, and the test conditions, that may not always be the case.

An alternative expression for primary batteries is based on the Peukert's law (Linden and Reddy 2002).

$$C_p = I^Y t \tag{4.108}$$

where C_p is the battery capacity at a one-ampere discharge rate (expressed in amp-hrs), while I is the discharge current in amps, Y is the Peukert constant (typically between 1.05 and 1.15), and t is the time of discharge in hours. The acceleration factor for two different cell types would then be

$$AF = \frac{\tau_2}{\tau_1} = \frac{C_{p2}}{C_{p1}} \left(\frac{I_1}{I_2} \right)^Y \qquad (4.109)$$

4.10.5 CORROSION CURRENT IN PRIMARY BATTERIES

Aqueous corrosion is a complex ageing process. A chemical battery exemplifies this process, which must consist of four elements for corrosion to occur: a metal anode, cathode, electrolyte, and a conductive path.

The exchange of matter can be described in non-equilibrium thermodynamics in terms of the currents at the electrodes. Corrosion involves two separate processes or half-reactions: oxidation and reduction. At the anode, oxidation reaction consumes metal atoms when they corrode, releasing electrons. These electrons are used up in the reduction reaction at the cathode. A common expression for the corrosion current in aqueous corrosion is from the Butler–Volmer equation (Linden and Reddy 2002), which indicates the anode and cathode currents I_a, I_c, respectively, for each electrode

$$i_{corr} = i_a - i_c = I_{oa} \exp\left(\frac{\Delta G_a}{RT} \right) - I_{oc} \exp\left(\frac{\Delta G_c}{RT} \right) \} \qquad (4.110)$$

Here we have identified the barrier height in terms of the Gibbs free energy. The thermodynamic work to change the free energy or equivalently cross the barrier equates to $\Delta G = -W = -qE$, where $q = zH$, H is the Faraday constant, E_o is the open circuit cell potential, and z is the stoichiometric number of electrons in the reaction, that is

$$M \Leftrightarrow M^{Z+} + ne \qquad (4.111)$$

where M is the metal forming M^{Z+} ions in solution. The anode and cathode works is distributed (i.e., $n_a = n\alpha$, $n_c = n(1 - \alpha)$, often $\alpha = 0.5$) so that the anode work amount is, for example,

$$\Delta G_a = -\alpha\, n\, H\, E \qquad (4.112)$$

For $E > 0$ the reaction is spontaneous and no applied potential is required for the reaction to occur. In the case of a battery in steady state with reasonable current flow, anodic reaction can dominate. The AF would be modified, so if a nonspontaneous reaction is forced, the potential is expressed in the acceleration factor as

$$AF \approx \frac{I_{f2}}{I_{f1}} = Exp\left\{-\frac{na\,HE}{R}\left(\frac{1}{T_2} - \frac{1}{T_1}\right)\right\} \tag{4.113}$$

4.10.6 EQUILIBRIUM THERMODYNAMIC CONDITION– NERNST EQUATION

When $I_a = I_c = I_o$ there is no net corrosion current and the equilibrium condition yields

$$C_{oa}K_a^o Exp\left(-\frac{na\,HE}{RT}\right) = C_c K_c^o Exp\left(\frac{(1-a)nHE}{RT}\right) \tag{4.114}$$

Here we substitute into Equation 4.110 the anode and cathode amplitudes (Linden and Reddy 2002) $I_{oa} = C_{oa}K_a^o\,nFA$ and $I_{oc} = C_{oc}K_c^o\,nFA$. Ka and Kc are temperature-dependent rate constants, C_0 is the concentration of the reducing agent at the anode, and C_R is that of the oxidizing agent at the cathode electrode surface.

Collecting terms yields the Nernst thermodynamic equilibrium condition

$$E = E^o + \frac{RT}{nH}\ln\left(\frac{C_R}{C_o}\right) \quad or \quad \Delta G = \Delta G^o - RT\ln\left(\frac{C_R}{C_o}\right) \tag{4.115}$$

where

$$E^o = \frac{RT}{nH}\ln\left(\frac{K_f^o}{K_b^o}\right), \Delta G_o = nHE^o = RT\ln\left(\frac{K_f^o}{K_b^o}\right) \tag{4.116}$$

The Nernst equation enables the calculation of the thermodynamic electrode potential when concentrations are known. It also can indicate the corrosive tendency of the reaction. When the thermodynamic free energy of the process is negative, there is a spontaneous tendency to corrode.

4.10.7 CORROSION RATE IN MICROELECTRONICS

In aqueous corrosion (above) the currents are easier to measure than they are in microelectronics occurring arbitrarily on a circuit board. In this instance, the concept of adding the local percent relative humidity at or near the surface has been found to aid in describing the potential for corrosion to occur. The rate of corrosion and the rate of mass transport are related then to the local percent relative humidity present at the surface, which enhances the electrolyte at the surface for conducting the corrosion currents. The corrosion current is not well defined but is proportional to the rate kinetics and this local relative humidity (%RH) as

$$I_{Corr} \propto (\%RH)^M K(T) \qquad (4.117)$$

In microelectronics, failures due to corrosion are accelerated under higher temperature and humidity conditions than normally occurring during use. In accelerated testing, the acceleration factor between the testing stress and use environment having different temperature and humidity conditions can be found from the ratio of the current in Equation 4.184 as

$$AF_{TH} = \frac{I_{Stress}}{I_{Use}} = \left(\frac{RH_{Stress}}{RH_{Use}} \right)^M \left(\frac{K(T)_{Stress}}{K(T)_{Use}} \right) = AF_H \, AF_T \qquad (4.118)$$

The temperature acceleration factor is deduced from Equation 4.176 as

$$AF_{T,V} = AF_T \, AF_V = Exp \left\{ \frac{E_a}{K_B} \left[\frac{1}{T_{Use}} - \frac{1}{T_{Stress}} \right] \right\} \qquad (4.119)$$

Here, E_a is the activation energy related to the failure mechanism. This is called the Peck's acceleration model (Peck 1986). Typically, E_a and M are not found in testing, but are estimated based on historical data and often $M = 2.66$, and $E_a = 0.7eV$ values are used.

4.11 THERMODYNAMIC DAMAGE IN MECHANICAL SYSTEMS

In this section we will provide examples of how to apply the concept of damage to a number of mechanical systems of creep and wear, as well as an analysis for cyclic vibration and thermal fatigue.

Example 4.9 Creep

Creep parameters include the strain (ε) length change ΔL ($\varepsilon = \Delta L/L$) due to an applied stress (σ) at temperature (T). In the elastic region, stress causes a strain that is recoverable (i.e., reversible) so that $\sigma = Y\varepsilon$, where Y is Young's modulus. When stress increases such that yielding results, the work is irreversible and the permanent plastic strain ε_p damage is termed plastic deformation. The most popular empirical creep rate equation is

$$\varepsilon_p = B_o \sigma^M t^p e^{-E_a/K_b T} \tag{4.120}$$

where B is a material strength constant, p is the time exponent where $0 < p < 1$ for primary and secondary creep stages, M (where $0 < (1/M) < 1$) is strain hardening exponent dependent on the material type, and E_a is the activation energy for the creep process. A linear time dependence is observed in the secondary creep phase where $p = 1$. Some tabulated values are given in Table 4.3 for M and B for the secondary creep rate.

The thermodynamic work causing damage from the creep process of the metal is found from the stress–strain creep area when the stress–strain relation is plotted. Assessing the damage is more accurately found if the data were available. Here, we can use the empirical creep rate expression to find this area by integrating the expression above to determine the work in terms of the applied stress, which is more easily known, so that

$$w = \int \sigma d\varepsilon_P = \int \sigma \frac{d\varepsilon_P}{dt} dt = \int B\sigma^{M+1} pt^{p-1} dt = B\sigma^{M+1} t^p \tag{4.121}$$

Table 4.3. Constants for stress–time creep law (Collins, Busby, and Staab 2010)

Material	Temperature °C	B (in²/lb)^N per day	M
1030 Steel	400	48×10^{-38}	6.9
1040 Steel	400	16×10^{-46}	8.6
2Ni-0.8Cr-0.4Mo Steel	454	10×10^{-20}	3.0
12Cr Steel	454	10×10^{-27}	4.4
12Cr-3W-0.4Mn Steel	550	15×10^{-16}	1.9

where $B = B_o \text{Exp}(-Q/RT)$ and we have assumed that stress is not time dependent. Note that we prefer to use the term $d\varepsilon/dt$ for a constant stress to integrate over time to obtain damage in terms of the applied stress.

In order to assess the damage, we need to have some knowledge of the critical damage at a particular stress, and temperature. Let's assume this occurs at time τ_1 at stress level σ_1, and temperature T_1. Then the thermodynamic damage ratio at any other stress σ_2 and temperature T_2 at time τ_2 along the same work path is

$$Damage = e^{-E_a/K_B\left(\frac{1}{T_2}-\frac{1}{T_1}\right)}\left(\frac{\sigma_2}{\sigma_1}\right)^{M+1}\left(\frac{t_2}{\tau_1}\right)^{p} \tag{4.122}$$

If damage is 1, failure occurs and we can then write this as

$$AF_D = \left(\frac{\tau_2}{\tau_1}\right) = \left[e^{-E_a/K_B\left(\frac{1}{T_1}-\frac{1}{T_2}\right)}\left(\frac{\sigma_1}{\sigma_2}\right)^{-(M+1)}\right]^{1/p} \tag{4.123}$$

Finally, it should be noted that if a number of different stresses are applied, the general creep damage ratio can be written by accumulating the thermodynamic damage along the same work path as

$$Damage = \sum_i\left(\frac{t_i}{\tau_i}\right)^{p} = \sum_i\left(\frac{t_i}{AF_D(1,i)\,\tau_1}\right)^{p} \tag{4.124}$$

Note on Creep Damage Acceleration Factor versus Reliability Acceleration Factor

We see from Equation 4.120 that the reliability strain acceleration factor can be found as

$$AF = \left(\frac{\tau_2}{\tau_1}\right) = \left[e^{-E_a/K_B\left(\frac{1}{T_1}-\frac{1}{T_2}\right)}\left(\frac{\sigma_1}{\sigma_2}\right)^{-(M)}\right]^{1/p} \tag{4.125}$$

This slightly differs in comparison to Equation 4.123 in the stress term power $M + 1$ compared to the above with stress power term M. This is due to the integration for finding the thermodynamic work in Equation 4.121. That is, work is the conjugate force times the conjugate distance. Thus, in terms of the energy damage, there will always be a higher power in

the conjugate force term. However, in reliability assessment, it depends on how one is testing and evaluating. If we use the damage approach, so that we know how much work we are doing on the system, we will need AF_D. If we are simply measuring the conjugate distance change, then we can use the reliability acceleration factor AF. Typically, in cyclic damage, the damage approach has advantages.

EXAMPLE 4.10   Wear

There are many different types of wear, these include abrasive, adhesive, fretting, and fatigue wear. The most common wear model used for adhesive and sometimes abrasive wear of the softer material between two sliding surfaces is the Archard's wear equation (Archard 1953; Archard and Hirst 1956)

$$V = \frac{kPl}{H} \tag{4.126}$$

Here, V is the removed volume of the softer material, P is the normal load (lbs), l is the sliding distance (feet), H is the hardness of the softer material in psi, and k is the Archard's wear coefficient (dimensionless). In the adhesive wear of metals (Archard 1953; Archard and Hirst 1956; Hirst 1957), wear coefficient k varies between 10^{-7} and 10^{-2}, depending on the operating conditions and material properties. It should be recognized that a wear coefficient k is constant typically only within a certain adhesive wear rate range.

We will take the same approach that we used for assessing creep damage. We write the wear volume $V = AD$, where A is the area and D is the depth of the removed. Then writing $l = v\,t$ for two sliding surfaces rubbing against each other at a constant velocity n over a period of wear time t, the thermodynamic damage for constant applied normal stress is

$$w = \int F\,dx = \int F \frac{dD}{dt} dt = \int \frac{kP^2 v}{HA} dt = \frac{kP^2 vt}{HA} \tag{4.127}$$

At some point, we may consider that too much damage has occurred and define this as a failure amount. Then we can assess the damage ratio if we have knowledge of the critical wear depth for a particular velocity v_1 and stress level P_1 (on the softer surface) and corresponding critical failure

time τ_1. Then the damage at any stress level relative to a known failure amount occurring, say, at stress level 1 is

$$Damage = \left(\frac{P_2}{P_1}\right)^2 \left(\frac{A_1 v_2 t_2}{A_2 v_1 \tau_1}\right) \qquad (4.128)$$

(Note that this is not an effective damage, because of the way Equation 4.126 is written in terms of measurable damage). When the damage ratio is 1, we obtain the wear time compression acceleration factor

$$AF_D(2,1) = \left(\frac{\tau_2}{\tau_1}\right) = \left(\frac{P_1}{P_2}\right)^2 \frac{A_2}{A_1} \frac{v_1}{v_2} \qquad (4.129)$$

If a number of different i stresses are applied along the same work path, the general wear damage ratio can be written by accumulating the thermodynamic damage at each stress level as

$$Damage = \sum_i \left(\frac{t_i}{\tau_i}\right) = \sum_i \left(\frac{t_i}{AF_D(1,i)\tau_1}\right) \qquad (4.130)$$

Note, since $l = vt$, we can substitute t for l, and τ for L, if we wish to write this in terms of the sliding distance rather than the wear contact time.

Note on Wear Damage Acceleration Factor versus Reliability Acceleration Factor

We see from Equation 4.126 that the reliability wear amount acceleration factor can be found as

$$AF_D(2,1) = \left(\frac{\tau_2}{\tau_1}\right) = \left(\frac{P_1}{P_2}\right)\frac{A_2}{A_1}\frac{v_1}{v_2} \qquad (4.131)$$

This slightly differs in comparison to Equation 4.129 in the stress term power 2 compared to the above with stress power term 1. This is due to the integration for finding the thermodynamic work in Equation 4.127. That is, work is the conjugate force times the conjugate distance. Thus, in terms of the energy damage, there will always be a higher power in the conjugate force term. However, in reliability assessment, it depends on how one is testing and evaluating. If we use the damage approach, so that we know how much work we are doing on the system, we will need AF_D. If we are simply measuring the conjugate wear amount, then we can use the reliability acceleration factor AF. Typically, in cyclic damage, the damage approach has advantages.

Example 4.11   Thermal Cycle Fatigue - Coffin-Manson Model

In thermal cycling, a temperature change ΔT in the environment, from one extreme to another, causes expansion and/or contraction (i.e., strain) in a material system. The plastic strain (ε) caused by the thermal cyclic stress (σ) in the material can be written similar to Equation 4.120

$$\varepsilon_P = \beta_o n^p \Delta T^j e^{-E_a/K_B T} \tag{4.132}$$

where we have substituted for the nonlinear stress $\sigma^M = \beta \Delta T^j$. The thermodynamic work causing damage from the thermal cycle process is found from the stress–strain creep area if the stress–strain relation could be plotted. The cyclic work is then found similar to Equation 4.121

$$w = \oint \sigma \, d\varepsilon = \int \sigma \frac{d\varepsilon}{dn} dn = \int \beta \Delta T^{j+1} pn^{p-1} dn = \beta \Delta T^k n^p \tag{4.133}$$

where $\beta = \beta_o \text{Exp}(-E_a/K_B T)$ and $k = j + 1$. In order to assess the damage we need to have some knowledge of the critical damage at a particular stress and temperature along this work path. Let's assume this occurs at N_1 at stress level ΔT_1. Then, similar to Equation 4.122, the thermodynamic damage ratio at any other stress ΔT_2 and at the nth cycle n_2 is

$$Damage = e^{-E_a/K_B\left(\frac{1}{T_2}-\frac{1}{T_1}\right)} \left(\frac{n_2}{N_1}\right)^p \left(\frac{\Delta T_2}{\Delta T_1}\right)^k \tag{4.134}$$

If damage is 1, failure occurs and we have

$$AF_D = \left(\frac{N_2}{N_1}\right) = e^{-Q/K_B\left(\frac{1}{T_1}-\frac{1}{T_2}\right)} \left(\frac{\Delta T_1}{\Delta T_2}\right)^K \tag{4.135}$$

where $k/p = K$ and $Q = E_a/p$. This ratio is the temperature cycle "Coffin–Manson" acceleration factors (Coffin 1974; Manson 1953; Manson 1966) when the activation energy Q is small and the Arrhenius effect neglected. For example, in solder joint testing, $K = 2$ for lead-free solder and about 2.5 for lead solder. The activation energy is about 0.123 that is typically used. For example, if the used condition is stress level 1 cycle between 20 and 60 ($\Delta T = 40°\text{C}$, $T_{max} = 60°\text{C}$), while the test stress condition is stress level 2 cycled between $-20°\text{C}$ and 100ºC ($\Delta T = 120°\text{C}$, $T_{max} = 100°\text{C}$), then Arrhenius $AF = 1.58$, while the Coffin–Manson $AF = 9$ with an overall AF

of 14.2. In the case where we have 1 cycle per day in use condition, we see that 10 years of use condition is about 260 test cycles.

The relation Equation 4.133 is also similar to the Norris–Lanzberg (Norris and Lanfzberg 1969) thermal cycle model, which also includes a thermal cycle frequency effect that is used mainly when parts do not fully reach extreme of each temperature.

Lastly, it should be noted that if we have a number of different stresses applied, the general damage ratio can be written by accumulating the thermodynamic damage along the same work path as

$$Damage = \sum_i \left(\frac{n_i}{N_i} \right)^p = \sum_i \left(\frac{n_i}{AF_D(1,i)\,N_1} \right)^p \tag{4.136}$$

Note on Fatigue Damage Acceleration Factor versus Reliability Acceleration Factor

Similar to the above two examples, the exponent for the stress variable ΔT is $j + 1$, where for the reliability acceleration factor it would just be j.

Example 4.12   A Mechanical Cycle (Vibration) Fatigue Time Compression Derivation

In a similar manner to the above argument for thermal cycle, we can find the equivalent mechanical vibration cyclic fatigue damage. In a vibration environment, a vibration level depends on the type of exposure. In testing, typically two types of environments are used, sinusoidal and random vibration profiles. In sinusoidal vibration, the stress level is specified in Gs, where G is a unitless quantity equal to the sinusoidal acceleration A divided by the gravitational constant g. In random vibration, a similar quantity is used, termed G_{rms} (defined below). Consider first the plastic strain (ε) caused by a sinusoidal vibration level G stress (σ) in the material. The strain can be written similar to Equation 4.132

$$\varepsilon = \beta_o n^p G^j \tag{4.137}$$

The cyclic work is then found similar to Equation 4.133 as

$$w = \oint \sigma\, d\varepsilon = \int \sigma \frac{d\varepsilon}{dn} dn = \beta G^{j+1} n^p \tag{4.138}$$

Similar to the above arguments, to assess the damage we need to have some knowledge of the critical damage at a vibration stress. Let's assume this occurs at N_1 at stress level G_1. Then as in Equation 4.134, the thermodynamic damage ratio at any other stress G_2 level at n_2 cycle is

$$Damage = \left(\frac{n_2}{N_1}\right)^p \left(\frac{G_2}{G_1}\right)^b \qquad (4.139)$$

where $b = j + 1$. If damage is 1, failure occurs and we have

$$AF_D = \left(\frac{N_1}{N_2}\right) = \left(\frac{G_2}{G_1}\right)^b \qquad (4.140)$$

As a side note, we again see the difference in the exponent $b = j + 1$ for the damage acceleration factor, where for the strain it would just be j.

AF_D is a commonly used relation for cyclic compression. Since the number of cycles is related to cycle frequency f and the period T according to $N = fT$, then if f is constant

$$AF_D = \frac{N_2}{N_1} = \frac{T_{n1}}{T_{n2}} = \left(\frac{G_2}{G_1}\right)^b_{Sinusoidal} \equiv \left(\frac{G_{rms\,2}}{G_{rms1}}\right)^b_{Random} \qquad (4.141)$$

This is a commonly used for the acceleration factor in sinusoidal testing. For random vibration above we substitute for G the random vibration G_{rms} level (Annex 2008). Note that if we write the cyclic equation for N_2 with $G \propto S$, where S is the stress, we have

$$N_2 = \left(\frac{G_1}{G_2}\right)^b N_1 = C_1 G_b^{-k} = C\ S^{-b} \qquad (4.142)$$

Therefore, we have equated $N_1(G_1)^b = C_1$, a constant. The relationship is generally used to analyze S–N data.

$$N = C\ S^{-b} \qquad (4.143)$$

The experimental number of cycles to failure N for a given stress S level constitutes the S–N curve data in fatigue testing of materials. Such data are widely available in the literature. The slope of the S–N curve provides an estimate of the exponent b above. S–N data are commonly determined using sinusoidal stress.

Most device vibration testing is typically either sinusoidal or random. The goal is to try and accelerate the type of vibration occurring under use conditions. For automotive, for example, this is a random vibration. For a piece of equipment undergoing cyclic motion, this is more likely sinusoidal. The relation for random data is $b/2$ when using the power spectral density (PSD) ($G_{rms}{}^2$/Hz) level instead of G_{rms}. This is evident from the fact that G_{rms} level is found as the square root of the area under the PSD spectrum.

$$G_{rms} = \left(\int_{f_1}^{f_2} PSD(f)df \right)^{\frac{1}{2}} \tag{4.144}$$

In the simple case of random vibration white noise, for example,

$$G_{rms} \cong \sqrt{W_{psd}\,\Delta f} \tag{4.145}$$

then the time compression expression above for random vibration is

$$AF_W = \frac{T_2}{T_1} = \left(\frac{W_{psd1}}{W_{psd2}} \right)^{b/2} = \frac{N_2}{N_1} \tag{4.146}$$

This is a common form used for the random vibration acceleration factor (Annex 2008).

4.12   THERMAL ACTIVATION FREE ENERGY APPROACH

We now ask: how does the free energy change as degradation work occurs? Sometimes, the system path to the free-energy minimum is smooth and downhill all the way to the bottom. For other systems, the path may descend to a relative minimum, but not an absolute minimum, something resembling a roller coaster. The path goes downhill to what looks like the bottom and faces a small uphill region. If that small hill could be scaled, then the final drop to the true minimum would be just over the top of the small hill. The small climb before the final descent to the true minimum is called a free energy barrier. The system may stay for a long period of time in the relative minimum, before the final decay to true equilibrium.

Often, the time spent in the neighborhood of the relative minimum is the lifetime of a fabricated product, and the final descent to the true

free energy minimum represents the catastrophic failure of the product. The estimated lifetime τ in which the system stays at the relative minimum obeys the Arrhenius law $(1/\tau) = 1/\tau_o \; exp \; (-\Delta\phi/K_BT)$, where $\Delta\phi$ is the height of the free energy barrier and is the activation energy needed to surmount the barrier.

__The activation energy__ can be thought of as the amount of energy needed for the degradation process. This defines a special relative equilibrium ageing state,

$$\varphi_B \leq \varphi \leq \varphi_B + \Delta\varphi \qquad (4.147)$$

where the subscript B refers to the barrier and Δf is the barrier height.

4.12.1 THERMAL ACTIVATION TIME-DEPENDENT DEGRADATION MODEL

When activation is the rate-controlling process, Arrhenius-type rate kinetics applies. In this section, the parametric time-dependence of an Arrhenius mechanism is addressed. This mechanism turns out to explain logarithmic-in-time ageing of many key device parameters described here. Such mechanism, with temperature as the fundamental thermodynamic stress factor, leads to this predictable logarithmic-in-time-dependent ageing on measurable parameters. Here the thermodynamic work is evident in the free energy path.

There are two thermally activated time-dependent (TAT) device degradation models (Feinberg and Crow 2001; Feinberg and Widom 2000). If degradation is graceful over time, then many instances occur in which logarithmic-in-time ageing results for one or more device parameters. This ageing and examples are described below. Degradation leading to catastrophic failure using the TAT model is presented in the appendix.

In Arrhenius processes, the probability rate dp/dt to surmount the relative minimum free energy barrier ϕ is as we mentioned in 4.12.1 can be written

$$\frac{dp}{dt} = \nu \; exp\left(-\frac{\varphi}{K_BT}\right) \qquad (4.148)$$

where ν is a rate constant, K_B is the Boltzmann's constant, T is the temperature, and t is the time. We wish to associate the thermodynamic ageing kinetics in the material with measurable parametric changes.

Therefore, we model the above as a fractional rate of parametric change given by

$$\frac{da}{dt} = \nu \, \exp\left(-\frac{\varphi(a)}{K_B T}\right) \qquad (4.149)$$

where a is the fractional change of the measurable parameter. We can let "a" be unitless as well as $a = \Delta P/Po$. For example, ΔP could be a parameter change that is of concern, such as resistance change, current change, mechanical creep–strain change, voltage transistor gain change, and so forth, such that a is then the fractional change.

In the TAT Model, the ageing process is closely related to the parametric change. The assumption is that the free energy itself will be associated with the parameter through the thermodynamic work. Thus, f will be a function of a as indicated in Equation 4.148.

Then the free energy can be expanded in terms of its parametric dependence using a Maclaurin series (with environmental factor held constant for the moment). The free energy reads

$$\varphi(a) = \varphi(0) + a y_1 + \frac{a^2}{2} y_2 + \dots \qquad (4.150)$$

where y_1 and y_2 are given by

$$y_1 = \frac{\partial \varphi(0)}{\partial a} \quad and \quad y_2 = \frac{\partial^2 \varphi(0)}{\partial a^2} \qquad (4.151)$$

4.12.2 ARRHENIUS AGEING DUE TO SMALL PARAMETRIC CHANGE

When $a \ll 1$, the first and second terms in the Maclaurin series yield

$$\frac{da}{dt} = \nu(T) \, \exp\left(-\frac{a y1}{K_B T}\right) \qquad (4.152)$$

where $\nu(T) = \nu_o \exp\left(-\frac{\varphi(0)}{K_B T}\right)$.

Rearranging terms, and solving for a as a function of t and integrating provides a *logarithmic-in-time ageing TAT model,* where

$$a = \frac{\Delta P}{P} \cong A \ \ln[1 + B \ t] \qquad \text{for a} << 1 \qquad (4.153)$$

Here A and B are

$$A = \frac{K_B T}{y_1} \quad and \ B = \frac{v(T) y_1}{K_B T} \qquad (4.154)$$

Logarithmic-in-time ageing is an important process since the origin of this ageing kinetics can mathematically be tied to the Arrhenius mechanisms of which numerous experimental examples exist (Feinberg and Crow 2001; Feinberg and Widom 2000). Figure 4.5 illustrates typical logarithmic in time ageing. One notes that ageing is highly nonlinear for

(a)

(b)
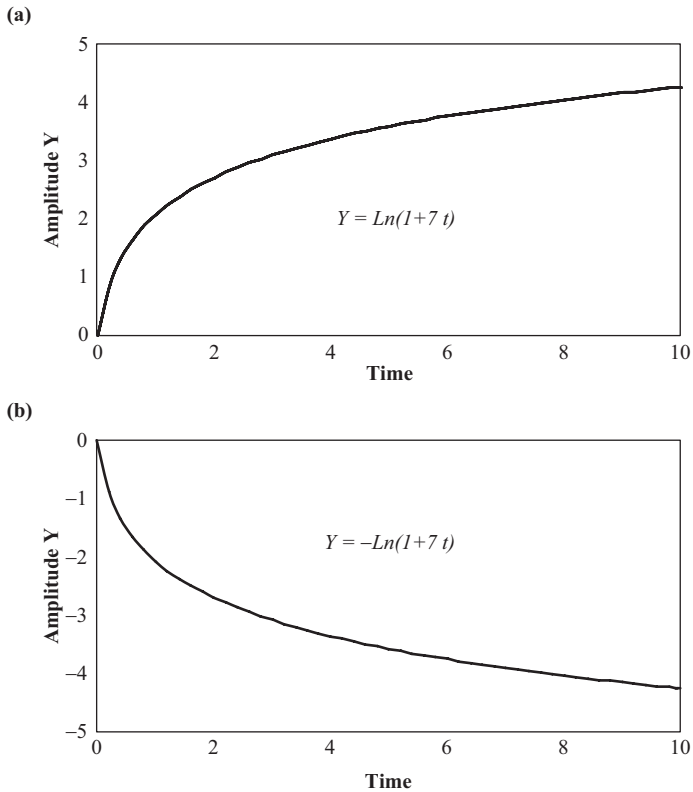
Figure 4.5. Examples of *Ln*(1+B time) ageing law with (a) similar to primary and secondary creep stages and (b) similar to primary battery voltage loss.

early time. This curve is representative of many ageing and kinetic processes such as crystal frequency ageing (Warner, Fraser, and Stockbridge 1965), corrosion of thin films, chemisorption processes (Ho 2006), early degradation of primary battery life (Linden 1980), activated creep (see Example 4.14), activated wear (see Example 4.13), transistor key parameter ageing (see Sections 4.11.5, 4.11.6), and so forth. The significance of parametric logarithmic-in-time ageing can further be put in perspective as it can be tied to catastrophic lognormal failure rates. This is discussed in the appendix (Feinberg and Crow 2001; Feinberg and Widom 2000).

Example 4.13   Activation Wear

Activation wear is a term we use here to associate wear to a thermally activated process. This approach offers an alternative to the Archard's empirical Type I wear model discussed above (see Equation 4.126). When logarithmic-in-time ageing occurs as illustrated in Type II wear shown in Figure 4.6 (often observed in metals [Bhushan 2001; Chiou, Kato, and Kavaba 1985]), the TAT model can apply. This is characterized by an initially high wear rate, then to a steady-state low wear rate. Type I shows the case of steady wear in time, where the surface roughness does not change from the initial value. Type II shows the case of steady wear, where surface roughness increases to a certain value and does not increase much
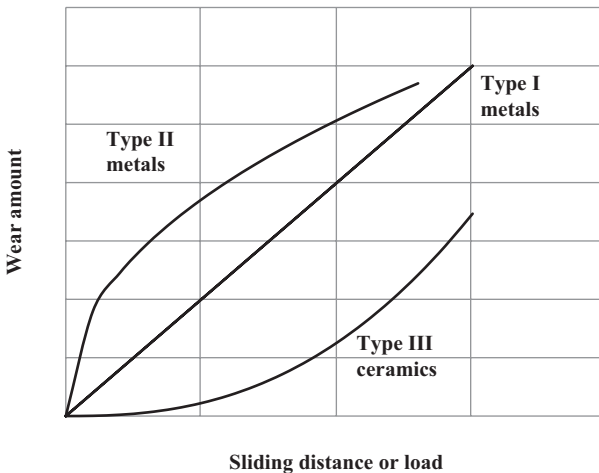
Figure 4.6. Types of wear dependence on sliding distance (time) (Bhushan 2001; Chiou, Kato, and Kavaba 1985).

after that over time. Type III typically observed in lapping and polishing for surface finishing of ceramics.

In activation wear, we consider surface mass trapped in a potential well. We need to apply enough thermal energy for mass to escape from the chemical bonds. When a normal force is applied to the surface due to contact wear under a constant velocity, the normal force P results in material removal. We will start by taking a naive kinetic friction approach

$$F_N = d/dt(Mv) = dM/dt \, v \, \alpha \, (\mu_{keff} P) \tag{4.155}$$

where μ_{keff} is an effective coefficient of sliding kinetic friction and the velocity v is constant. Here we start with a simple kinetic frictional model to get the general form. For an activated process of mass removal, we write this

$$\frac{dM}{dt} \, v = \mu_{KEff}(T) \, P = \mu_{KEff} \, P \exp\left(-\frac{\varphi(M)}{K_B T}\right) \tag{4.156}$$

Here, $\mu_{KEff}(T)$ is an temperature-dependent effective kinetic coefficient of sliding friction, which must be generalized

$$\frac{dM}{dt} = \gamma \exp\left(-\frac{\varphi(M)}{K_B T}\right) \tag{4.157}$$

In this case $\gamma = \dfrac{\mu_k P}{v}$. We have to now depart from our naïve approach as kinetic friction, although insightful, is not helpful enough in accurately depicting the amplitude in wear. To be consistent and specify the model a bit further, we can use historical information such as the Archard's wear model in Equation 4.126. We have a few options, we write the amplitude

$$\gamma = \frac{kP v \rho}{H} = \frac{\bar{M}_o}{\tau} \tag{4.158}$$

That is, we can either make use of the Archard's amplitude or write the amplitude in experimental terms of an initial average mass removal $\bar{M}_o$, and τ, an average time constant for removal. This amplitude can be found experimentally as we will describe. In addition, we have added ρ, the density of the softer material to 4.125 for consistent units with mass removal and the other parameters are defined in 4.125. Expanding the free energy as a function of the mass removal, we have

$$\varphi(M) \approx \varphi(0) + \frac{d\varphi}{dM} M + \dots = \varphi(0) + \mu M + \dots \qquad (4.159)$$

We identify μ as the *activation chemical potential per unit mass,* which now gives

$$\frac{dM}{dt} = \gamma(T) = \gamma_o \exp\left(-\frac{\varphi(0) + \mu M}{K_B T}\right) = \gamma(T) \exp\left(-\frac{\mu M}{K_B T}\right) \quad (4.160)$$

and ϕ is the activation energy for the wear process, which can be found experimentally in testing at different temperatures, as explained below. Rearranging terms, and solving for the mass as a function of time t and integrating provides a logarithmic-in-time ageing TAT model in terms of mass removal over time

$$M = A\, ln(1 + Bt) \qquad (4.161)$$

Here A and B are

$$A = \frac{K_B T}{\mu} \quad and \quad B = \frac{\gamma(T)\mu}{K_B T} \qquad (4.162)$$

Note the logarithmic-in-time dependence in the activation wear case differs from the Archard's linear dependence (i.e., $l = vt$). Here we find that when Bt in Equation 4.161 is less than or of the order of 1, the removal amount is large at first, then is less as time accumulates—nonlinear in time removal. However, when $Bt \gg 1$, the removal is in $ln(t)$ dependence. Also, since $ln(1 + X) \sim X$ for $X \ll 1$, then Equation 4.161 can be approximated as $M \approx ABt = \gamma(T)t$ for $Bt \ll 1$, which upon substitution of Equation 4.158 agrees with Equation 4.126, the Archard wear equation. Note that one could make use of this early time approximation to help determine $\gamma_o = \frac{\overline{M}_o}{\tau}$ and the wear activation energy f experimentally.

Example 4.14   Activation Creep Model

Activation creep is a term we use here to associate creep due to a thermally activated process where the activation free energy is a function of the strain itself. This approach offers an alternative approach to modeling the creep process. The general model indicates that for creep to occur under temperature activation that a number of dislocations will occur over time in the metal lattice. This is when, due to temperature stress, dislocations N

have a probability to hop over a potential barrier and weaken the crystal metal lattice. Its hopping rate of occurrence is

$$\frac{dN}{dt} = N_o \exp\left(-\frac{\varphi(N)}{K_B T}\right) \tag{4.163}$$

Here, in the TAT model, the activation energy is a function of N. In the case of creep, N would be proportional to strain. However, this relation must of course be modified when a load is applied and the material is also subjected to mechanical means. We will therefore start with a popular form of the empirical creep rate equation that is typically written

$$\frac{d\varepsilon}{dt} = \varepsilon_o \sigma^N \exp\left(-\frac{\varphi(\varepsilon)}{K_B T}\right) \tag{4.164}$$

Note the difference from Equation 4.120 on the right hand side regarding the time dependence of this popular form for the creep rate. Also the key difference in this approach is that we associated the creep rate process with the activation free energy as a function of the strain itself. We will at this point use the TAT model by expanding it in terms of the strain dependence in a Maclaurin series

$$\varphi(\varepsilon) \approx \varphi(0) + \frac{d\varphi}{d\varepsilon} \varepsilon + \dots \tag{4.165}$$

We identify that the change in the free energy is due to damage that occurred from mechanical thermodynamic conjugate work (see Table 4.1)

$$\frac{d\varphi}{d\varepsilon} = \frac{dW}{d\varepsilon} = \sigma \tag{4.166}$$

This then yields

$$\frac{d\varepsilon}{dt} = \varepsilon_o \sigma^N \exp\left(-\frac{\varphi(\varepsilon)}{K_B T}\right) = \varepsilon_0(T)\sigma^N \exp\left(-\frac{\sigma\varepsilon}{K_B T}\right) \tag{4.167}$$

This is now in the form of the TAT model (see Equations 4.153 and 4.154) we find in integration

$$\varepsilon = \frac{\Delta L}{L} \cong A \ \ln[1 + B \ t] \qquad \text{for } \varepsilon << 1 \tag{4.168}$$

and

$$A = \frac{K_B T}{\sigma} \ \text{and} \ B = \frac{\varepsilon_o(T)\sigma^{N+1}}{K_B T} \tag{4.169}$$

or

$$\varepsilon = \frac{\Delta L}{L} \cong \frac{K_B T}{\sigma} \ \ln[1 + \frac{\varepsilon_o(T)}{K_B T} \sigma^{N+1} \ t] \tag{4.170}$$

where $\varepsilon_o(T) = \varepsilon_o \exp\left(-\frac{\varphi(0)}{K_B T}\right)$.

We note that we actually had to start with an empirical expression in Equation 4.164 for the creep rate. Once we expanded the activation free energy in terms of the strain, the creep rate indicated a nonlinear time dependence $\ln(1 + Bt)$. The logarithmic-in-time dependence model has the common curvature found in primary and secondary creep stages. A power law dependence $\varepsilon \ a \ \tau^b$, where $0 < b < 1$ is also popular as pointed out in 4.120, and both the $\ln(1 + Bt)$ and power law form have been used by other authors (see e.g., Lubliner 2008). Interesting enough we see the stress amplitude is an inverse relation to strain, which of course by itself would be incorrect. However, the argument in the natural log function includes strain to the $N + 1$ power, which as long as N is larger than 0, the overall dependence will actually show that the strain increase with stress. Finally, the activation energy should be the same $\phi(0)$ as reported in the literature. The third stage of creep can also be modeled using the TAT model. This concept is provided in the chapter appendix.

4.12.3 Transistor Ageing

Here we illustrate how we can extend the TAT model to transistor ageing (Feinberg et al. 2000). We are primarily concerned with key transistor device parameters. In the bipolar case for the common-emitter configuration, the key transistor parameter is beta ageing, showing it to be directly proportional to the fractional change in the base-emitter leakage current. In the Field-Effect Transistor (FET) case, the key transistor parameter considered is transconductance ageing that results from a change in the drain–source resistance and gate leakage current. Then the TAT model is used to provide an ageing expression that accounts for the time degradation of

these parameters found in life test. These expressions provide insight into degradation that links ageing to junction temperaturedependent mechanisms. The mechanisms for leakage can be thought of similar to a corrosion process having corrosion current. We have all the similar components— an anode, cathode, a conducting path, and an effective type of electrolyte. Some typical life test data on HBTs and Metal Semiconductor Field-Effect Transistor (MESFETs) are illustrated.

4.12.4 BIPOLAR BETA AGEING MECHANISM

Generally, there are two common bipolar ageing mechanisms: an increase in emitter ohmic contact resistance and an increase in base leakage currents. Both can be thought of as we mention as a corrosion occurring process. Since β is given by I_{ce}/I_{be} in the common emitter configuration, any base leakage degradation in I_{be} will degrade β. In this section, we describe a model for β degradation (Feinberg et al. 2000) over time due to leakage. We start by considering a change in the base current gain for the common emitter configuration as

$$\beta(t) = \beta_o - \left|\Delta\beta(t)\right| \tag{4.171}$$

where β_o is the initial value (prior to ageing) of I_{ce}/I_{be}. The time-dependent function $\Delta\beta(t)$ can be found through the time derivative

$$\dot{\beta}(t) = \frac{d\beta}{dt} = \frac{d}{dt}\left(\frac{I_{ce}}{I_{be}}\right) = \frac{\dot{I}_{ce}}{I_{be}} - \frac{I_{ce}}{I_{be}^2}\dot{I}_{be} = \beta_o\left(\frac{\dot{I}_{ce}}{I_{ce}} - \frac{\dot{I}_{be}}{I_{be}}\right) \tag{4.172}$$

Approximating d/dt by $\Delta/\Delta t$ and noting that Δt is common to both sides of the equation and cancels, then

$$\Delta\beta(t) \approx \beta_o\left(\frac{\Delta I_{ce}(t)}{I_{ce}} - \frac{\Delta I_{be}(t)}{I_{be}}\right) \approx -\beta_o\left(\frac{\Delta I_{be}(t)}{I_{be}}\right) = \beta_o\left(\frac{\left|\Delta I_{be}(t)\right|}{I_{be}}\right) \tag{4.173}$$

In the above equation, ΔI_{ce} has been set to zero as no change in this parameter is usually observed experimentally. We added the absolute value sign since beta overall will degrade ($\beta(t) = \beta_o - \Delta\beta(t)$) as the leakage current to the base increases. Thus, the first result is that the change in β is directly proportional to the fractional change in the base-emitter leakage current.

At this point, base charge storage is discussed in order to develop a useful capacitive model. When a transistor is first turned on, electrons penetrate into the base bulk gradually. They reach the collector only after a certain delay time τ_d. The collector current then starts to increase, in relation to the current diffusion rate. Concurrent with the increase of the collector current is excess charges build up in the base. As a first approximation, the collector current and excess charge increases in an exponential manner with time constant τ_b. This transient represents the process of charging a "capacitor" in the simplest of RC circuits shown in Figure 4.7. We use this approximation to provide a simple model for base leakage. The steady-state value of excess charge build up in base-emitter bulk Q_k is then

$$Q_k = (Q_{be})_k \cong (C_{be} V_{be})_k = (C_{be} (I_{be} R_{be}))_k = (I_{be} \tau_b)_K \qquad (4.174)$$

where $\tau_b = R_{be} C_{be}$ is the time constant for steady-state excess charges in the base-emitter junction. As discussed above, the base-emitter junction primarily contributes to ageing effects. Along with this bulk effect is parasitic surface charging Q_s and leakage. We can also treat these using a simple RC charging model. In this view, the surface leakage can be expressed as

$$Q_s = (Q_{be})_s \cong (C_{be} V_{be})_s = (C_{be} (I_{be} R_{be}))_s = (I_{be} \tau_b)_s \qquad (4.175)$$

The total excess charging at the base is

$$Q_{be} = Q_s + Q_k \qquad (4.176)$$

As the transistor ages, Q_{be} increases along with I_b. Some of the increase in Q_{be} is caused by the increase in impurities and defects in the base surface and bulk regions due to operating stress.

The impurities and defects cause an increase in electron scattering, and an increase in the probability for trapping and charging and eventual recombination in the base. The above features lead to an increased leakage current. In the capacitive model shown in Figure 4.7, incremental changes are

$$dQ = C \, dV = C R \, dI = \tau \, dI \qquad (4.177)$$

Here, we view Q, V, and I as time varying with age, that is

$$\Delta\beta(t) \cong -\beta_o \left(\frac{\Delta I_{be}(t)}{I_{be}} \right) = -\beta_o \left(\frac{\Delta Q_{be}(t)}{Q_{be}} \right) = -\beta_o \left(\frac{\Delta V_{be}(t)}{V_{be}} \right) \qquad (4.178)$$
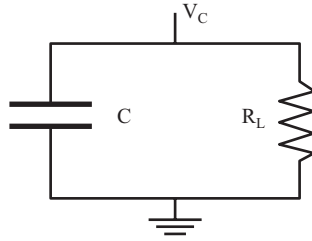
Figure 4.7. Capacitor leakage model.
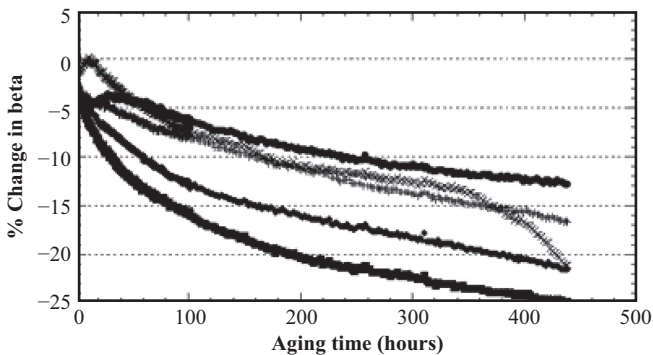
Figure 4.8. Life test data on C-doped MBE HBT devices at 235°C at 10 kA/Cm².

Thus, a second result is that the change in β is proportional to the fractional change in the base-emitter leakage current, charge, and voltage.

Experimentally, β degradation is observed to follow a logarithmic-in-time ageing equation (Feinberg et al. 2000) as shown in the Figure 4.8 and in the next section is described by the model

$$\frac{\Delta\beta(t)}{\beta_o} = \frac{\Delta I_{be}(t)}{I_{be}} = \frac{\Delta q_{be}(t)}{q_{be}} = -A\,Log(1+B\,t) \qquad (4.179)$$

4.12.5 TAT MODEL FOR TRANSISTORS AND DIELECTRIC LEAKAGE

The leakage current often leading to dielectric breakdown has historically been explored and there are numerous mechanisms, or a combination of several, providing physical explanation of the origin of the current that

flows through the dielectric. These models stem from the fact that the work required to create defects that increase leakage current is thermally activated and there is a reduction in the free energy barrier for defect generation due to the electric field lowering the barrier for defect creation. For example, in a bond breakage model, the E-field is related to the bonds breakage through the dipole energy. The energy barrier for creating defects has the form

$$\varphi(q) \approx \varphi(0) + \frac{d\varphi}{dq} q + \dots \tag{4.180}$$

We identify that the change in the free energy is due to damage that occurs from electrical thermodynamic conjugate work Table 4.1

$$\frac{d\varphi}{dq} = \frac{dW}{dq} = V \tag{4.181}$$

Common dielectric leakage mechanisms that relate to various forms of the free energy with different explanations are the Poole–Frenkel Effect, Schottky Effect, Thermochemical E-model, Tunneling, and Fowler–Nordheim Tunneling. The most general expression for current density j for these models is

$$j = CE^K \exp\left\{\frac{-(\varphi_o - a\, E^N)}{K_B T}\right\} \tag{4.182}$$

where E is the electric field, a and C are constants. For the Poole–Frenkel (Harrell and Frey 1999) model $K = 1$, $N = 1/2$, for the Schottky effect (Sze 1981) model $K = 0$, $N = 1/2$, for the Thermo-chemical model $K = 0$, $N = 1$, for the tunneling models $K = 2$, $E = -1$ (Sze 1981).

We can start our leakage TAT model consistent with Equation 4.182 and use Equations 4.180 and 4.181 for the free energy with $N = 1$, and the general constant K as

$$\frac{dq}{dt} = q_o V^K \exp\left(-\frac{(\varphi(0) + qV)}{K_B T}\right) = q_0(T)V^K \exp\left(-\frac{qV}{K_B T}\right) \tag{4.183}$$

This is in the form of the TAT model results (see Equations 4.153 and 4.154), where we find

$$\Delta\beta = \frac{\Delta q_{be}}{q_{be}} = \frac{\Delta I_{be}}{I_{be}} \cong A\, \ln[1 + B\,t] \quad \text{for } \Delta b_{\text{loss}} \ll 1 \tag{4.184}$$
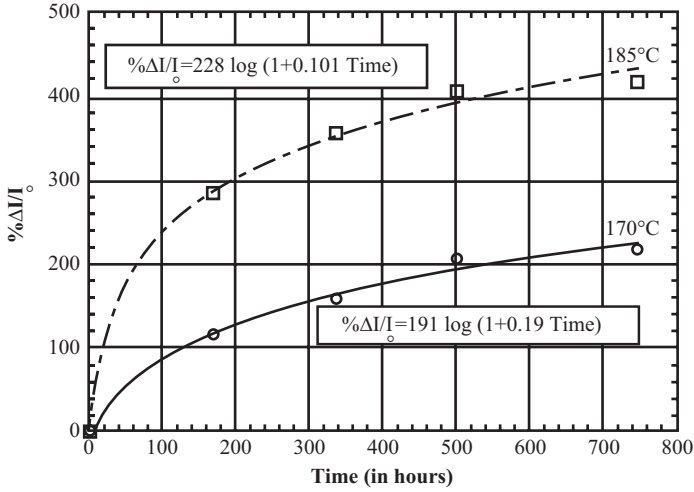
Figure 4.9. Life test data of gate-source MESFET leakage current over time fitted to the $ln(1 + Bt)$ ageing model. Temperatures are ambient. Junction rise is about 30°C (Feinberg et al. 2000).

where $A = \dfrac{K_B T}{I_{be} V}$ and $B = \dfrac{q(T) V^{K+1}}{K_B T}$ or

$$\Delta \beta(t) \cong \frac{K_B T}{V I_{be}} \ln\left[1 + \frac{q_o(T)}{K_B T} V^{K+1} t\right] \qquad (4.185)$$

where $q_o(T) = q_o \exp\left(-\dfrac{\varphi(0)}{K_B T}\right)$.

We expect K > 0 for Beta degradation $\beta(t) = \beta_o - |\Delta\beta(t)|$ to increase with $V (=Ed)$. The result of this logarithmic-in-time model yields a good fit to the Beta degradation data in Figure 4.8 and as well for the FET data in Figure 4.9.

4.12.6   FET PARAMETER DEGRADATION

In this section, transconductance degradation over time is described to help understand ageing in FET devices, such as MESFETs, and what role problems like leakage current play (Feinberg et al. 2000). We start by modeling a change in the transconductance g_m as

$$g_m(t) = g_o - \left|\Delta g_m(t)\right| \qquad (4.186)$$

where g_o, the initial value, is taken in the linear portion of the transconductance curve, is

$$g_o = \frac{I_{DS}}{V_{GS} - V_o} \tag{4.187}$$

Here, we use the linear portion of the curve for simplicity. Similar results will follow for other portions of the curve. The time-dependent function $\Delta g_m(t)$ is found from its derivative as

$$\dot{g}_m(t) = \frac{dg_m}{dt} = \frac{d}{dt}\left(\frac{I_{DS}}{V_{GS} - V_o}\right) = \frac{\dot{I}_{DS}}{V_{GS} - V_o} - \left(\frac{I_{DS}}{(V_{GS} - V)^2}\right)\dot{V}_{GS} \tag{4.188}$$

or

$$\dot{g}_m(t) = \frac{I_{DS}}{V_{GS} - V_o}\left(\frac{\dot{I}_{DS}}{I_{DS}} - \frac{\dot{V}_{GS}}{V_{GS} - V_o}\right) = g_o\left(\frac{\dot{I}_{DS}}{I_{DS}} - \frac{\dot{V}_{GS}}{V_{GS} - V_o}\right) \tag{4.189}$$

We assume that the drain–source current change occurs as $dI/dt \sim (V/R^2)(dR/dt)$ with V_{DS} constant and voltage-gate changes as $dV_{GS}/dt \sim d/dt(IR) = RdI_{GS}/dt$. Approximating d/dt by $\Delta/\Delta t$ and noting that Δt is common to both sides of the equation and canceling, the expression simplifies to

$$\Delta g_m(t) = g_o\left(\frac{\Delta R_{DS}}{R_{DS}} - \frac{\Delta I_{GS}}{I_{GS} - I_{GS_o}}\right) \tag{4.190}$$

Thus, the primary result for FETs is that transconductance ageing arises from a change in the drain–source resistance and gate leakage. However, it is commonly found that resistance ageing dominates the reaction (Feinberg et al. 2000). As far as R_{DS} is concerned, resistance is related to scattering inside the drain–source channel $\Delta R_{DS}/R_{DS} = \Delta \rho_{DS}/\rho_{DS} = \Delta l_{DS}/l_{DS}$, where ρ is the resistivity, and l is the average mean-free path that the electrons in the channel travel between collisions. This distance

decreases as ageing occurs and more defects occur in the channel causing increased scattering.

Here, we wish to point out that similar to β degradation (a mechanism that we have modeled as dominated by leakage), MESFET gate leakage data, as shown, commonly follows a logarithmic-in-time ageing form as well. This time dependence is illustrated in Figure 4.9. The TAT model for gate leakage similar to Equation 4.185 is

$$\frac{\Delta I_{GS}(t)}{I_{GS}} \cong -\frac{K_B T}{V I_{GS}} \ln\left[1 + I_o(T)\frac{V^{K+1}}{K_B T}\ t\right] \qquad (4.191)$$

where $I_o(T) = I_o \exp\left(-\frac{\varphi(0)}{K_B T}\right)$.

APPENDIX

PARAMETRIC AGEING AT END OF LIFE DUE TO THE ARRHENIUS MECHANISM

A second TAT model can be obtained for both the initial ageing period and end of life using both terms in the Maclaurin expansion in Equations 4.149 and 4.150 and performing the integration. The results obtained in Feinberg and Crow (2001) and Feinberg and Widom (2000) are

$$a = \xi + b\ erf^{-1}\left[\exp(-K^2)\left(\beta\frac{\nu(T)}{b}\right)t + erf(K)\right]$$

where erf and erf^{-1} are the error function and its inverse and

$$\xi = y1 / y2, \ \ b = -2K_B T / y2, \ \ K = \xi / b, \ \ and \ \beta = 2 / \sqrt{2\pi}$$

This model is a parametric ageing phenomenon that ages similar to logarithmic-in-time models and quickly goes catastrophic at the end-of-life due to Arrhenius degradation. This is illustrated in Figure 4.10. The figure shows that ageing starts off similar to logarithmic-in-time ageing and then quickly goes catastrophic at the critical corresponding time t_c. The reader is referred to Feinberg and Widom (1996) for details.

Figure 4.10 illustrates a number of rate processes. Some examples are batteries: (Feinberg and Widom 1996), the three phases of creep, and cold-worked metals recrystallizing exhibit forms of this dependence over

$$\frac{\Delta P}{P_0} \alpha \; Erf^{-1}(K_1 Time + K_2)$$

$$\frac{\Delta P}{P_0} = ALog(1 + b \; Time)$$

$t_{Critical}$

t_c

Time

$\frac{\Delta P}{P_0}$
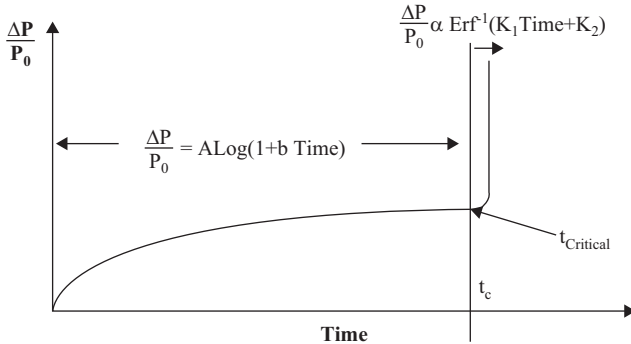
Figure 4.10. Ageing over all time t with critical values t_c prior to catastrophic failure 3.

time. What is interesting in this model is that the rate of initial ageing is mathematically connected to its rate of final catastrophic behavior in this model. This suggests that if the initial ageing process is truly understood, a catastrophic prognostic may be possible!

REFERENCES

Archard, J.F. 1953. "Contact and Rubbing of Flat Surface." *Journal of Applied Physics* 24, no. 8, pp. 981–88. doi: http://dx.doi.org/10.1063/1.1721448

Archard, J.F., and W. Hirst. 1956. "The Wear of Metals under Unlubricated Conditions." *Proceedings of the Royal Society* A-236, pp. 397–410. doi: http://dx.doi.org/10.1016/0043-1648(57)90017-0

Bhushan, B. ed. 2001. *Modern Tribology Handbook*, K. Kato, and K. Adachi. CRC Press.

Chiou, Y.C., K. Kato, and T. Kayaba. 1985. "Effect of Normal Stiffness in Loading System on Wear of Carbon Steel—Part 1: Severe-mild Wear Transition." *ASME, Journal of Tribology* 107, pp. 491–95. doi: http://dx.doi.org/10.1115/1.3261114

Coffin, L.F. 1954. "A Study of the Effects of Cyclic Thermal Stresses on a Cuctile Metal." *Transaction ASME* 76, pp. 923–50. doi: http://dx.doi.org/10.1080/01495739208946130

Coffin, L.F. 1974. "Fatigue at High Temperature—Prediction and Interpretation." James Clayton Memorial Lecture. *Proceedings of the Institution of Mechanical Engineers (London)* 188, pp. 109–27. doi: http://dx.doi.org/10.1243/pime_proc_1974_188_014_02

Collins, J.A., H. Busby, and G. Staab. 2010. *Mechanical Design of Machine Elements and Machines.* 2nd ed. New York: Wiley.

Feinberg, A. January 2015. Thermodynamic Damage Measurements of an Operating System. *Reliability and Maintainability Symposium.*

Feinberg, A., and A. Widom. March 1996. "Connecting Parametric Ageing to Catastrophic Failure through Thermodynamics." *IEEE Transaction on Reliability* 45, no. 1, p. 28. doi: http://dx.doi.org/10.1109/24.488913

Feinberg, A., and A. Widom. January 2000. "Thermodynamic Extensions of Miner's Rule to Chemical Cells." *Reliability and Maintainability Symposium.* pp. 341–44. doi: http://dx.doi.org/10.1109/rams.2000.816331

Feinberg, A., and D. Crow, eds. 2001. *Design for Reliability.* CRC Press.

Feinberg, A., P. Ersland, V. Kaper, and A. Widom. 2000. "On Ageing of Key Transistor Device Parameters." *Proceedings - Institute of Environmental Sciences and Technology*, p. 231.

Feinberg, A., and A. Widom. June 2000. "On Thermodynamic Reliability Engineering." *IEEE Transaction on Reliability* 49, no. 2, p. 136. doi: http://dx.doi.org/10.1109/tr.2000.877330

Harrell, W.R., and J. Frey. 1999. "Observation of Poole–Frenkel Effect Saturation in SiO_2 and other Insulating Films." *Thin Solid Films* 352, pp. 195–204. doi: http://dx.doi.org/10.1016/s0040-6090(99)00344-2

Hazewinkel, M., ed. 2001. Differential Entropy. *Encyclopedia of Mathematics. Springer*.

Hirst, W. 1957. *Proceedings of the Conference on Lubrication and Wear*. Institution of Mechanical Engineers, London, p. 674.

Ho, Y-S. 2006. "Review of Second-order Models for Adsorption Systems." *Journal of Hazardous Materials* B136, pp. 681–89. doi: http://dx.doi.org/10.1002/chin.200648222 http://en.wikipedia.org/wiki/Differential_entropy

Linden, D., ed. 1980. *Handbook of Batteries and Fuel Cells*. New York: McGraw-Hill.

Linden, L. and T.B. Reddy. 2002. *Handbook of Batteries*. 3rd ed. New York: McGraw-Hill.

Lubliner, J. 2008. "The Physics of Plasticity." *Plasticity Theory.* (see Chapter 2), Dover.

Manson, S.S. 1953. *Behaviour of Materials under Conditions of Thermal Stress. NACA-TN-2933 from NASA*. Cleveland, OH: Lewis Research Center.

Manson, S.S. 1966. *Thermal Stress and Low Cycle Fatigue*. New York: McGraw-Hill.

Miner, M.A. 1945. "Cumulative Damage in Fatigue." *Journal of Applied Mechanics* 12, pp. A159–64.

Norris, K.C., and A.H. Landzberg. 1969. "Reliability of Controlled Collapse Interconnections." *IBM Journal of Research and Development* 13, no. 3, pp. 266–71. doi: http://dx.doi.org/10.1147/rd.133.0266

Peck, D.S. 1986. "Comprehensive Model for Humidity Testing Correlation." *Proceedings of the IEEE International Reliability Physics Symposium* 24, pp. 44–50. doi: 10.1109/IRPS.1986.362110 *MIL-STD-810G*. Method 514.6, Annex, A. October 2008.

Sze, S.M. 1981. *Physics of Semiconductor Devices*. John Wiley and Sons.

Warner, A.W., D.B. Fraser, and C.D. Stockbridge. 1965. "Fundamental Studies of Ageing in Quartz Resonators." *IEEE Transaction on Sonics and Ultrasonics* 12, no. 2, pp. 52–58. doi: http://dx.doi.org/10.1109/t-su.1965.29361

CHAPTER 5

MONITORING DEGRADATION IN THE FIELD

Xiandong Ma

5.1   INTRODUCTION

The range of material degradation mechanisms can be classified into four basic categories based on direct mechanical action, heat or radiation, the presence of chemical reagents, and a combination of the above mechanisms (Batchelor, Nee Lam, and Chandrasekaran 2011). In general, materials are continually changing at either molecular or macroscopic scale. All materials including wood, ceramics, and plastics deteriorate at the surface to varying degrees when they are exposed to certain combinations of ultraviolet (UV) light from sunshine, liquids, gases, or contact with other solids. For metals, deterioration is usually caused as a result of a reaction with its environment, especially with oxygen. Most metals corrode because they react with oxygen in the atmosphere, particularly under moist conditions, which is called oxidation. Ferrous metals such as steel are particularly susceptible to oxidation and require ongoing maintenance or they will suffer inevitable structural failure. Some nonferrous metals are particularly resistant to corrosion, for example, copper. The corrosion has a significant effect on the mechanical and physical properties of the metallic materials. These include the reduction of metal thickness leading to loss of strength or complete structural failure, localized corrosion leading to cracks in the structure, a disproportionate weakening in comparison to the amount of metal lost, and fatalities and damages leading to structural failure.

The monitoring of degradation in the field has been a difficult and challenging task over the past decades. Among the monitoring techniques commonly used, nondestructive testing (NDT) and evaluation has been widely recognized as an effective approach for monitoring degradation. This technique is primarily concerned with the development of measurement technologies and analysis techniques for the quantitative characterization of materials, tissues, and structures by noninvasive means. Ultrasonic, radiographic, thermographic, electromagnetic, and optic methods are employed to probe the interior microstructure and characterize the subsurface features. Good reviews on these NDT techniques can be referred in Drinkwater and Wilcox (2006) and Sposito et al. (2010). Traditionally, the areas of application are flaw detection, material characterization, and structural health condition monitoring for the purpose of dynamically measuring their property and performance over the long term. The NDT applications have been now extended to cover medical diagnosis, intelligent robotics, on-line manufacturing process control, and security screening (Ma 2012).

Historically, NDT and analysis was performed by highly experienced maintenance individuals who applied lessons learned through years of practice. Over the past decades, there has been much interest in the development and deployment of new technologies and intelligent systems for use in a wide range of fields. These techniques embrace electrical, mechanical, thermal, electromagnetic, acoustic, optical, chemical phenomena to yield monitoring data for collection and desirable instruments for processing and interpreting the measured data.

This chapter reviews and analyzes measurement and monitoring technologies for metallic material characterization, condition monitoring of insulation degradation and the associated instruments. In Section 5.2, electromagnetic sensors used for eddy current testing and imaging are presented, focusing on measurement of the electrical property of metals, the decarburization of high-carbon steels, and imaging liquid metal flow profile during continuous casting. The remaining sections then discuss condition monitoring techniques for evaluating and diagnosing high-voltage insulation deterioration of power apparatuses. In Section 5.3, the effects of multiple physical stresses on insulation degradation are discussed, alongside a model representing such degradation. In Section 5.4, partial discharge (PD) detection techniques for monitoring insulation conditions are presented by means of both nonelectrical and electrical techniques. The methods employed to infer the nature, form, and quantities of discharges activity will also be presented in this section.

5.2   ELECTROMAGNETIC NDT

For a fundamental eddy current testing, two inductive coils are usually sufficient—with one acting as an exciter and the other a detector. The excitation coil with a number of turns is excited by an alternating current (AC) source, which generates a changing magnetic field in its vicinity. This time-varying magnetic field interacts with the test sample and then induces eddy currents. The eddy currents in the sample in turn generate a secondary magnetic field, which opposes the primary magnetic field. The interaction between these two fields alters the distribution of the magnetic flux, thus resulting in an apparent change of the coil voltage. By measuring the coil voltage change, the properties such as the electrical conductivity and magnetic permeability of the material, structure of the metals, flaws, or fatigue cracks in metallic samples can be determined by using analytical and/or experimental methods.

With more inductive coils placed along the periphery of an object space, it is possible to tomographically image the distribution of materials inside a region of interest. The sensor array used can contain excitation and detection coils, which can either be dedicated to a particular function (exciter or detector) or operate in both modes. By energizing an excitation coil with an AC signal, a measurement can be obtained from detection coils. This kind of projection is continued until the last excitation coil is excited and measurement is taken. The measured data are then manipulated using appropriate mathematical inversion techniques to create an image of the internal object distribution.

5.2.1   EDDY CURRENT TESTING

Different shapes of sensor coils are normally constructed purposely to accommodate different variously shaped samples. For example, pancake-type surface coils are used to inspect plate, sheet, or irregularly shaped samples, while encircling coils are used primarily for inspecting rods, tubes, cylinders, or wire in manufacturing applications. When the area to be tested is large, pancake-type surface coils are chosen to reduce the testing time, whereas coils as small as practical are required to detect small cracks. In practical measurements, double-coil arrangements are preferred, where one coil is used for excitation with a separate secondary coil used for detection. The separation of excitation coil and detection coil can prevent thermal drift problems in the measurement signal due to overheating of the excitation coil.

Eddy currents induced inside the test sample by the applied magnetic field attenuate with depth below the test sample surface. This attenuation is mainly governed by the test object's electrical conductivity σ, magnetic permeability μ, and the applied frequency f for a given test geometry. The standard depth of penetration δ can be used to characterize this diffusion phenomenon, which, for plane geometry, is mathematically given by $\delta = 1/\sqrt{\pi f \mu \sigma}$. Apparently, high-frequency measurements give information regarding the properties adjacent to the surface, whereas low-frequency testing probes the subsurface deeper inside the test sample. Multi-frequency eddy current testing has been therefore commonly employed.

One of the applications has been to the characterization of microstructural changes of cellular metal foams. Cellular metal foams are a relatively new class of material and have increasingly been recognized due to their exceptional combination of mechanical, thermal, acoustic, electrical, and chemical properties. To identify the healthy condition of porous metals for an industrial application, evaluation of the foam electrical properties is of paramount interest, which can be achieved through nondestructive eddy current measurements. The induced eddy currents flowing in the metal foams are affected significantly by the foam properties and, consequently, measurement of the impedance change, defined as the voltage change with respect to the excitation current, on the detection coil(s) permits the metal foams to be characterized.

Figure 5.1 shows solenoidal coils designed to examine the cylindrically shaped samples (Ma, Peyton, and Zhao 2005a). It has been found that the phase-frequency response of the normalized eddy current signal of the sensor is relatively immune to the coil-to-sample spacing and fill-factor variations, from which the equivalent conductivity of the foams are measured. As an example, Figure 5.2 shows the variations of electrical conductivity of the aluminum foams with both porosity and pore size. The results show that higher porosity leads to lower equivalent electrical conductivity due to the decreasing the metal volume ratio in the foams. The pore size determines the amount of air trapped, the average wall thickness between pores, and the degree of interconnectivity between pores, thus affecting the equivalent conductivity of the foams as well. Consequentially, through monitoring of the equivalent conductivity, the mean porosity and hence health conditions of the foams can be estimated (Ma and Peyton 2006a). The sensors and technique can also be used for measurement of magnetic permeability of porous Fe samples (Ma, Peyton, and Zhao 2006b).

Decarburization is a well-known phenomenon that can change the structure and content of steel. It occurs usually when steel is heated in an environment where oxygen is present, leading to oxidation and loss of car-
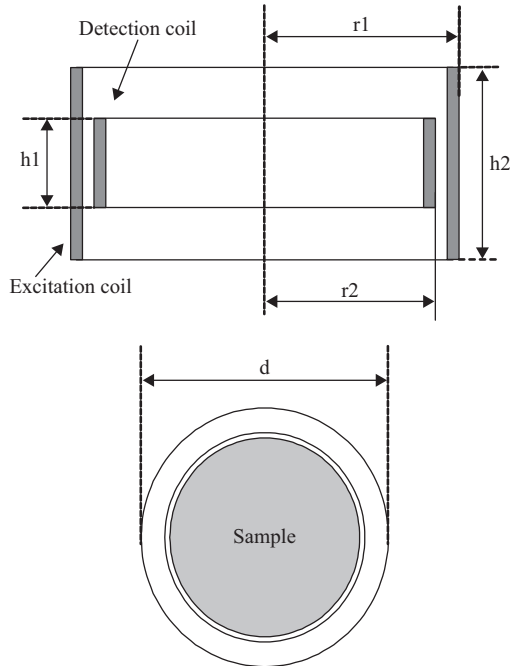
Figure 5.1. The schematic diagram of the solenoidal coils, where $r1$ and $r2$ are the outer radii of excitation and detection coils, $h1$ and $h2$ are the heights of excitation and detection coils, and d is the outer diameter of tube of which the coils are wound axially.

bon. As a result of decarburization, the steel can lose some of its strength and ductility, and bring dramatic modifications to its mechanical properties, like being vulnerable to breaking hence decreasing the fatigue lifetime of the steel. Eddy current techniques have been demonstrated to be very sensitive and well related to surface carbon content (Amiri and Kashefi 2009) and surface modifications associated with the decarburizing (Mercier et al. 2006). The measurement of three different electromagnetic properties (eddy current, tangential magnetic field, and incremental permeability) was used to evaluate decarburization profiles in spring steel (Burzic, Zamberger, and Kozeschnik 2010). The results, however, were highly influenced by the calibration procedure. A multifrequency, multi-output electromagnetic sensor was proposed with which hardness profile beneath the surface can be determined by correlating an optimum frequency to each measured output voltage (Kahrobaee and Kashefi 2011).
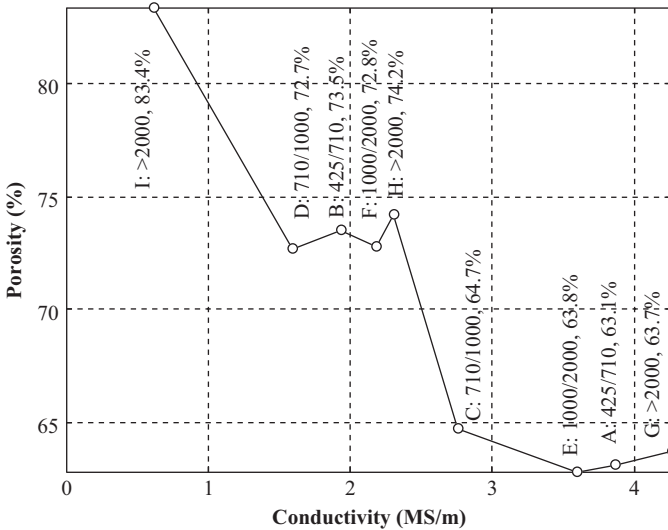
Figure 5.2. Results showing the variation of electrical conductivity with both porosity and pore size of the aluminum foams. The foam samples are categorized into groups with pore sizes of 425–710 μm, 710–1,000 μm, 1,000–2,000 μm, and larger than 2,000 μm and with porosities ranging from 63.1 percent to 83.4 percent.

Among the sensors, an H-shaped electromagnetic sensor has been used to evaluate decarburization of high-carbon steels and rail decarburization depth due to its unique design. As shown in Figure 5.3, the sensing head has a primary excitation coil and four secondary flux sensing coils. The sensor was designed with two U cores joined back to back with a common excitation coil, one of which is being used as the active sensing element and the other as a dummy reference element. Adjacent pairs of secondary pickup coils are wired in series, therefore maximizing the signal of interest while helping to reduce the common mode interference caused by ambient magnetic fields. Multi-frequency analysis is performed by exciting the sensing head with a composite waveform containing the required harmonics. The measurements are sampled and analyzed by the digital signal processor (DSP) using a fast Fourier transform (FFT) to extract the harmonic phase and magnitude information. To maximize the dynamic range of the system, a three-tier approach was used to remove the background voltage from active sensing element's output voltage. More details are seen in Dickinson et al. (2007).

Efforts have been made to measure decarburization depth using this H-shaped electromagnetic sensor. Hao et al. have tried to measure
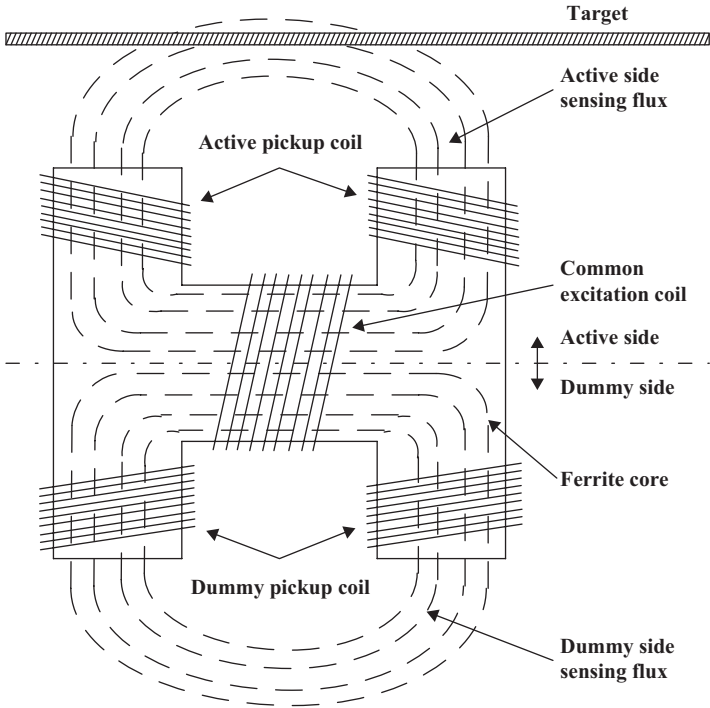
Figure 5.3. Magnetic head before encapsulation and the main elements of the magnetic sensing head.

decarburization of high-carbon steel with pearlitic microstructure, where the relationship between the sensor output and decarburized layer thickness was modeled using finite element software (Hao et al. 2008). A potential online nondestructive measurement using this EM sensor was reported in Hao et al. (2009) under favorable operation frequencies. The work demonstrates that the demagnetizing field from online samples has a significant effect on the sensor output. An H-shaped ferrite core EM sensor has also been applied to phase transformation detection in steel rolling application, where the response of the sensor can be described by a simple analytical model (Zhu, Yin, and Peyton 2011). The EM sensor was recently used to evaluate rails with different levels of decarburization due to different bloom reheat times (Zhu et al. 2012).

As one of the most extensively used NDT techniques, eddy current testing permits detection of cracks and internal discontinuities in conductive materials, for both ferromagnetic and non-ferromagnetic materials. Defects in the metallic materials will cause a change in the eddy current flow. It is this change that can be detected by the EM sensors and the

associated electronics. The advantage of the eddy current method over other techniques is that testing can be made without any direct physical contact between the sensor and the inspected material. An optimized sensor must induce the greatest eddy current density near the crack to obtain the greatest sensor response (Bavall 2002; Mook, Lange, and Koeser 2001). Another eddy current testing technique is called pulsed eddy current (PEC) testing. Compared to the conventional eddy current inspection techniques, PEC testing allows multi-frequency operation by using only a step function voltage signal to excite the coil. As a result, the electromagnetic response to several different frequencies can be measured simultaneously. Since the skin depth of penetration is dependent on the frequency of excitation, information from a range of depths can be obtained simultaneously. A good review of the state-of-the-art methods and current challenges for crack detection using PEC testing can be found in Bai et al. (2013). It has also been found that the use of data fusion methods like principal component analysis (PCA) to mine the transient PEC signals can significantly reduce the computation time, hence allowing rapid identification of the cracks (Horan, Underhill, and Krause 2013; Theodoulidis, Wang, and Tian 2012).

5.2.2 ELECTRICAL TOMOGRAPHIC METHODS

The mathematical concept of tomography was first proposed early in the 19th century. The basic aim of modern tomography is to image and determine the cross-sectional distribution of materials of interest by a set of measurements using sensors that are distributed around the periphery of a process. Despite their relatively modest image resolution, electrical tomography has opened up many applications. This is mainly because this sensing technique is nonintrusive and noninvasive; has a relatively high imaging speed, low cost, and can be used in the harsh operating conditions, for example, in the cases where temperature and pressure are extremely high.

Electrical tomography techniques can themselves be further divided into different modalities based on the measurement of the passive electrical quantities—namely, resistance (impedance), capacitance, and inductance. Electrical impedance tomography (EIT) is based on the measurement of resistance and reactance patterns and can produce conductivity and permittivity images. Electrical capacitance tomography (ECT) involves the measurement of capacitance profiles and generates images of permittivity distributions. Electromagnetic inductance tomography (EMT), or
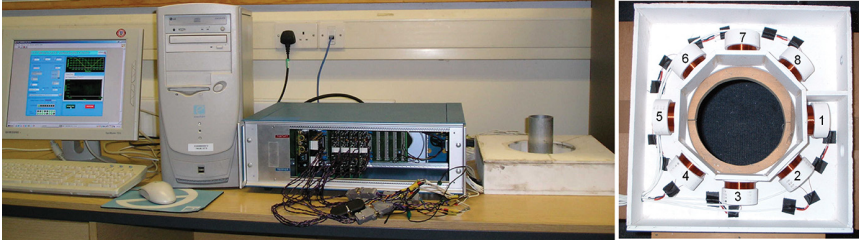
Figure 5.4. Photograph of the EMT system developed by Lancaster University, from right to left: Inductive sensor array, conditioning electronics unit, and PC installed with a data acquisition system.

magnetic induction tomography (MIT), employs mutual inductance measurements between inductive sensors to provide images that represent the conductivity and permeability distributions of the material. The applications of ECT include the measurement of multiple-phase flows in pipes (Gamio et al. 2005) and solid concentration in an industry process (Sun et al. 2008). The EIT techniques have been widely used for medical imaging and geophysical investigations, and also for crack identification in electrically conducting materials (Lazarovitch, Rittel, and Bucher 2002) and air-bubble detection in pipelines (Ijaz et al. 2008).

As shown in Figure 5.4, a typical EMT system consists of (i) an array of inductive coils arranged around the object periphery being tested, (ii) a custom conditioning electronics unit and a data acquisition system, and (iii) a host computer running image reconstruction algorithms. An industrial case study of this system was to image molten steel flow and solidification. The control of steel flow through a submerged entry nozzle during continuous casting is critical to ensure steel cleanness and surface quality. Existing methods of assessing steel flow in the nozzle and mold are based on mathematical and physical modeling. Despite these studies, there is no real method of determining the actual conditions within the nozzle, which is fundamentally an opaque tube. The imaging system has been successfully used to monitor molten steel flow conditions using real data acquired during continuous casting at Corus (Ma et al. 2005b; Ma, Peyton, and Higson 2006c; Ma et al. 2008). An example of image results is shown in Figure 5.5. The first and second rows in the figures show the steel flow was initially positioned in the middle of the nozzle. The third row shows the steel flow then splashed on to the sidewalls of the nozzle. Images in the last row show the deposition of steel remnants on the inside of the nozzle wall after the pouring.
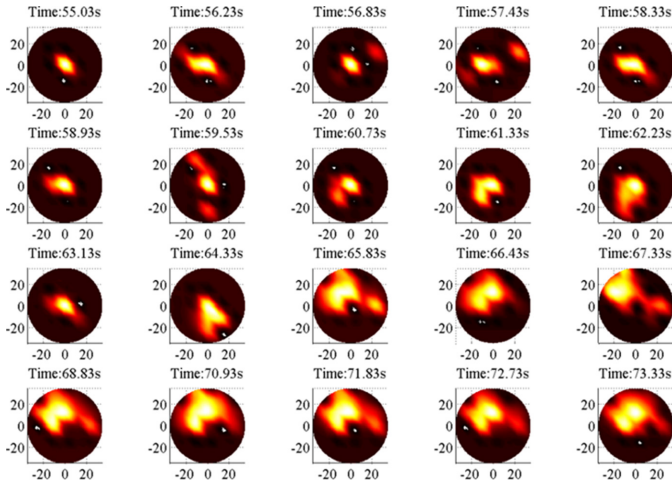
Figure 5.5. EMT images of steel flow profiles through the nozzle at different time instants.

In addition, the planar electromagnetic tomography systems were proposed for conductivity measurements (Riedel et al. 2004) and detection of crack faults on thin metallic plates (Yin 2006). The proposed systems differ from traditional EMT systems in that the planar sensors are more strongly coupled through samples rather than directly. The sensors can be placed above the plate under inspection, which is often the case where only one surface of the testing plate is accessible for many NDTs. This sensor arrangement also facilitates their implementation as pancake coils or flat coils manufactured on the printed circuit board (PCB).

For electrical tomographic methods, the measurement sensors, such as metallic electrodes used in ECT and inductive coils placed in EMT, should be sufficiently large compared to the object region being tested to give a measureable change in terms of the relevant electrical quantity. This means that not many sensors are used and 8 or 12 sensors are common in these sensing modalities. An N-sensor system can only provide $N(N-1)/2$ independent measurements. This means that the techniques provide relatively low-resolution images in contrast to other conventional tomographic methods, such as X-rays, in which high-resolution images are usually formed of slices of a material. However, X-rays suffer from the inherent hazards of radiation and size of the equipment. As highlighted above, the use of electrical tomography techniques for fault detection and process visualization are fast and relatively inexpensive. Furthermore, due to the flexible sensor design and the significant advantages of measurement principle, there

are strong possibilities of using these techniques in the emerging area of great importance, such as monitoring of CO_2 migration in carbon capture and storage (Hattenberger et al. 2013) and security checking body scanner (Marsh et al. 2013).

5.3 INSULATION DEGRADATION

5.3.1 MULTIPLE PHYSICAL STRESSES

Deterioration of insulation systems is subject to thermal, mechanical, and electrical stressing, or their combinations. Thermal ageing in electrical machines is accelerated under impulse conditions including overvoltage, switch transients, and lightning strikes, under excessively high flux densities or eddy currents and under overload condition. Thermal ageing causes molecular decomposition and oxidation of the insulating materials, resulting in delamination, cracking, embrittlement, and depolymerization of the insulating materials (Simons 1980). Thermal ageing is enhanced by the different thermal expansion coefficients for copper conductor and the insulation, yielding a considerable shear stress between the conductor and the insulation.

Thermal ageing in transformers causes degradation of the solid paper dielectrics by deteriorating the molecular bonds in the cellulose chains from which it is made, and degradation of oil by dissolving various chemical gases depending on the conditions inside the transformer such as the temperature and oil. Degradation of the oil insulation not only serves the chemical degradation of oil itself but also increases the number of impurities and eventually produces bubbles in oil.

Although thermal ageing has little impact on the voltage breakdown compared with mechanical and electrical stress, it accelerates the deterioration of the insulation. Borsi investigated the relation between temperature changes during thermal ageing and PD behavior of epoxy–resin transformers (more details about PDs are given in the subsequent sections). The work concluded that the PD inception voltage of coils will drop to ~50 percent in a temperature range from −30°C to 140°C (Borsi 1993). The reduction of PD inception voltage is caused by significant increase in the relative permittivity of insulating materials with temperature, and by contrast, by decrease in the relative gas density with temperature in the cylindrical voids existing in the insulation.

Mechanical stress results from thermomechanical periods when electrical machines are in start or stop operation or in shocking load period, and from electromagnetic forces. It causes bar vibration producing fretting

between copper, insulation, and core, thus deteriorating the insulation integrity. Kimura and Kaneda (1995) observed the development of defects under mechanical stress by scanning electrode microscope and found that the delamination between insulating materials, that is, between mica and epoxy, occurs first, then cracks follow, thus lowering the breakdown voltage. Normally, the bars inside the machine can withstand a magnitude of mechanical force up to 10 kg cm^{-1}.

Localized electrical stress is enhanced by transient pulsing from switch surges and system disturbances such as direct-online starting of machines and tripping of circuit breakers on overload. For transformer windings, typical nominal electrical stress is within 1.5 kV mm^{-1}, while the electrical stress of stator windings varies between 2 and ~2.5 kV mm^{-1} (Mayoux 2000).

As mentioned previously, thermal and mechanical stresses cause defects. These defects lead to localized high electrical stresses being formed around them, thus resulting in PDs. The interaction between thermal, mechanical, and electrical stress in turn leads to increased PD activity, causing further deterioration and eventual failure of the insulating systems. Once the degradation is prevalent, PDs then represent the dominant mechanism of degradation. It is therefore imperative that PDs in such systems be detected and quantified at an early stage in their evolution to correctly assess the integrity of insulation systems.

5.3.2 A DIELECTRIC DISCHARGE MODEL

As the name implies, PD is the electrical discharge that involves only a portion of the dielectric between two electrodes and does not completely bridge the gap (Danikas 1993). In general, the principal discharge types are internal discharges, surface discharges, and corona discharges. Internal discharges occur in gas-filled and oil-filled cavities. Surface discharges appear in gases or in oil in the cases where a strong stress component parallel to the dielectric surface exists. Corona discharges appear around sharp points in gases and in liquids, subject to a high voltage. Although the magnitude of the aforementioned discharges is usually small, they cause progressive deterioration and lead to ultimate failure. It is therefore essential to analyze and understand the mechanisms of PDs.

The model for a discharging air-filled void in a dielectric has been found to be similar to a void in a solid dielectric between two metal electrodes. The physics of the recurrence of discharges under an externally applied sinusoidal AC voltage can be explained using Figure 5.6
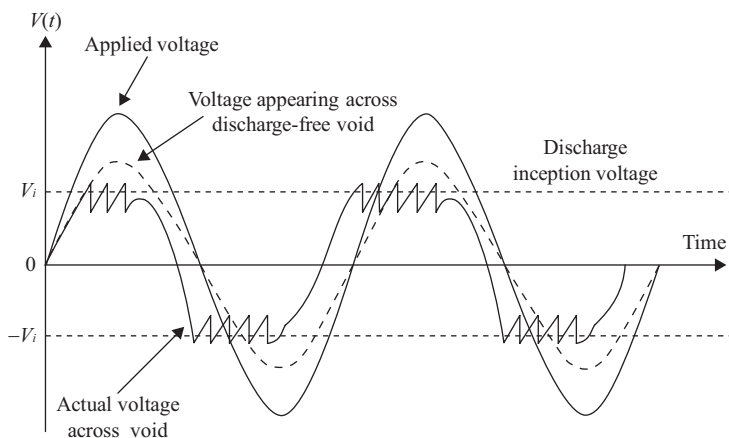
Figure 5.6. Voltages and current pulses at a void or cavity in the dielectric material.

(Arora and Mosch 2011). When the voltage $V(t)$ is sufficiently low, no discharges occur. As $V(t)$ is gradually raised to the breakdown voltage V_i, also known as discharge inception voltage, a discharge begins. The discharge will be self-extinguished as the charge initially transferred across the void builds on the void surfaces and reduces the electric field. Further discharges occur when the decrease in field is counteracted by a continuing rise in voltage until the peak of $V(t)$, after which the field will be dropping within the remaining period of this half cycle. A similar discharge sequence occurs during the negative half cycle. However, due to the impact of the accumulated charge from the positive half cycle and the applied voltage $V(t)$, the electric field will be intensified in the void during the negative half cycle, thus accelerating the breakdown that occurs earlier on the applied wave. This discharge sequence happens continuously with the applied voltage $V(t)$.

A three-capacitor model has been widely used to model a discharge resulting from a gas-filled void inside solid or fluid dielectric materials, as shown in Figure 5.7. Electric fields within the void are represented by C_c and those starting at the void walls toward the dielectric materials in series with the void form the equivalent capacitance C_b. All fields outside the void, that is, in parallel with the void, are represented by the equivalent capacitance C_a. Normally, $C_a \gg C_c \gg C_b$ for realistic geometric dimensions.

The charge accumulated at the measuring leads is referred to as the apparent charge q transferred from the actual charge in a void. The apparent
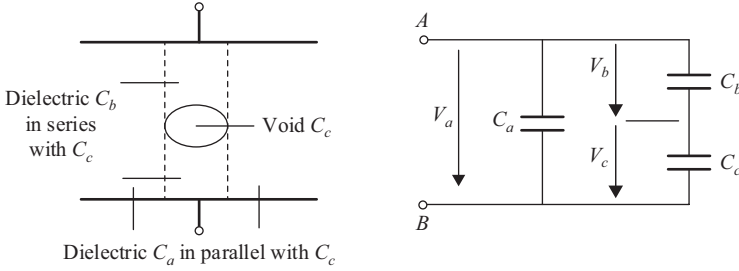
Figure 5.7. A three-capacitor model of discharge in a void surrounding by dielectric.

charge q results in a voltage drop ΔV_a across the terminals A and B of the test specimen with the following relationship between them.

$$\Delta V_a = C_b \Delta V_c / (C_a + C_b) = q / (C_a + C_b) \tag{5.1}$$

where $q = C_b \Delta V_c$, ΔV_c is the voltage drop across the void. It has been found that there is a direct relationship between the apparent charge q and the energy involved in the discharge, which causes the deterioration of the dielectric, and a relationship between q and the size of the void. Therefore, the apparent charge q is chosen as a measure for discharges, which can readily be measured with electrical discharge detectors through measuring ΔV_a at external leads of the test specimen. The apparent charge q is measured typically in picocoulomb pC. However, C_a and C_b need to be known in order to calculate q.

5.4   PD MEASUREMENT

5.4.1   NONELECTRICAL DETECTION

- Thermography techniques
  Given that PDs are generally "hotter" than their surrounding media, it might reasonably be thought that the thermal imaging techniques could be applied in their detection and measurement. Unfortunately, most PDs are enclosed in some way, for example, within solid insulation or within metal-clad enclosures. As the relatively low temperatures of PDs and the high thermal impedances are likely present between the discharges and the imaging technique, the use of this type of testing is very limited. These techniques are

only applied in the cases where the discharges are external to associated plant items, for example, on overhead lines/busbars or post/string-type insulators. However, such techniques are carried out not in any quantitative way.

• Chemical detection

PD activity can cause chemical reactions in insulating media. In power transformers, PDs manifest themselves in discharge-generated gases in the oil insulation and in degradation bi-products in paper insulation.

Under the action of PDs, oil degrades through bond scission to form characteristic gases dissolved in the oil. Typical gases produced are hydrogen, methane, ethane, ethylene, and acetylene. The quantity and types of gases produced depend on the nature of PD, its severity, and the associated temperature. By sampling oil in the transformer and further analyzing the types of the fault-generated gases and their corresponding constituents, dissolved gas analysis (DGA) can be utilized to infer PD activity in the transformers (Lelekakis et al. 2011). The DGA technique might be the simplest and cheapest technique available for online testing of transformers and has been used as the primary indicator of malfunctions leading to overheating, arcing, and PD.

Several interpretation techniques have been developed and are used in the interpretation of dissolved gases in oil. Although the presence of PD in oil-filled plant item can be inferred from the absolute levels of the measured different gases, PD is indicated primarily through the ratios of these gases. Gas ratios have been utilized to differentiate between fault types since 1970, typical example of which is the Rogers' ratios. Rogers (1978) used the order of gas evolution to form the gas ratios of methane/hydrogen, ethane/methane, ethylene/ethane, and acetylene/ethylene. A diagnosis table was created based on nearly 10,000 DGA results, together with examination of the failed units and the units with suspected faults. The table went through several evolutions and was produced in two formats; in the first, diagnosis was based on codes generated by ratios, while, in the second, diagnosis was made based on the value of the ratio. Different types of PD can be identified from the gas ratios.

However, the diagnosis obtained using gas-ratio methods was often not comparable between techniques. Other additional graphical techniques were, therefore, introduced to complement the interpretation of dissolved gases in oil. Duval (1989) developed a triangle based on the relative percentage of methane, ethylene, and

acetylene gas. The triangle is classified into six regions representing high-energy arcing, low-energy arcing, corona discharge, and hot spots. The triangle was used for fault diagnosis in conjunction with individual levels of gases rather than the gas ratios. Mostly recently, as have been expected, new intelligent techniques are finding increasing applications in interpreting DGA results, including genetic programming (Shintemirov, Tang, and Wu 2009), fuzzy logic (Abu-Siada, Hmood, and Islam 2013), and neural network (Yang and Hu 2013).

For the detection of paper insulation, high-performance liquid chromatography (HPLC) was developed to evaluate the insulation integrity in transformers by analyzing degradation by-products. Degradation of the paper under fault conditions results in cellulose chain scission, thus yielding glucose and degraded forms of glucose. Monitoring the glucose levels in the impregnated oil may indicate the extent of paper degradation. The PD activity can be inferred using this technique by the analysis of subsequent decomposition products under glucose degradation. HPLC is now being widely used in field analysis (Andrews et al. 2006; Li, Chen, and Gao 2012).

• Acoustic detection
Acoustic detection is performed by detecting the acoustic wave that initiates at the discharge site and then propagates to detection points through the insulation media. Therefore, the transducers, for example, piezo-electric sensors, are normally attached to the outer surface of plant items under measurement to acquire the wave signal. The intensity of the emitted acoustic waves is proportional to the energy released in the discharge. On this basis, the amplitude of the wave is proportional to the square root of the energy of the discharge and there should be a linear relationship between discharge magnitude and acoustic signal. However, acoustic measurements are mainly used to detect the presence of discharges and locate the discharges within the plant items.

An advantage of acoustical detection is its immunity to electromagnetic noise. However, the wave signal after propagation is heavily distorted due to a range of factors including pathway diversion, frequency-dependent velocity effects, and transmission losses, hence hampering the use of the data for directly inferring the nature, form, and quantities of discharges activity. The remarkable contribution of this technique to PD detection is its ability to locate PDs. For example, using a minimum of three sensors fixed to

the earthed tank of a large transformer and a reference signal, it is possible to determine both the presence and the accurate location of discharges. By measuring the relative times of arrival of the pulses from the discharging site at the three sensors, respectively, and by assuming a constant velocity of acoustic propagation through the transformer geometry, the relative distance from each of the sensors to the discharging site can be computed and triangulated. The details regarding the fundamentals of this approach can be seen in Lundgaard (1992). Recently, this approach has been proven to be extremely effective in locating discharging sources occurring in transformers in conjunction with digital processing techniques such as the wavelet analysis (Markalous, Tenbohlen, and Feser 2008) and pattern recognition method (Kraetge et al. 2013).

• Optical detection

  Optical analysis of the light emitted by the discharge enables location of the exact source of a discharge and hence the area where deterioration of the dielectric takes place. The work using optical techniques in PD analysis was placed on relating typical optical images to typical discharge pulse shapes through simultaneous recordings of both electrical and optical signals produced by the discharges in cavities. The optical observation can be utilized as a means to identify single discharges due to different spatial distributions being generated across the surface of the cavity by different types of discharges. However, the direct use of this technique is still restricted in the laboratory measurements mainly due to the cost of the measurement set-up, low optical sensitivity, and the obscurity of most apparatus under test to light (Biswas et al. 2012; Fracz 2013).

5.4.2 ELECTRICAL DETECTION

Electrical methods have been developed to detect the presence of PD activity, the PD intensity, PD source location, and the PD pulse with regard to the voltage phase angle. Radio noise measurements in mV amplitude on high-voltage equipment produced the first PD-related information. This approach was a measure of integrated PD activity. Therefore, it could not provide any information on PD events, PD sequence in time, and its effects on insulation.

An ideal insulation system should behave as a perfect capacitor, that is, there should only be capacitive current flowing through the system. However, a small component of resistive current is generated if insulation

integrity of the system begins to lose. As losses increase, the resistive current becomes larger and an angle δ created between the capacitive current and the resultant resistive current increases, as does the tan δ. By monitoring the tan δ of the system, relative discharge activity can be inferred. Traditionally, the tan δ measurement employed a Schering Bridge type circuit connected across the high-voltage terminals of the plant item, but latterly, instruments have made direct measurements of the different current components. The tan δ measurement is dependent on the capacitance and resistance values in the equivalent circuit of test samples, the type of equivalent circuit used, and the frequency of the external test voltage (Kuffel, Zaengl, and Kuffel 2000). Any sudden change in tan δ would be considered indicative of PD inception and the rate of change with increasing voltage would be indicative of the relative severity of the PD. Because tan δ measures the integrated quantities associated with the dielectric loss, this approach is only capable of indicating the integrity of the insulating system degradation of the specimens under test. It cannot distinguish between a few large discharges that might be detrimental and a lot of small ones that might be quite innocuous. Therefore, most commercial instruments have been developed to incorporate a peak magnitude detector to deal with this problem.

One of the instruments is the partial discharge analyzer (PDA) developed by the Canadian Ontario Hydro. For rotating machines, a capacitance coupler can be permanently connected to the machine phase terminals to allow winding discharge measurements. In large hydraulic generators, two coupling capacitors, each having a typical capacitance of 80 pF, are required to be permanently installed at the ends of the ring bus of each winding phase, respectively. The PDA instrument (Stone et al. 1990) can take the signals from the pair of couplers and sort the PD pulses by removing external noise by means of differential amplifier, thus making the analysis of discharges in stator windings more reliable. The lengths of the coaxial cables connecting the couplers at the ends in the same phase to the terminals of PDA instrument are purposely designed to ensure that the electrical distance from the machine terminal along the ring bus to the instrument terminal is the same for the two couplers. The connection scheme is shown in Figure 5.8, where two identical capacitive couplers C_1 and C_2 (80 pf) are permanently installed to capture signals. To ensure that the test presents an equivalent electrical path to the inputs of the differential amplifier, the coaxial cables on both sides are designed with such a specific relationship

$$a + \frac{b}{0.65} = c + \frac{d}{0.65} \tag{5.2}$$
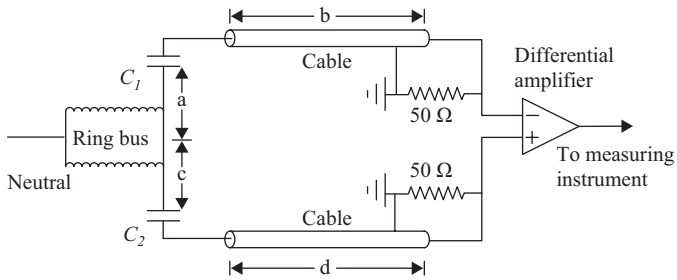
Figure 5.8. Measurement arrangement for rejection of common-mode electrical noise for generator PD detection.

The coefficient 0.65 is referred to as the ratio of the pulse propagation velocity in the coaxial cable to that in the vacuum. With this arrangement, extraneous noise to the generator is rejected as it arrives almost simultaneously (within 1.5 ns) at the inputs of the differential amplifier. However, a discharge within the winding yields a nonzero output as the electrical signal arrives at the inputs of the differential amplifier with a sufficiently separated time interval, thus being recognized and then sorted for further digital analysis.

Due to high levels of noise appearing on the site and also the very compact structure of the circuit ring bus in the turbine generators, the PDA technique is not applicable to discharge detection in turbine generators, where a new coupling device has been developed. This new coupling device, known as the stator slot coupler (SSC) developed by the Canadian IRIS Power, is a form of ultrawide-band directional electromagnetic coupler in nature, which was designed to be installed in the slot region of the stator core during manufacturing. The SSC has a very broad frequency band extending from 10 MHz to >1 GHz. Although PD signals and noise are detected by this coupler simultaneously, they can be distinguished by pulse shape recognition through modern digital electronics, enhancing the coupler's sensitivity to PD. PD signals manifest themselves on the coupler with high-frequency components as they occur close to the coupler, while the external noise manifests itself on the coupler with low-frequency components due to attenuation effects of high-frequency components when noise propagates some distance to reach the coupler. For a large turbo generator measurement, six SSC couplers are needed in order to cover most of the highly stressed portion of the winding, thus increasing the fault probability of the generators during operation.

A discharge source can also produce an electromagnetic wave that propagates away from the PD site and travels toward the earthed metal

enclosure of the plant and further to the atmosphere outside the switchboard through a gap somewhere in the enclosure. A transient earth voltage (TEV) is produced by the action of the wave connecting with the earthed metalwork and can be detected by a capacitive probe (Brown 1996). This method gives an indication of the PD severity with the magnitude of a detected signal, which lies normally in the milli-volts to volts range. With multiple capacitive probes used, it is possible to gain some knowledge of location using the time of flight principle, as the probe nearest to the source should detect the discharge first. The attractive advantage of this type of testing is its nonintrusive nature, that is, no disruption or outage is required to the plant under test. However, as with other types of PD measurement, the measurement readings of using TEV principle suffer readily from external electromagnetic noise. It is therefore essential that a background survey should be completed prior to measurements so that discharges from the plant or from external sources can be differentiated.

Antenna is another type of technique that could be used for on-line monitoring of the PD severity of the plant (Kaneko et al. 2009). An antenna measures the radiated radio interference generated by discharges based on the principle that the frequency band of an antenna can be tuned to a specific radio frequency bandwidth to best respond to the discharges. As with the TEV method, antenna techniques are of nonintrusive nature. However, PD signals, when measured with this type of testing, are also highly susceptible to noise from environments in substation. Therefore, an effective noise reduction technique is desirable when the antenna technique is applied. As a direct connection to high-voltage terminals is not required, the PD measurements using the antennas can be made and carried out without planned machine outage.

Current transformer (CT) has been used in particular for motors and, to a lesser extent, for transformers as a "first pass" technique by engineers. By connecting a clamp-on CT to the neutral strap of the plant item, the output readings can be taken to an oscilloscope or a frequency spectrum analyzer for further PD analysis. The use of the CT approach is extremely cheap, simple, and safe, and no disconnections need to be made during test. Unfortunately, this type of testing suffers from the inherent disadvantages mainly including the lack of an effective calibration method to determine the magnitude of any present discharges, susceptibility to extraneous interference, for example, pulses from power electronics circuitry and corona discharging from somewhere in the system, and no effective phase information on the location of discharges with respect to the AC voltage power cycle. In addition, there may not always be a neutral available on the plant item under test.

Rogowski coils have been used for the detection and measurement of PD currents for decades (Hashmi, Lehtonen, and Nordman 2011; Zhang, Xiao, and Li 2009). They work on Ampere's law and are used to measure currents via the magnetic field produced by these currents. An air-cored coil is placed around the conductor in a toroidal fashion. The current flowing through the conductor produces an alternating magnetic field around the conductor, resulting in a voltage being induced in the coil. The rate of change in this voltage is proportional to the rate of change in current. This voltage is then integrated electronically to provide an output proportional to the current. Rogowski coils are normally connected to the high-voltage terminals of the plant items to measure the PD current output. In use, no direct connection to the high-voltage terminals and no coupling capacitor are required. However, this detection method does not work through shields.

Another approach connecting the discharge transducer to the high-voltage terminals of the plant item is the use of a capacitor divider type assembly (Haddad and Warne 2007). Typically, this discharge transducer consists of a discharge-free high-voltage capacitor connected in series with or in parallel with a low-voltage impedance circuit (RC or RLC), which in turn is connected to an oscilloscope or similar data acquisition instrument. The high voltage can be reduced to a safe level at the low-voltage impedance (typically 1,000:1 ratio) by careful choice of component values. Individual pulses from discharges can be displayed superimposed on the AC power cycle voltage.

5.4.3   QUANTIFICATION AND CHARACTERIZATION

A digital PD measurement system can register the discharge magnitude and the discharge position in relation to an AC power cycle. The basic quantities that have been used to characterize the recurrence of discharges are (i) the magnitude q_i of the apparent charge of an individual dischargen in pC, (ii) the phase position φ_i in degrees in a power cycle when the discharge q_i occurs, and (iii) the instantaneously applied voltage v_i normally in kV at which the discharge q_i has occurred. Figure 5.9 shows the PD pulse sequences and the parameters used to characterize the recurrence of discharges in a complete AC power cycle of 50 Hz.

The analysis and interpretation of measured data has been a difficult task. Different PD measurement systems have been developed to allow the detection and recording of PDs. In particular, the phase-resolved methods have become very popular and widely used for the pattern recognition of
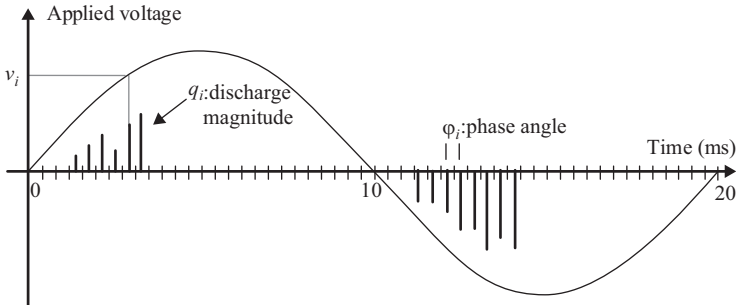
Figure 5.9. The PD pulse sequences and the relative parameters used to characterize the recurrence of discharges.
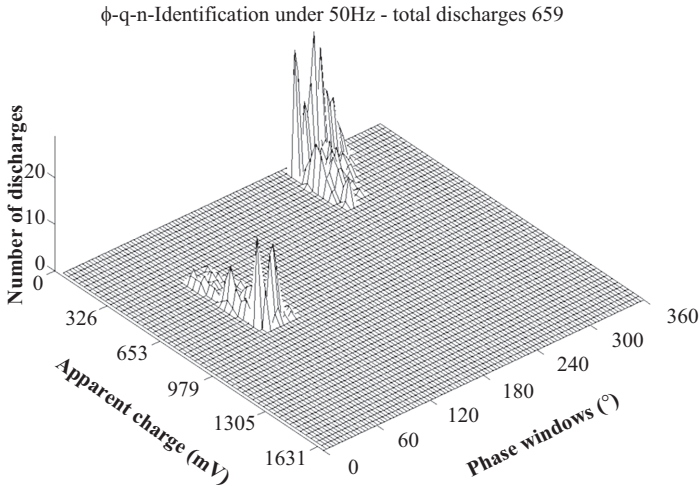
Figure 5.10. An φ–q–n pattern as observed in a point–plane discharging configuration.

PD sources as measured with PD detectors. This is because the approach can correlate those physical parameters at the applied voltage, as shown in Figure 5.10, and therefore renders a characterization of the discharge pattern by physical parameters.

The idea behind the phase-resolved methods is first to divide the power frequency cycle into a number of small phase windows and then calculate the additional discharge quantities with respect to each phase window, and finally plot them as a function of the phase windows φ. The commonly used univariate distribution are φ–n, φ–q_a, and φ-q_p, where n is

the number of the discharges, q_a the average discharge magnitude, and q_p the peak discharge magnitude. The quantities n, q_a, and q_p are all measured in each phase window within test recording period T. Based on these, the so-called φ–q–n pattern can be developed and considered as the most complete form of a graphical PD representation because φ–q–n patterns provide a three-dimensional visualization realized by a two-dimensional array, where n is the number of identical or similar discharges recorded within each phase interval. One example as observed in a point–plane test configuration is shown in Figure 5.10, where the point electrode was connected to the high-voltage terminal (Ma, Zhou, and Kemp 2004). In this example, the test circuit utilized is a conventional narrowband detector with a resonant frequency of 260 kHz under a sampling frequency of 2.5 MHz, and a recording duration of 10 minutes. The PD patterns clearly show that a large number of discharges with small magnitude appear in the negative half cycle, whereas a small number of discharges with large magnitude appear in the positive half cycle.

The φ–q–n patterns within an adequate recording duration have been utilized to identify the origin of the PD source. An extensive investigation was presented in Hudon and Belec (2005) in order to determine the phase-resolved PD pattern of several types of discharge sources including internal discharges, slot discharges, end winding discharges, and delamination discharges in modern rotating machine insulation. It was shown that the use of specific patterns of its PD source can serve to allow a better diagnosis of the generators. Such discharge patterns have also recently been used to detect, identify, and locate internal and surface discharge sources in oil–paper insulation systems (Rubio-Serrano et al. 2012). It was proved that different PD sources in power transformers can be better identified by means of the energy content in different frequency bands.

In contrast, in an ultrawide-band PD measurement, the recorded signals are interpreted using another kind of 3D display that characterizes all the discharge pulses occurring within a very limited time duration of, for example, one second due to the extremely high sampling frequency utilized, usually >1 GHz (Pearson et al. 1995; Rudd, Mcarthur, and Judd 2010). Unlike Figure 5.10, the PD patterns in such a measurement are represented by the cycle number in x-axis, phase angle at which PD pulses occur with respect to the AC power cycle in y-axis, and amplitude in z-axis. This type of phase-resolved method has been widely used alongside the ultra-high frequency (UHF) measurements to detect electromagnetic waves produced by PD pulses with frequencies that extend well into the UHF range under a number of different configurations such

as for gas-insulated switchgears (GIS) (Gao, Ding, and Liu 2011) and power transformers (Beltle, Muller, and Tenbohlen 2012). However, the phase-resolved PD methods are normally employed at field sites where the phase information is available, which is not always the case, for example, the PD patterns of free-moving conducting particles in GIS. A chaotic analysis of PD was proposed in a study by Koo et al. (2010) to identify the type of defects by means of PD patterns without employing the phase information.

5.4.4 NOISE ISSUES AND DENOISING

Measurements of PD magnitudes are often hampered by electrical noise. According to IEC60270 (IEC 2000), the disturbances that might corrupt the indication of PD reading during off-line PD measurements fall into two categories, those which occur even if the test circuit is not energized and those which occur when the circuit is energized. The former disturbance sources include inherent thermal noise of the measuring instrument itself and externally induced environment noise caused, for example, by switching operations in other circuits, high-voltage tests in the vicinity and radio transmissions, and so forth. The latter disturbances may be caused, for example, by sparking due to the imperfectly earthed objects in the vicinity and imperfect screen connections in the test set-up or by the higher harmonics contained in the applied test voltage, and so forth. For on-line PD measurements, except for the aforementioned possible noise sources, the disturbances are still possibly caused by operation of the power system including the communicator operation, and thyristor firing generated inside electrical machines, and so forth.

Typical methods utilized to reduce noise include the use of specially designed measurement circuitries and digital signal processing methods. A balanced measurement circuit has been recommended in [IEC2000] owing to its advantage of suppressing external interference, such as discharges coming from the high-voltage source and the high-voltage leads within the test circuit. Another circuit capable of discriminating polarity, as shown in Figure 5.8, can be used to eliminate common-mode electrical noise in PD detection for generators. PD pulses of opposite polarity are acquired at the two arms of the test circuit when a PD appears within the test object, whereas PD pulses of same polarity are captured when a PD appears outside the test circuit. Consequently, only those PD events originating from the test object are accepted and quantified, whereas those PD events having "common mode," that is, same polarities, caused by the

discharges occurring outside the test circuit are rejected. The other benefit associated with this system is its capability of suppressing some internal noise, such as static exciter transients, because this type of noise has a relatively low-frequency range and its traveling time suffers little change due to difference in signal path lengths.

Many researches have been reported on denoising of PD signals using digital processing methods, especially for on-site PD measurements. Traditionally, the techniques used for digital signal processing are realized in either the time or frequency domain. In the frequency domain, denoising has been widely implemented by FFT coefficients to a certain threshold for different frequency bandwidths (Ma, Zhou, and Kemp 2000; Sriram et al. 2005). However, the FFT does not provide information on time, such as the time at which changes in frequency occur. The windowed Fourier transform, also known as the short-time Fourier transform (STFT), in attempting to overcome this deficiency, can provide a two-dimensional representation of a signal in the time-frequency domain through windowing the signal. However, resolution in both time and frequency remains constant due to the same window being employed in STFT across the entire frequency range. With regard to the PD signals, there always exist nonperiodic and fast transient features, which tend to be ignored and cannot be revealed efficiently and explicitly. For these reasons, the applications of Fourier transform method to denoising PD signals are limited. Adaptive digital filters have been used in PD measurements directly in the frequency domain for suppressing narrow-band sinusoidal interference (Kopf and Feser 1995; Sher et al. 1995). However, white noise still continues to remain in the PD signals following digital filtering of the given signals.

More recently, the wavelet-based denoising techniques have been found to be more promising for detecting PD signals buried in noise. As we know, the objective of denoising is to remove noise from the PD signals as effectively as possible while preserving the signal features essential to the application. The wavelet transforms can decompose/reconstruct a given signal in varying scales and characterize the signal in both the time and frequency domain simultaneously with variable resolution. Wavelet-based denoising is accomplished by selecting a wavelet that can be scaled to represent effectively the signal features, performing a decomposition of the signal using the selected wavelet, thresholding the wavelet coefficients to eliminate those associated with noise while preserving those that are required to describe the signal of interest, and then reconstructing the denoised signal. For this reason, it is particularly suitable for the analysis of transient, irregular, and non-periodic signals such as in the case of

PD signals (Dey et al. 2010; Ma, Zhou, and Kemp 2002a; Ma, Zhou, and Kemp 2002b). As an example, Figure 5.11(a) gives the wavelet denoising results of a practical PD data set where PD signals at levels comparative to and below noise can be fully extracted. Figure 5.11(b) shows the PD pattern that is dominated by noise when raw data are directly used, whereas Figure 5.11(c) shows the PD pattern based on the wavelet-based noise suppression. The wavelet denoising technique has both eliminated noise and detected more PD activity with low apparent discharge magnitude, and is thus much more effective in characterizing the nature of the discharges.

However, denoised signals using wavelet methods can achieve good results if the PD pulses are residing at different resolution levels with noise and proper selection of mother wavelet, the number of decomposition levels, and threshold value. It can be therefore envisaged that the method may remove some useful information related to PD pulses if these conditions are not fulfilled. Adaptive noise cancelation methods were used for preventing degradation of the denoised PD signals. For example, the method proposed in Kopf and Feser (1995) defines and weights the disturbing frequency components involved in the measured PD signal and then remove them. Although the proposed filters are self-configurable, they are normally realized by a high-order finite impulse response (FIR) filter and thus are time consuming. An integration of artificial intelligence methods with the wavelet denoising method has been proving a more robust approach for PD detection and characterization. The applications include the use of neural networks or neuro-fuzzy networks for pattern recognition to automatically classify PD sources (Ma et al. 2013; Mazzetti et al. 2006).

5.5   CONCLUSION

Electromagnetic sensors and related instruments have been effectively used to examine the electrical properties and further the health condition of the metallic materials. Different sensors need to be designed to accommodate the types of sample shapes. The appropriate selection of an effective operation frequency range is critical in achieving the reliable results, which depends mainly on the coil configuration and the properties of the sample. The lower operating frequency should be selected at which electromagnetic skin depth is comparable with the thickness or radius of the sample to ensure that a sufficient volume of material is probed. The upper operating frequency should be selected to guarantee useful information regarding the properties adjacent to the surface. With more inductive coils, it is possible to tomographically image the distribution of materials inside a region of interest with a suitable inverse reconstruction algorithm.

(a)

(b)

(c)
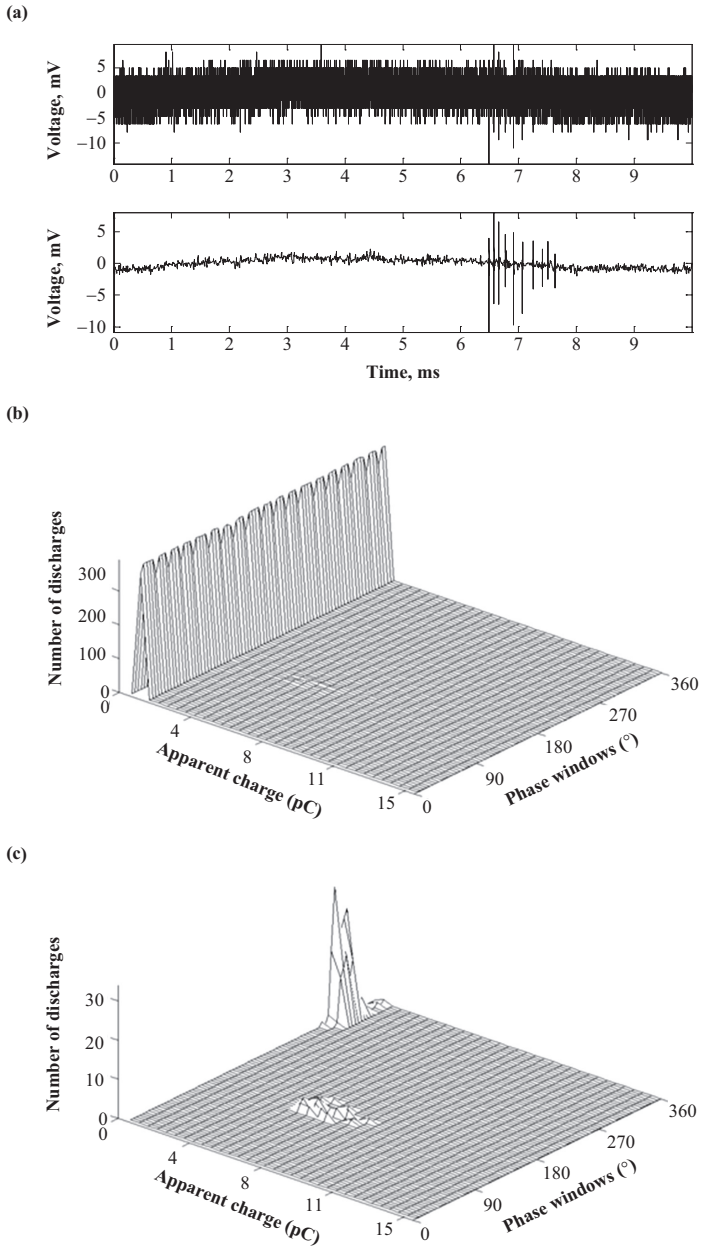
Figure 5.11. (a) Raw data and denoised signal using wavelet, (b) PD pattern of original data, and (c) PD pattern with wavelet-based noise suppression.

Deterioration of insulation systems is particularly subject to thermal, mechanical, and electrical stressing, or their combinations. Once the degradation is prevalent, a phenomenon called PD then represents the dominant mechanism of degradation. The electrical methods capable of measuring individual discharge pulses have been widely used for PD feature extraction and pattern recognition. In practical measurements, PD pulse as acquired by detectors is normally distorted and deformed due to the complex path it passes through such as the windings in transformer and motor stator. Reduction methods of interference are critical for accurate analysis of the discharges and further reliable pattern identification. The recognition to wave-shapes of the discharge pulses and the difference between PD signal characteristics and noise will be greatly helpful for the noise suppression in order to increase the PD signal-to-noise ratio significantly through digital processing methods.

The monitoring methods presented in this chapter have been demonstrated to be able to provide information on the nature, form, and severity of the degradations, and thus minimizing the risk of unexpected failure. It is believed that the reliability and availability of future industrial processes and advanced electrical power systems can be increased by the application of improved eddy current testing and condition monitoring techniques.

REFERENCES

Abu-Siada, A., S. Hmood, and S. Islam. 2013. "New Fuzzy Logic Approach for Consistent Interpretation of Dissolved Gas-in-oil Analysis." *IEEE Transactions on Dielectrics and Electrical Insulation* 20, no. 6, pp. 2343–49. doi: http://dx.doi.org/10.1109/tdei.2013.6678888

Amiri, M.S. and M. Kashefi. October 2009. "Application of Eddy Current Nondestructive Method for Determination of Surface Carbon Content in Carburized Steels." *NDT & E International* 42, no. 7, pp. 618–21. doi: http://dx.doi.org/10.1016/j.ndteint.2009.04.008

Andrews, T., R.N. Hampton, A. Smedberg, D. Wald, V. Waschk, and W. Weissenberg. 2006. "The Role of Degassing in XLPE Power Cable Manufacture." *IEEE Electrical Insulation Magazine* 22, no. 6, pp. 5–16. doi: http://dx.doi.org/10.1109/mei.2006.253416

Arora, R. and W. Mosch. 2011. *High Voltage and Electrical Insulation Engineering*. IEEE Press.

Bai, L., G. Yun Tian, A. Simm, S. Tian, and Y. Cheng. March 2013. "Fast Crack Profile Reconstruction Using Pulsed Eddy Current Signals." *NDT & E International* 54, pp. 37–44. doi: http://dx.doi.org/10.1016/j.ndteint.2012.11.003

Batchelor, A.W., L. Nee Lam, and M. Chandrasekaran. 2011. *Materials Degradation and Its Control by Surface Engineering*. 3rd ed. Singapore: Imperial College Press.

Bavall. L. 2002. "Determination of the Thickness of Copper Coatings on Steel by Measuring the Impedance of a Thin Elliptic Coil." *Measurement Science and Technology* 13, no. 4, pp. 510–19. doi: http://dx.doi.org/10.1088/0957-0233/13/4/313

Beltle, M., A. Muller, and S. Tenbohlen. 2012. "Statistical Analysis of Online Ultrahigh-frequency Partial-discharge Measurement of Power Transformers." *IEEE Electrical Insulation Magazine* 28, no. 6, pp. 17–22. doi: http://dx.doi.org/10.1109/mei.2012.6340520

Biswas, S., C. Koley, B. Chatterjee, and S. Chakravorti. 2012. "A Methodology for Identification and Localization of Partial Discharge Sources Using Optical Sensors." *IEEE Transactions on Dielectrics and Electrical Insulation* 19, no. 1, pp. 18–28. doi: http://dx.doi.org/10.1109/tdei.2012.6148498

Borsi, H. December 1993. "The Relation between Thermal and Electrical Stress and the PD Behavior of Epoxy-rein Transformers." *IEEE Transactions on Electrical Insulation* 28, no. 6, pp. 1007–15. doi: http://dx.doi.org/10.1109/14.249374

Brown, P. December 1996. "Non-intrusive Partial Discharge Measurements in High Voltage Switchgear." *IEE Colloquium on Monitors and Condition Assessment Equipment*. IET, pp. 10/1–10/5, doi: 10.1049/ic:19961073.

Burzic, D., J. Zamberger, and E. Kozeschnik. July 2010. "Non-destructive Evaluation of Decarburization of Spring Steel Using Electromagnetic Measurement." *NDT & E International* 43, no. 5, pp. 446–50. doi: http://dx.doi.org/10.1016/j.ndteint.2010.04.006

Drinkwater, B.W. and P.D. Wilcox. October 2006. "Ultrasonic Arrays for Non-destructive Evaluation: A Review." *NDT & E International* 39, no. 7, pp. 525–41. doi: http://dx.doi.org/10.1016/j.ndteint.2006.03.006

Danikas, M.G. December 1993. "The Definitions Used for Partial Discharge Phenomena." *IEEE Transactions on Electrical Insulation* 28, no. 6, 1075–81. doi: http://dx.doi.org/10.1109/14.249381

Dey, D., B. Chatterjee, S. Chakravorti, and S. Munshi. 2010. "Cross-Wavelet Transform as a New Paradigm for Feature Extraction from Noisy Partial Discharge Pulses." *IEEE Transactions on Dielectrics and Electrical Insulation* 17, no. 1, pp. 157–66. doi: http://dx.doi.org/10.1109/tdei.2010.5412014

Dickinson, S.J., R. Binns, W. Yin, C. Davis, and A.J. Peyton. 2007. "The Development of a Multi-frequency Electromagnetic Instrument for Monitoring the Phase Transformation of Hot Strip Steel." *IEEE Transactions on Instrumentation and Measurement* 56, no. 3, pp. 879–886. doi: http://dx.doi.org/10.1109/TIM.2007.894183

Duval, M. November 1989. "Dissolved Gas Analysis: It Can Save Your Transformer." *IEEE Electrical Insulation Magazine* 5, no. 6, pp. 22–27. doi: http://dx.doi.org/10.1109/57.44605

Fracz, P. 2013. "Measurement of Optical Signals Emitted by Surface Discharges on Bushing and Post Insulator." *IEEE Transactions on Dielectrics and Electrical Insulation* 20, no. 5, pp. 1909–14. doi: http://dx.doi.org/10.1109/tdei.2013.6633724

Gamio, J.C., J. Castro, L. Rivera, J. Alamilla, F. Garcia-Nocetti, and L. Aguilar. April–June 2005. "Visualisation of Gas–oil Two-phase Flows in Pressurised Pipes Using Electrical Capacitance Tomography." *Flow Measurement and Instrumentation* 16, no. 2–3, pp. 129–34. doi: http://dx.doi.org/10.1016/j.flowmeasinst.2005.02.011

Gao, W., D. Ding, and W. Liu. 2011. "Research on the Typical Partial Discharge Using the UHF Detection Method for GIS." *IEEE Transactions on Power Delivery* 26, no. 4, pp. 2621–9. doi: http://dx.doi.org/10.1109/tpwrd.2011.2166089

Haddad, A. and D. Warne. 2007. *Advances in High Voltage Engineering*. 2nd ed. The Institution of Engineering and Technology.

Hao, X.J., W. Yin, M. Strangwood, A.J. Peyton, P.F. Morris, and C.L. Davis. 2008. "Off-line Measurement of Decarburization of Steels Using a Multifrequency Electromagnetic Sensor." *Scripta Materialia* 58, no. 11, pp. 1033–36. doi: http://dx.doi.org/10.1016/j.scriptamat.2008.01.042

Hao, X.J., W. Yin, M. Strangwood, A.J. Peyton, P.F. Morris, and C.L. Davis. 2009. "Characterization of Decarburization of Steels Using a Multifrequency Electromagnetic Sensor: Experiment and Modeling." *Metallurgical and Materials Transactions A* 40, no. 4, pp. 745–56. doi: http://dx.doi.org/10.1007/s11661-008-9776-y

Hashmi, G.M., M. Lehtonen, and M. Nordman. "Calibration of On-line Partial Discharge Measuring System Using Rogowski Coil in Covered-conductor Overhead Distribution Networks." *IET Science Measurement and Technology* 5, no. 1, pp. 5–13. doi: http://dx.doi.org/10.1049/iet-smt.2009.0124

Schmidt-Hattenberger, C., P. Bergmann, D. Bösing, T. Labitzke, M. Möller, S. Schröder, F. Wagner, and H. Schütt. 2013. "Electrical Resistivity Tomography (ERT) for Monitoring of CO_2 Migration—From Tool Development to Reservoir Surveillance at the Ketzin Pilot Site." *Energy Procedia* 37, pp. 4268–75. doi: http://dx.doi.org/10.1016/j.egypro.2013.06.329

Horan, P., P.R. Underhill, and T.W. Krause. April 2013. "Pulsed Eddy Current Detection of Cracks in F/A-18 Inner Wing Spar Without Wing Skin Removal Using Modified Principal Component Analysis." *NDT & E International* 55, pp. 21–7. doi: http://dx.doi.org/10.1016/j.ndteint.2013.01.004

Hudon, C. and M. Belec. 2005. "Partial Discharge Signal Interpretation for Generator Diagnostics." *IEEE Transactions on Dielectrics and Electrical Insulation* 12, no. 2, pp. 297–319. doi: http://dx.doi.org/10.1109/tdei.2005.1430399

IEC. 2000. *IEC60270: High-Voltage Test Techniques—Partial Discharge Measurements*. 3rd ed. International Standard.

Ijaz, U.Z., A.K. Khambampati, J.S. Lee, S. Kim, and K.Y. Kim. July 20, 2008. "Nonstationary Phase Boundary Estimation in Electrical Impedance Tomography Using Unscented Kalman Filter." *Journal of Computational Physics* 227, no. 15, pp. 7089–112. doi: http://dx.doi.org/10.1016/j.jcp.2007.12.025

Kahrobaee, S. and M. Kashefi. July 2011. "Hardness Profile Plotting Using Multi-Frequency Multi-Output Electromagnetic Sensor." *NDT & E International* 44, no. 4, pp. 335–8. doi: http://dx.doi.org/10.1016/j.ndteint.2011.01.005

Kaneko, S., S. Okabe, M. Yoshimura, H. Muto, C. Nishida, and M. Kamei. 2009. "Detecting Characteristics of Various Type Antennas on Partial Discharge Electromagnetic Wave Radiating Through Insulating Spacer in Gas Insulated Switchgear." *IEEE Transactions on Dielectrics and Electrical Insulation* 16, no. 5, pp. 1462–72. doi: http://dx.doi.org/10.1109/tdei.2009.5293961

Kimura, K. and Y. Kaneda. June 1995. "The Role of Microscopic Defects in Multistress Aging of Micaceous Insulation." *IEEE Transactions on Dielectrics and Electrical Insulation* 2, no. 3, pp. 426–32. doi: http://dx.doi.org/10.1109/94.395419

Kopf, U. and K. Feser. 1995. "Rejection of Narrow-band Noise and Repetitive Pulses in On-site PD Measurements." *IEEE Transactions on Dielectrics and Electrical Insulation* 2, no. 6, pp. 1180–91. doi: http://dx.doi.org/10.1109/94.484323

Koo, J.Y., S.Y. Jung, C.H. Ryu, S.W. Lee, and B.W. Lee. 2010. "Identification of Insulation Defects in Gas-insulated Switchgear by Chaotic Analysis of Partial Discharge." *IET Science, Measurement & Technology* 4, no. 3, pp. 115–24. doi: http://dx.doi.org/10.1049/iet-smt.2009.0028

Kraetge, A., S. Hoek, M. Koch, and W. Koltunowicz. 2013. "Robust Measurement, Monitoring and Analysis of Partial Discharges in Transformers and Other HV Apparatus." *IEEE Transactions on Dielectrics and Electrical Insulation* 20, no. 6, pp. 2043–51. doi: http://dx.doi.org/10.1109/tdei.2013.6678852

Kuffel, E., W.S. Zaengl, and J. Kuffel. 2000. *High Voltage Engineering: Fundamentals*. Oxford, UK: Butterworth-Heinemann.

Lazarovitch, R., D. Rittel, and I. Bucher. July 2002. "Experimental Crack Identification using Electrical Impedance Tomography." *NDT & E International* 35, no. 5, pp. 301–16. doi: http://dx.doi.org/10.1016/s0963-8695(01)00054-8

Lelekakis, N., D. Martin, W. Guo, and J. Wijaya. 2011. "Comparison of Dissolved Gas-in-Oil Analysis Methods Using a Dissolved Gas-in-Oil Standard." *IEEE Electrical Insulation Magazine* 27, no. 5, pp. 29–35. doi: http://dx.doi.org/10.1109/mei.2011.6025366

Li, X., J. Chen, and L. Gao. 2012. "Enhanced Degradation of Phenol by Carbonate Ions with Dielectric Barrier Discharge." *IEEE Transactions on Plasma Science* 40, no. 1, pp. 112–17. doi: http://dx.doi.org/10.1109/tps.2011.2173356

Lundgaard, L.E. September/October 1992. "Partial Discharge—Part XIV: Acoustic Partial Discharge Detection—Practical Application." *IEEE Electrical Insulation Magazine* 8, no. 5, pp. 34–43. doi: http://dx.doi.org/10.1109/57.156943

Ma X., C. Zhou, and I.J. Kemp. September 17–20, 2000. Investigation into the use of Wavelet Theory for Partial Discharge Pulse Extraction in Electrically Noisy Environments. *Proceeding of the 8th International Conference Materials, Measurements and Application*. Edinburgh, UK.

Ma, X., C. Zhou, and I.J. Kemp. 2002. "Interpretation of Wavelet Analysis and its Application in Partial Discharge Detection." *IEEE Transactions on Dielectrics and Electrical Insulation* 9, no. 3, pp. 446–57. doi: http://dx.doi.org/10.1109/tdei.2002.1007709

Ma, X., C. Zhou, and I.J. Kemp. 2002. "Automated Wavelet Selection and Thresholding for PD Detection." *IEEE Electrical Insulation Magazine* 18, no. 2, pp. 37–45. doi: http://dx.doi.org/10.1109/57.995398

Ma, X., C. Zhou, and I.J. Kemp. May 18–20, 2004. Novel Computer-based Processing System for Partial Discharge Detection and Diagnosis. *Proceedings of the 21st IEEE Instrumentation and Measurement Technology Conference*, Vol. 1, pp. 684–8. Como, Italy.

Ma, X., A.J. Peyton, and Y.Y. Zhao. July 2005. "Measurement of the Electrical Conductivity of Open-celled Aluminium Foam using Non-contact Eddy Current Techniques." *NDT & E International* 38, no. 5, pp. 359–67. doi: http://dx.doi.org/10.1016/j.ndteint.2004.10.003.

Ma, X., A.J. Peyton, R. Binns, and S.R. Higson. April 2005. "Electromagnetic Techniques for Imaging the Cross-section Distribution of Molten Steel Flow in the Continuous Casting Nozzle." *IEEE Sensors Journal* 5, no. 2, pp. 224–32. doi: http://dx.doi.org/10.1109/jsen.2004.842443

Ma, X. and A.J. Peyton. April 2006. "Eddy Current Measurement of the Electrical Conductivity and Porosity of Metal Foams." *IEEE Transactions on Instrumentation and Measurement Technology* 55, no. 2, pp. 570–6. doi: http://dx.doi.org/10.1109/tim.2006.873549

Ma, X., A.J. Peyton, and Y.Y. Zhao. 2006. "Eddy Current Measurements of Electrical Conductivity and Magnetic Permeability of Porous Metals." *NDT & E International* 39, no. 7, pp. 562–8. doi: http://dx.doi.org/10.1016/j.ndteint.2006.03.008

Ma, X., A.J. Peyton, S.R. Higson, A. Lyons, and S.J. Dickinson. 2006. "Hardware and Software Design for an Electromagnetic Induction Tomography (EMT) System Applied to High Contrast Metal Process Applications." *Measurement Science and Technology* 17, no. 1, pp. 111–18. doi: http://dx.doi.org/10.1088/0957-0233/17/1/018

Ma, X., A.J. Peyton, S.R. Higson, and P. Drake. 2008. "Development of Multiple Frequency Electromagnetic Induction Systems for Steel Flow Visualization." *Measurement Science and Technology* 19, no. 9. doi: http://dx.doi.org/10.1088/0957-0233/19/9/094008

Ma, X. 2012. "Investigations of the State-of-the-Art Methods for Electromagnetic NDT and Electrical Condition Monitoring." *Insight-Non-Destructive Testing and Monitoring* 54, no. 9, pp. 482–8. doi: http://dx.doi.org/10.1784/insi.2012.54.9.482

Ma, H., J.C. Chan, T.K. Saha, and C. Ekanayake. 2013. "Pattern Recognition Techniques and Their Applications for Automatic Classification of Artificial Partial Discharge Sources." *IEEE Transactions on Dielectrics and Electrical Insulation* 20, no. 2, pp. 468–78. doi: http://dx.doi.org/10.1109/tdei.2013.6508749

Markalous, S., S. Tenbohlen, and K. Feser. 2008. "Detection and Location of Partial Discharges in Power Transformers using Acoustic and Electromagnetic Signals." *IEEE Transactions on Dielectrics and Electrical Insulation* 15, no. 6, pp. 1576–83. doi: http://dx.doi.org/10.1109/tdei.2008.4712660

Marsh, L.A., C. Ktistis, A. Jarvi, D.W. Armitage, and A.J. Peyton. March 2014. "Three-dimensional Object Location and Inversion of the Magnetic Polarizability Tensor at a Single Frequency Using a Walk-through Metal Detector." *Measurement Science and Technology* 24, no. 4. doi: http://dx.doi.org/10.1088/0957-0233/24/4/045102

Mayoux, C. October 2000. "Degradation of Insulating Materials Under Electrical Stress." *IEEE Transactions on Dielectrics and Electrical Insulation* 7, no. 5, pp. 590–601. doi: http://dx.doi.org/10.1109/tdei.2000.879355

Mazzetti, C., F.M.F. Mascioli, F. Baldini, M. Panella, R. Risica, and R. Bartnikas. 2006. "Partial Discharge Pattern Recognition by Neuro-fuzzy Networks in Heat-shrinkable Joints and Terminations of XLPE Insulated Distribution Cables." *IEEE Transactions on Power Delivery* 21, no. 3, pp.1035–44. doi: http://dx.doi.org/10.1109/tpwrd.2006.875861

Mercier, D., J. Lesage, X. Decoopman, and D. Chicot. December 2006. "Eddy Currents and Hardness Testing for Evaluation of Steel Decarburizing." *NDT & E International* 39, no. 8, pp. 652–60. doi: http://dx.doi.org/10.1016/j.ndteint.2006.04.005

Mook, G., R. Lange, and O. Koeser. May 2001. "Non-destructive Characterisation of Carbon-fibre-reinforced Plastics by Means of Eddy-currents." *Composites Science and Technology* 61, no. 6, pp. 865–73. doi: http://dx.doi.org/10.1016/s0266-3538(00)00164-0

Pearson, J.S., O. Farish, B.F. Hampton, M.D. Judd, D. Templeton, B.M. Pryor, and I.M. Welch. October 1995. "Partial Discharge Diagnostics for Gas Insulated Substations." *IEEE Transactions on Dielectrics and Electrical Insulation* 2, no. 5, pp. 893–905. doi: http://dx.doi.org/10.1109/94.469984

Riedel C.H., M. Keppelen, S. Nani, R.D. Merges, and O. Dössel. 2004. "Planar System for Magnetic Induction Conductivity Measurement using a Sensor Matrix." *Physiological Measurement* 25, no. 1, pp. 403–11. doi: http://dx.doi.org/10.1088/0967-3334/25/1/043

Rogers, R.R. 1978. "Concepts used in the Development of the IEEE and IEC Codes for the Interpretation of Incipient Faults in Power Transformers by Dissolved Gas in Oil Analysis." *IEEE Winter Meeting* EI-13, no. 5. doi: http://dx.doi.org/10.1109/TEI.1978.298141

Rubio-Serrano, J., M.V. Rojas-Moreno, J. Posada, J.M. Martínez-Tarifa, G. Robles, and J.A. Garcia-Souto. 2012. "Electro-acoustic Detection, Identification and Location of Partial Discharge Sources in Oil-paper Insulation Systems." *IEEE Transactions on Dielectrics and Electrical Insulation* 19, no. 5, pp.1569–78. doi: http://dx.doi.org/10.1109/tdei.2012.6311502

Rudd, S., S.D.J. Mcarthur, and M.D. Judd. 2010. "A Generic Knowledge-based Approach to the Analysis of Partial Discharge Data." *IEEE Transactions on Dielectrics and Electrical Insulation* 17, no. 1, pp.149–56. doi: http://dx.doi.org/10.1109/tdei.2010.5412013

Sher, Z.K., D. Zhu, X. Jin, and K. Tan. 1995. "A New Adaptive Technique for On-line Partial Discharge Monitoring." *IEEE Transactions on Dielec-*

trics and Electrical Insulation 2, no. 4, pp. 700–707. doi: http://dx.doi.org/10.1109/94.407034

Shintemirov, A., W. Tang, and Q.H. Wu. 2009. "Power Transformer Fault Classification Based on Dissolved Gas Analysis by Implementing Bootstrap and Genetic Programming." *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews* 39, no. 1, pp. 69–79. doi: http://dx.doi.org/10.1109/tsmcc.2008.2007253

Simons, J.S. May 1980. "Diagnostic Testing of High-voltage Machine Insulation." *IEE Proceedings, Part B: Electric Power Applications* 127, no. 3, pp. 139–54. doi: http://dx.doi.org/10.1049/ip-b.1980.0055

Sriram S., S. Nitin, K.M.M. Prabhu, and M.J. Bastiaans. 2005. "Signal Denoising Techniques for Partial Discharge Measurements." *IEEE Transactions on Dielectrics and Electrical Insulation,* pp. 1182–91. doi: http://dx.doi.org/10.1109/tdei.2005.1561798

Sposito, G., C. Ward, P. Cawley, P.B. Nagy, and C. Scruby. October 2010. "A Review of Non-destructive Techniques for the Detection of Creep Damage in Power Plant Steels." *NDT & E International* 43, no. 7, pp. 555–67. doi: http://dx.doi.org/10.1016/j.ndteint.2010.05.012

Stone, G.C., H.G. Sedding, N. Fujimoto, and J.M. Braun. 1990. "Practical Implementation of Ultrawideband Partial Discharge." *IEEE Transactions on Electrical Insulation* 27, no. 1, pp. 70–81. doi: http://dx.doi.org/10.1109/14.123442

Sun, M., S. Liu, Z. Li, J. Lei. 2008. "Application of Electrical Capacitance Tomography to the Concentration Measurement in a Cyclone Dipleg." *Chinese Journal of Chemical Engineering* 16, no. 4, pp. 635–9. doi: http://dx.doi.org/10.1016/s1004-9541(08)60133-0

Theodoulidis, T., H. Wang, and G.Y. Tian. April 2012. "Extension of a Model for Eddy Current Inspection of Cracks to Pulsed Excitations." *NDT & E International* 47, pp. 144–9. doi: http://dx.doi.org/10.1016/j.ndteint.2012.01.005

Yang, M.-T., and L.-S. Hu. 2013. "Intelligent Fault Types Diagnostic System for Dissolved Gas Analysis of Oil-immersed Power Transformer." *IEEE Transactions on Dielectrics and Electrical Insulation* 20, no. 6, pp. 2317–24. doi: http://dx.doi.org/10.1109/tdei.2013.6678885

Yin, W., and A. J. Peyton. 2006. "A Planar EMT System for the Detection of Faults on Thin Metallic Plates". *Measurement Science and Technology*, 17, no. 8, pp. 2130–5.

Zhang, Z.S., D.M. Xiao, and Y. Li. May 2009. "Rogowski Air Coil Sensor Techniques for On-line Partial Discharge Measurement of Power Cables." *IET Science Measurement and Technology* 3, no. 3, pp. 187–96. doi: http://dx.doi.org/10.1049/iet-smt:20080025

Zhu, W., W. Yin, and A.J. Peyton. November 2011. "Modelling and Experimental Study of an Electromagnetic Sensor with an H-shaped Ferrite Core used for Monitoring the Hot Transformation of Steel in an Industrial Environment."

NDT & E International 44, no. 7, pp. 547–52. doi: http://dx.doi.org/10.1016/j.ndteint.2011.05.005

Zhu, W., S. Cruchley, W. Yin, X.J. Hao, C.L. Davis, and A.J. Peyton. March 2012. "Evaluation of Rail Decarburisation Depth using a H-Shaped Electromagnetic Sensor." *NDT & E International* 46, no. 1, pp. 63–9. doi: http://dx.doi.org/10.1016/j.ndteint.2011.10.005

CHAPTER 6

PHYSICS OF DEGRADATION IN FERROELECTRIC DEVICES

Paul M. Weaver

6.1 INTRODUCTION

Ferroelectric describes a class of materials that are the electrical analogs of ferromagnets where magnetization is replaced by electrical polarization and the magnetic field is replaced by the electric field. It is this commonality that is behind the "ferro" term in the name, rather than the presence of iron. In common with their magnetic cousins, ferroelectric materials display a remanent polarization that persists even in the absence of an electric field, and can be reversed by the application of a sufficiently strong electric field (Figure 6.1). Also in common with ferromagnets, they have been adopted in a very wide variety of important technological applications such as medical ultrasound, pyroelectric motion detectors, and diesel engine fuel injectors and many others.

Piezoelectric effect	Actuators, sensors, transducers
Pyroelectric effect	Motion sensors, pyrometers
High permittivity	Capacitors, energy storage
Tunable permittivity	Voltage-controlled capacitors
Reversible polarization	Ferroelectric memory
Electrocaloric effect	Solid-state cooling
Magnetoelectric effect	Sensors, data storage
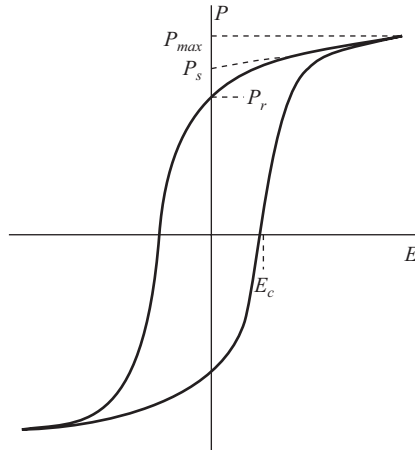
Figure 6.1. P–E hysteresis loop parameters for a ferroelectric material [after (B. Jaffe, Cook, and H. Jaffe 1971)].
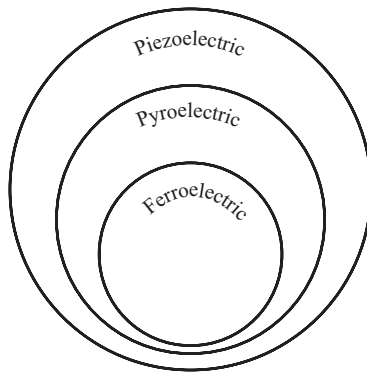
Figure 6.2. Relationship between piezoelectric, pyroelectric, and ferroelectric materials.

Ferroelectrics display both piezoelectric and pyroelectric effects that are responsible for many of their technological applications (see Figure 6.2). Piezoelectricity is the conversion of mechanical to electrical energy, and vice versa. A subset of piezoelectric materials possesses a permanent electric polarization. Because the polarization varies with temperature, these materials are also pyroelectric.

Ferroelectrics are distinguished from the broader class of pyroelec-
tric materials by the fact that the electric polarization can be reversed by
the application of an electric field, as described above. In principal, the
polarization of any of the pyroelectric materials could be reversed by the
application of a sufficiently large electric field, but in the case of non-
ferroelectric pyroelectrics, this field exceeds the breakdown strength of
the material, so the polarization reversal is not practically achievable. An
important technological consequence of the reversibility of the polariza-
tion in ferroelectrics is that it makes possible the fabrication of ceramics
with piezoelectric and pyroelectric properties. A ceramic is formed by sin-
tering together grains of the material. Because the grains are randomly
oriented, the piezoelectric and pyroelectric response averages to zero in
the bulk ceramic. The most widely used ferroelectric material for piezo-
electric applications is lead zirconate titanate (PZT). Nonferroelectric
pyroelectric and piezoelectric devices must therefore be made from single
crystals or other methods that create orientation of the crystallographic
axes. These methods are necessarily expensive, and restrict the geometries
and fabrication methods that can be employed.

Because applications of ferroelectrics invariably involve their elec-
trical properties, performance depends on electrical and mechanical fac-
tors, as well as the coupling between them. This chapter focuses on the
factors that affect the coupling and functional performance. The effects
of humidity and temperature on performance are described in detail. The
important topics of electrical breakdown and fatigue in ferroelectrics have
been treated extensively elsewhere (Cain 2014; Lupascu 2004), and so are
not covered in detail here. To develop our understanding of the physics
affecting performance in harsh environments, it is necessary to be able
to reliably measure material and device characteristics under challenging
conditions, so a section on measurement is also included.

6.2   HUMIDITY

This section describes factors affecting the performance of piezoelectric
devices in conditions of high humidity under (direct current) d.c. electric
fields. This does not appear to directly affect the piezoelectric properties
of the material, but rather the electrical properties through the creation of
leakage current pathways through the material. Significant levels of leak-
age current increase power consumption, and can lead to device failure.

Leakage currents have been shown to be associated with the devel-
opment of discrete conducting sites distributed across the whole area of

the material subject to the electric field. These sites appear to be related to pores, and it is likely that pores in the ceramic provide a route for ingress of moisture. This is supported by the observation that high-density, low-porosity ceramics show significantly longer lifetime in humid conditions. However, pores are not the only mechanism for water vapor to affect the ceramic. Conduction also occurs at edges (and surfaces exposed by cracking). Edge conduction becomes a more prominent contributor in low-porosity ceramics. In this case, surface contamination and poor surface condition can significantly increase the rate of growth of the leakage current. These observations suggest that porosity is a major factor in the development of leakage currents in ceramics exposed to high humidity and high electric fields, although electrode material and processing conditions also have an effect.

Polymer barrier coatings provide protection against gross contamination from water drops and condensation, but provide little impediment to the passage of water vapor from the air. They are therefore generally ineffective in preventing long-term protection against the effects of humidity. Combination coatings such as metal or diamond-like carbon (DLC) on a polymer base may extend lifetime in humid conditions, but processing to achieve full-area coverage and sealing at electrical connections present significant challenges. Control over the manufacturing processes from ceramic through to device integration is crucial for achieving cost-effective performance in harsh environments.

6.2.1 CONDUCTION MECHANISMS

Undoped PZT is known as a wide bandgap semiconductor with a bandgap of between 2.6 and 3.5 eV (Nagaraj et al. 1999). It also exhibits p-type conductivity due to the presence of low-valence impurities substituting for higher valence Pb ions. This causes A-site (Pb ion) vacancies, which act as electron acceptors, leading to the production of holes (B. Jaffe, Cook, and H. Jaffe 1971). This means that a Schottky barrier can be formed with the electrode material, which can promote conduction in thin films (Dawber, Rabe, and Scott 2005; Nagaraj et al. 1999). Tunneling can also be a mechanism for conduction in very thin films (Dawber, Rabe, and Scott 2005; Scott 2006). However, for films thicker than a few tens of nm, conduction is determined by the bulk of the material through Poole–Frenkel emission, where thermally excited electrons are assisted by the electric field into the conduction band (Nagaraj et al. 1999).

The large piezoelectric coefficients in this class of "soft" PZT materials is achieved by A-site vacancy substitution (e.g., La doping), which also

confers high electrical resistivity (B. Jaffe, Cook, and H. Jaffe 1971). Data from B. Jaffe, Cook, and H. Jaffe (1971) would indicate a volume resistivity of 10^9 Ωcm at temperatures around room temperature. For the sample size and field strength used in this study we would expect a resistance of over 300 GΩ, and a leakage current of less than 1 nA. The currents measured after exposure to humid conditions under d.c. bias are in the μA to mA range, so it cannot be explained by these electronic conduction mechanisms. In addition, there is no obvious reason why bulk conduction mechanisms should be directly affected by humidity, or the duration of the stress condition (humidity + d.c. field), although secondary effects such as humidity-induced chemical changes could play a part.

There is evidence that the conduction mechanism in humid environments is ionic in character. This includes the observation of anodic breakdown features and sensitivity of the leakage current to the relative humidity and electrode material (Lipscomb 2009). The large leakage currents observed under humid conditions are not observed under dry atmosphere (<250 nA at 55°C), confirming that the conductivity is caused by the humidity. In fact, a reduction of leakage current is usually observed on heating to this temperature from room temperature due to the expulsion of moisture. The effects of the bulk conduction mechanisms referred to above are only easily measurable above around 200°C.

Humidity ingress has long been recognized as a potential failure mode in ceramic capacitors, which are often based on ferroelectric materials such as barium titanate ($BaTiO_3$) and operate under similar levels of electric field to ferroelectrics in other applications such as actuators. There have been a number of studies of the effects of DC fields on electrical breakdown in multilayer capacitors (MLCs). Waser, Baiatu, and Hardtl (1990) have termed resistance degradation as a limiting factor in the lifetime of MLCs. In MLCs, the effects of humidity have also been investigated, particularly with silver electrodes (Ling and Jackson 1989). Subsequent studies on the reliability of multilayer ceramic actuators have also examined the behavior of silver electrodes under d.c. operation in humid conditions (Nagata and Thongrueng 1998; Pritchard, Bowen, and Lowrie 2001; Thongrueng, Tsuchiya, and Nagata 1998). More recently, the design of multilayer actuators has been optimized to reduce the effects of humidity by burying the electrode within the actuator, and through the use of protective ceramic coatings (Bindig and Helke 2000; Pertsch et al. 2006).

More recently, the effects of water vapor have been studied in the context of memory and electronic applications Baniecki et al. (2002) where it has been shown to have a strong influence on the leakage characteristics of PZT thin films. The application of large d.c. electric fields to piezoelectric

actuators in humid environments is known promote the development of leakage currents that increase with time (Lipscomb et al. 2009). This results in an increase in power consumption and may ultimately lead to failure of the device.

Changes in insulation resistance are accelerated by higher electric fields (Lipscomb et al. 2009) and it appears that both humidity and an electric field are required (Chen et al. 2003). This increase in conductivity is generally thought to be caused by electromigration of metallic cations formed at the anode (Lipscomb et al. 2009; Thongrueng, Tsuchiya, and Nagata 1998; Zheng, Swingler, and Weaver 2010). The ions migrate from the anode to the cathode where they recombine to form metallic conductive pathways.

Silver electrodes are well known for promoting this behavior (Pertsch et al. 2006; Thongrueng, Tsuchiya, and Nagata 1998; van den Ende et al. 2009), but this type of electromigration is also known with nickel electrodes (Kostic, Rensch, and Sturm 1995). There is also some evidence that hydrogen, evolved electrolytically at the electrodes, can also play a role in increasing the conductivity of both PZT (Chen et al. 2002) and barium titanate (Chen et al. 2003) ceramics, probably through reduction of the oxides that form the ceramic itself. Degradation in electrical resistance can also be accelerated by crack formation, particularly in cyclic applications (van den Ende et al. 2009). The development of cracks is also affected by humidity, and there are reports of crack development being both delayed (Jiang et al. 2008) and accelerated (van den Ende et al. 2009) in high humidity conditions, with the mechanism depending on both the ceramic and electrode materials.

Ionic silver migration from silver electrodes is commonly implicated in humidity-related conduction in otherwise insulating materials such as thick film insulators (Ling and Jackson 1989; Mock and Lubitz 2008; Thongrueng, Tsuchiya, and Nagata 1998). A similar mechanism has also been proposed previously for electrical changes in PZT with Ni electrodes (Lipscomb et al. 2008, 2009). This mechanism (Figure 6.3) involves ingress of water at the positive anode through a pore or defect, followed by oxidation of the cathode metal to form ions in solution. The ions then migrate under the electrical field until they reach the cathode. At the cathode, the ions receive electrons and are reduced back to the metal. This newly deposited metal forms filaments through the grains and defects in the ceramic, which grow back toward the anode. As they approach the anode, the distance becomes insufficient to support the electric field and breakdown occurs.

The extremity of the metal filament close to the anode would be partially destroyed by this breakdown, thus switching off the current
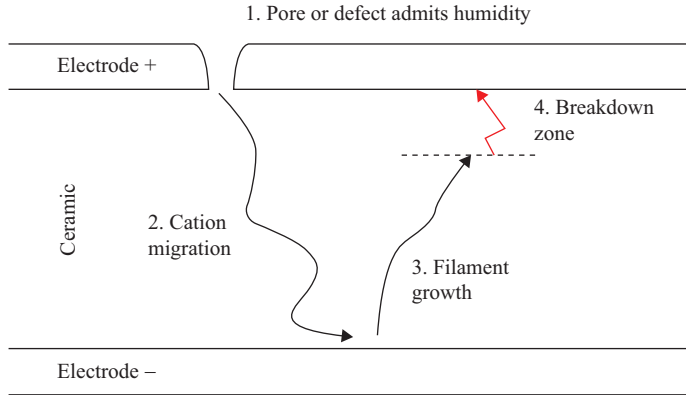
Figure 6.3. Electromigration model for conduction in humid environments.
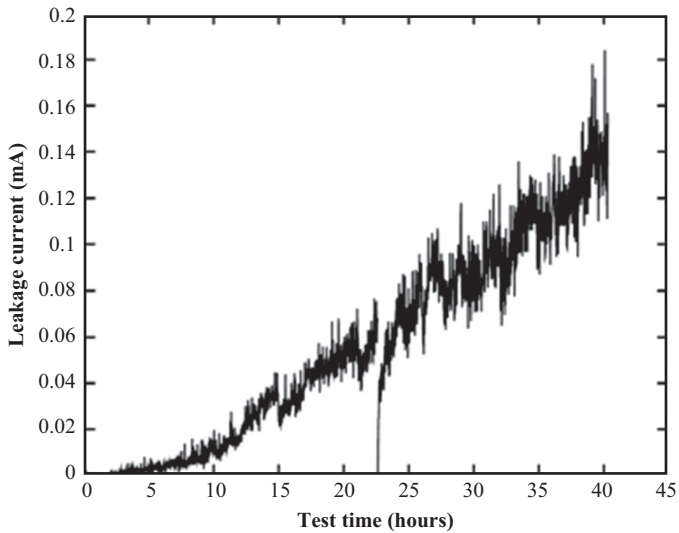
Figure 6.4. Time evolution of leakage current 80% RH 80°C.

until it regrows. This leads to the current flowing through a succession of rapid transients, highly characteristic of this breakdown mechanism (Zheng, Swingler, and Weaver 2010). Because the conducting channels take time to form, there is an initial lag before any leakage current is observed (Figure 6.4). As the numerous conducting channels all grow at different rates, the conductivity builds up over time—a process that can

be explained by a statistical model of the growth (Lipscomb et al. 2009). Eventually all the channels are fully formed, and the leakage current reaches a steady state or only slowly increasing value.

The amount of metallic conducting material required to account for the conductivity observed in measurements of leakage currents described here is extremely small and therefore difficult to detect. This means that in piezoelectric ceramics, the evidence for this is usually indirect, although there have been reports of migrated silver metal in capacitor structures after exposure to humidity and electric field (Ling and Jackson 1989).

The development of the conductivity in the ceramic is often accompanied by the evolution of visible features on the cathode (Thongrueng, Tsuchiya, and Nagata 1998; Weaver et al. 2012). Such a feature on PZT after exposure to humidity and electric field is shown in Figure 6.5. Most of the surface is clearly metallic, but around the breakdown site there is an area where the electrode material appears to be no longer present. The structural features normally present on the electrode or ceramic surface (grains, etc.) are not visible in this region, indicating that melting has occurred in the area. There are also a number of small cracks visible around the feature, consistent with fast cooling. These features are indicative of an energetic high-temperature breakdown such as an arc discharge. Further evidence for the energetic nature of the process can be seen in the cross section of Figure 6.6 where significant cracking through the depth of the ceramic is observed. The evolution of cracks could be exacerbated by the piezoelectric electromechanical response to a sudden reduction in electric field in the vicinity of the breakdown. It is likely that breakdown events occur repetitively at the same site. This could also cause gradual
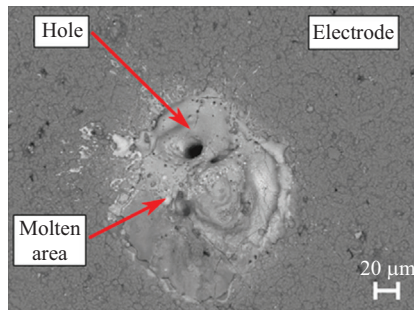
Figure 6.5. Scanning electron micrograph of the breakdown feature and erosion of electrode material on the anode surface (Weaver et al. 2012).
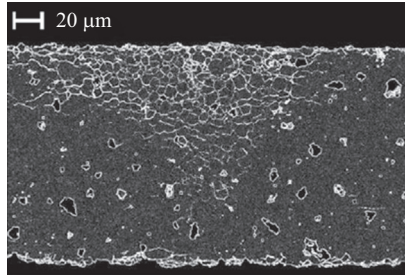
Figure 6.6. Cross section of the ceramic (with electrode removed after test) in the region of a breakdown feature. The anode is at the top of the picture. This image has been processed to enhance the visibility of edge features (Weaver et al. 2012).

damage to the ceramic structure. The resistivity of the thin electrode material (approximately 2 μm) will serve to limit the overall current, although a localized high-current density on breakdown is likely. It is thought that this current-limited breakdown transient is responsible for the current fluctuations (Zheng, Swingler, and Weaver 2010). The fact that these features only occur on the anode is strong evidence for the ionic migration/metal filament mechanism for the growth of the conductive channels.

At the center of the fused area is a hole. It is possible that this was originally a surface pore that provided a location for entry of the water vapor required to promote electromigration. This observation supports the idea that pores play an important role in the susceptibility to humidity and that the porosity of the ceramic is an important factor.

6.2.2 GEOMETRIC FACTORS

The conduction model described above (Lipscomb et al. 2009) is a bulk process, but surface conduction is also a possible contributor to the overall leakage current. An important question, therefore, is whether surface conduction at the edges of the ceramic plays a significant role. Trials with electrodes that left a margin from the edge did not show any improvement in the lifetime (defined as the time taken for the resistance to drop to 10 percent of its maximum value (Lipscomb et al. 2009). In addition, discolorations that were possible breakdown sites have been observed at various points over the surface and were not confined to the edges (Lipscomb et al. 2009). Experiments where the edges were removed part way through

a test showed that the leakage current was rapidly restored to its value before removal, confirming that the edges were not a major contributor to the conductivity (Weaver et al. 2012). In another experiment, a ceramic sample was tested for approximately 20 hours, then removed from test and broken into two approximately equal halves. Both halves were retested separately. After a short time for recovery of the leakage current, the two halves showed approximately equal leakage currents that were approximately half the current before breaking in two. The breaking of the ceramic pieces creates two new edges with a combined length of approximately 33 percent of the original perimeter. This increase in edge area was not reflected in a corresponding increase in leakage current. There is evidence, however, that edge effects become significant at longer timescales in ceramics where the leakage by other mechanisms has been suppressed.

6.2.3   CERAMIC POROSITY

These studies suggest a number of possible ways, at the ceramic component level, of improving the lifetime of piezoelectric devices in humid conditions. Porosity in the ceramic provides a pathway for ingress of moisture into the body of the ceramic, so reducing the porosity would be expected to improve lifetime. Typical pore features in a commercial soft PZT ceramic are shown in Figure 6.7

Most ceramics are available in lower porosity versions, although higher production costs make them more expensive. Figure 6.8 shows

(a)   **(b)**

Figure 6.7. Scanning electron micrograph showing porosity (a) through the thickness of the ceramic and (b) on the electrode surface of the ceramic (Weaver et al. 2012).

(a)   (b)

Figure 6.8. SEM cross sections showing (a) porosity in the standard ceramic and (b) low-porosity version of the ceramic (Weaver et al. 2012).

SEM cross sections of two ceramics of the same composition with different porosity.

Test results demonstrated very significant reductions in leakage current in the high-porosity version of the ceramic. The longer lifetimes achieved with low-porosity ceramics means that factors, such as edge effects, that did not play a significant role in conduction through previously now become a limiting factor. Optical microscope study of a tested low-porosity ceramic (after electrode removal) showed that eventually discolorations of the ceramic due to breakdown events were observed, but that these were mostly located at the edges rather than through the bulk. It is thought that this edge conduction can ultimately lead to short circuit failure but at much longer timescales. Contamination was also found to play a role in the conduction processes in low-porosity ceramics that was masked in the more porous ceramics. Aqueous cleaning to remove ionic contaminants and careful handling was found to improve lifetime and consistency in performance.

6.2.4 ELECTRODE MATERIALS

Electrode materials and processing conditions affect the level of leakage current observed. As conduction occurs through the bulk of the material away from the edges, it must be able to penetrate the metallic electrode material through coverage defects, cracks, or incompletely covered surface pores. This would imply that thicker electrodes may improve lifetime in humidity. However, this approach increases the stiffness of the device possibly leading to unacceptable degradation of device performance, particularly for thinner ceramics. Gold electrodes have been demonstrated to significantly improve performance (Lipscomb et al. 2009). However, Au electrodes can be expensive particularly if thickness is required to overcome porosity.

6.2.5   BARRIER COATINGS

One approach to improving the lifetime of an actuator in humid conditions is to prevent the moisture from reaching the ceramic by means of a moisture barrier. For actuator applications, any moisture barrier must not impede the movement of the actuator. This requirement for flexibility precludes the use of thick ceramic or epoxy encapsulation that could be used to protect, for example, an electronic component. This is particularly true for thin bender actuators where even relatively thin polymer coatings can reduce the actuator movement, especially at low temperatures when the coating can become stiff. Thin flexible polymer coatings are all permeable (Massey 2003) and provide insufficient protection against water vapor transmission (although they do provide important protection against water, dust, and other contaminants).

As discussed above, all materials are, to some degree, permeable to water, and polymers are more permeable than solid materials. Combination coatings or coatings of semi-crystalline materials such as DLC may provide more effective protection than simple polymer barriers (Jones et al. 2010).

6.2.6   PIEZOELECTRIC PROPERTIES

It is clear that humidity has a significant effect on the electrical properties of the ceramic, but electrical measurements do not tell us if the piezoelectric properties are affected. Measurements of piezoelectric displacement at 80°C and 80 percent RH with 250 volt bias at the start of a test then after 2.5 hours, 5.4 hours, 8.3 hours, and 11.3 hours showed no significant change in piezoelectric performance.

6.3   TEMPERATURE EFFECTS ON RELIABILITY OF PIEZOELECTRIC ACTUATORS

6.3.1   THERMAL EXPANSION IN FERROELECTRICS

To illustrate the importance of thermal expansion in the reliability of a piezoelectric device, we consider a typical application such as a piezoelectrically actuated valve. A schematic of such a device is shown in Figure 6.9. To operate correctly the actuator needs to move between two positions that open and close the valve. In the open position, there must

Actuator

Actuator movement

Open

Closed

Valve opening
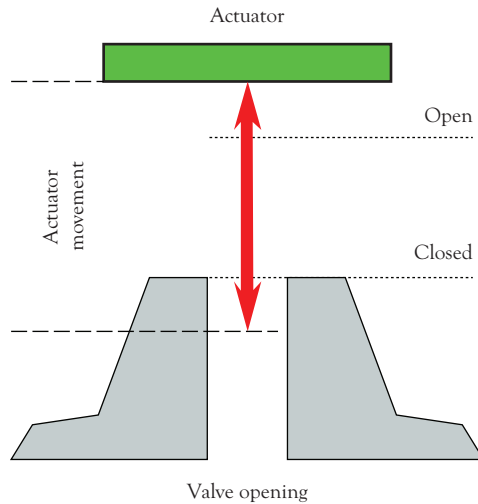
Figure 6.9. Schematic diagram of piezoelectrically operated valve.

be sufficient clearance between the valve seal and the orifice to achieve the required flow rate. In the closed position, the valve seal needs to at least cover the orifice and, usually, some overtravel is needed to create sufficient force to create an effective seal. These two positions define the range of movement required from the actuator between the ON condition and the OFF condition. Note that these two positions are relative to the valve orifice. If there is any drift in the position of the actuator, then it could encroach on these limits and compromise the performance of the valve. This is why thermal expansion is an important factor in the design and reliability of many piezoelectric devices. In fact, it is the difference in thermal expansion between the piezoelectric material and the supporting structure that leads to thermal drift. For example, a typical actuation strain may be around 1,000 ppm. A thermal expansion difference of 10 ppm would create a drift equal to the entire actuation strain over a 100°C temperature span. Such a valve would only work across a limited temperature range.

The piezoelectric materials most widely used for actuator applications such as the valve are ceramics of ferroelectric materials, usually lead zirconate titanate (PZT), because of the very coupling between their electrical polarization and mechanical strain. As described above, these materials are also pyroelectric, with strong coupling between electrical
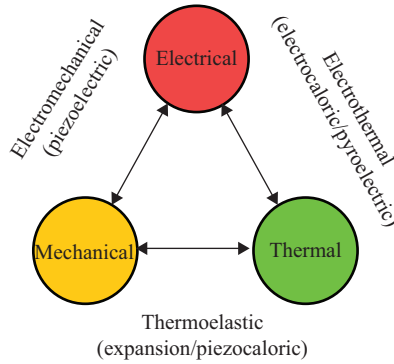
Figure 6.10. Coupling between electrical, mechanical, and thermal properties in a ferroelectric material.

polarization and temperature. There is, therefore, also a strong coupling between temperature and mechanical strain as shown in Figure 6.10.

Because of this coupling between the temperature-dependent polarization and mechanical strain, the thermal expansion of a ferroelectric is more complex than that of most materials, with strong nonlinearity (temperature dependence of the expansion coefficient), thermal hysteresis, and anisotropy in the thermal expansion, as well as being influenced by electric field or mechanical stress.

When considering thermal expansion in ferroelectric materials, it is useful to think of the relationship between ferroelectric and piezoelectric properties in terms of electrostriction. Electrostriction is a property of all materials whereby a strain is induced by an electric field. Most materials are nonpolar so the strain response is symmetric with respect to the electric field, that is, strain occurs in the same direction irrespective of the sign of the electric field. Because of this symmetry, the strain, S, depends on only even powers of electric field and for most materials the effect is quadratic and described by an electrostriction coefficient Q

$$S = QP^2 \tag{6.1}$$

where P is the polarization. For an isotropic material, this induced polarization is in the same direction as the applied field and is given by

$$P = \chi E \tag{6.2}$$

where χ is the susceptibility.
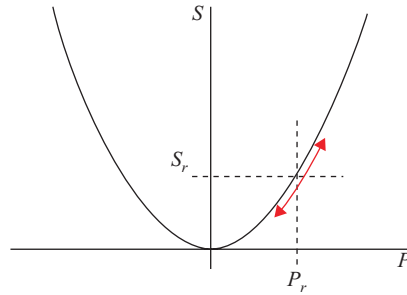
Figure 6.11. Schematic diagram of quadratic electrostriction. The approximately linear small-field piezoelectric response of a polar material is indicated by the arrowed line.

Ferroelectric materials are, however, polar and possess a polarization even in the absence of an electric field. In this case the electrostriction equation in terms of polarization still applies, but is offset by the remanent polarization. This is shown schematically in Figure 6.11.

The polarization at zero electric field is the remanent polarization, P_r, and is associated with a remanent strain, S_r, through the electrostriction equation. This is illustrated by experimental results for a commercial soft PZT shown in Figure 6.12.

Above the Curie temperature, the response matches very closely a quadratic electrostriction. Below the Curie temperature, it still follows a quadratic electrostriction where the electric field is aligned with the poling field, with a remanent polarization of just under 0.3 C m^{-2} at room temperature.

The stability in the position of a piezoelectric material is therefore determined by a combination of the remanent strain (coupled by electrostriction to the remanent polarization) and the underlying thermal expansion of the material. Drift in position of a material with temperature is shown in Figure 6.13.

The effect of this as drift in the zero field (remanent) strain is shown in Figure 6.14, which shows thermal expansion orthogonal to the poling direction. It is clear that there is no simple single value for the thermal expansion coefficient.

For expansion orthogonal to the poling direction, the expansion due to changes in polarization act in the same direction as the underlying thermal expansion. For expansion in the poling direction, the two effects act in opposite directions. Decreasing polarization leads to contraction in poling direction which opposes, and can dominate, the underlying thermal

(a)

(b)
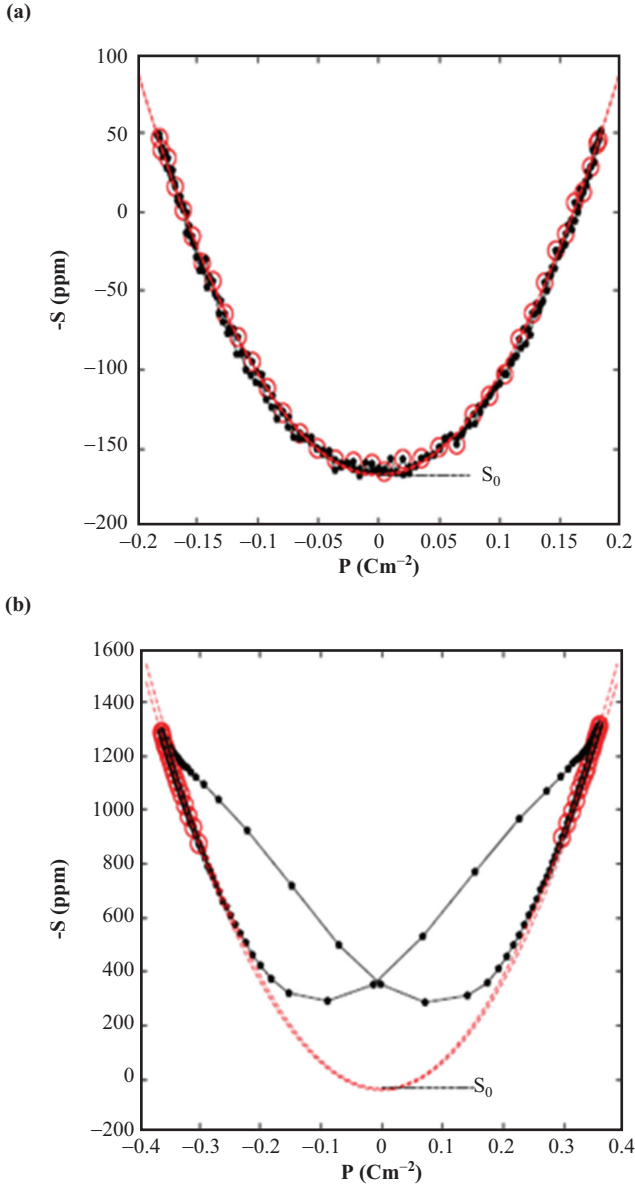
Figure 6.12. Experimental results showing strain in a ferroelectric (a) above the Curie temperature (180°C) and (b) below the Curie temperature (19°C) (Weaver, Cain, and Stewart 2010). Points where the electric field is aligned with the poling direction are marked as circles.
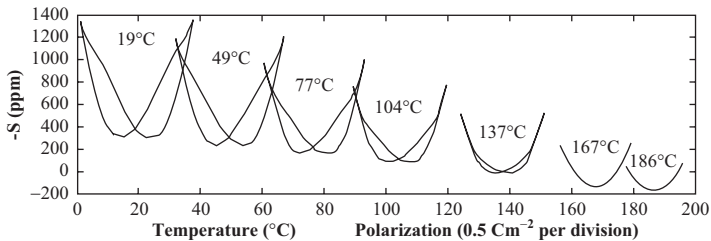
Figure 6.13. Strain–polarization loops at different temperatures. The x-axis values for each loop (electric field or polarization) were offset to center each loop on the temperature at which it was measured.
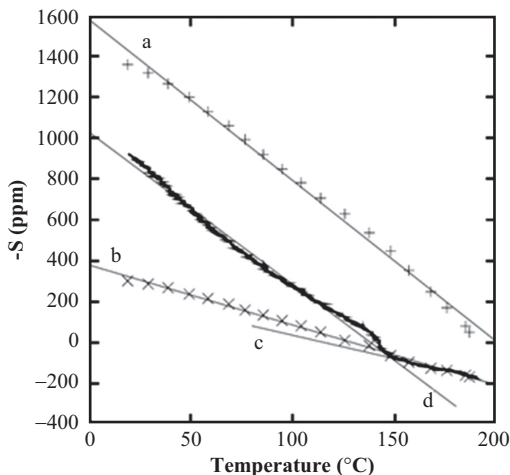
Figure 6.14. Thermal expansion orthogonal to the poling direction during temperature cycling with bipolar loops (a) is a straight-line fit to the loop maxima (+) and (b) is a straight-line fit to the loop minima (×) and (c) is a straight-line fit to the rema-nent strain (solid line) for T > 148°C, while (d) is a fit to the portion of the curve T < 138°C.

expansion. This results in an apparently negative thermal expansion below the Curie temperature (Rusek et al. 2008).

The interaction between changes in polarization and the underlying thermal expansion can be expressed as (Weaver, Cain, and Stewart 2010):

$$S(T, P) = S_0(T) + QP^2 \tag{1.3}$$

where S is the strain, T is the temperature, and P is the polarization. S_0 is the underlying lattice strain, which is subject to normal thermal expansion, and Q is the electrostriction coefficient. In general, Q will be temperature dependent, but it is often treated as independent of temperature. The values of S_0 and Q can be obtained by fitting to the experimental strain curves such as those shown in Figure 6.12. This can be experimentally challenging as measurements of both strain and polarization must be made relative to some value at a reference temperature. Materials that are good electrical insulators at room temperature can become significantly conductive at high temperatures, resulting in charge drift due to leakage currents. This can present serious challenges to tracking the ferroelectric polarization at high temperature. However, at the moderate temperatures used in Figure 6.13, charge drift was not a major problem, and it was possible to obtain values of the electrostriction coefficient and underlying expansion for this particular material as shown in Figure 6.15.

6.3.2 THERMAL EXPANSION IN BIMORPH ACTUATORS AND ELECTRONIC COMPENSATION

These thermal expansion effects will be observed directly in monolithic or stack actuators, but for many applications, some means of amplifying the piezoelectric strain is used to obtain larger movement from the actuator. A very common configuration is to join the ceramic to a metal, called a unimorph actuator (Figure 6.16), or to another ceramic to make a bimorph actuator. In either case, application of an electric field induces curvature in the composite structure, resulting in a deflection at the tip of the actuator that can be several mm.

For a unimorph actuator, the choice of substrate material is crucial in achieving the best temperature stability and usually some form of low expansion alloy is used to match the expansion of the ceramic. Due to the complexity of the ceramic expansion, this can only ever be approximately achieved and only over a limited temperature range. A further limitation is the increase of remanent polarization with decreasing temperature resulting in a smaller and smaller stroke for unipolar actuation (the electric field applied in the poling direction) as the temperature reduces.

Curve (a) of Figure 6.17 shows good stability in the ON position of the actuator, achieved by accurate matching of the thermal expansion of the substrate. Curve (b) shows the position after returning to zero field, that is, the remanent position. This shows considerable drift with temperature, and a significant reduction in stroke at low temperature. This

a)
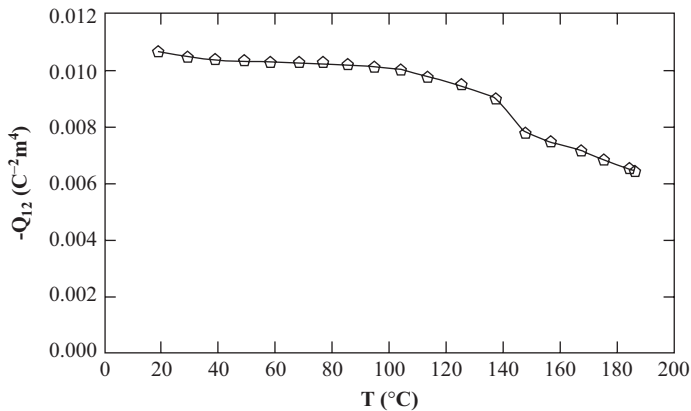
b)
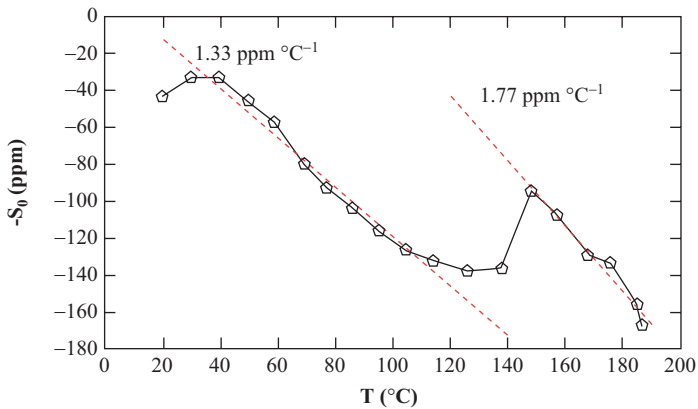
Figure 6.15. Electrostriction coefficient, (a) $Q12$, and (b) strain zero S_0 as functions of temperature (bipolar cycles). In (b) the dashed lines represent straight line fits to the approximately linear portions of the curves.
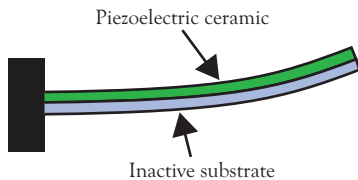
Figure 6.16. Piezoelectric unimorph actuator operating principle.
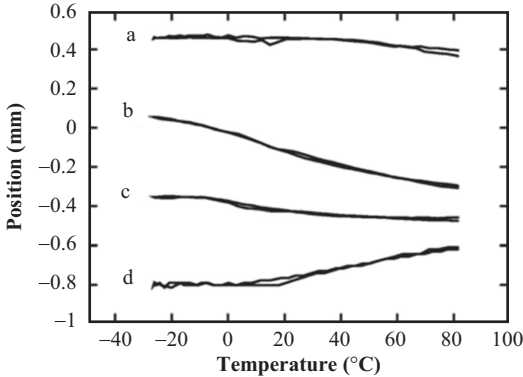
Figure 6.17. Temperature dependence of the position of a unimorph actuator under different electronic compensation regimes (Weaver 2011), (a) is with full field applied in the poling direction, (b) is after returning to zero field, (c) is with a reverse charge applied, while (d) is after once more returning to zero field.

low-temperature stroke would be the limiting factor in the design of, for instance, a valve actuator. Application of an electric field opposing the polarization direction can remove some of this build-up of remanence at low temperature. Figure 6.18 shows how such a reverse field affects the strain in the ceramic and the actuator position.

A moderate reverse field can significantly increase the movement, particularly where the remanence is large. However, too high an electric field exceeds the coercive voltage, resulting in semi-permanent repolarization of the actuator causing it to drive in a direction opposite to that intended. Because the coercive field is strongly temperature dependent (Figure 6.19), applying a fixed reverse field would only work over a fairly restricted temperature range.

What's needed is some kind of adjustable reverse voltage that takes account of the temperature variation of the coercive field to avoid repoling the ceramic. Such a system based on charge measurements is described in Weaver (2011). The application of a temperature-varying reverse field is shown in curve (c) in Figure 6.17 showing a significant reduction in the remanence without repoling at high temperature. Even after the reverse field has been removed, there is still some benefit as shown in curve (d) in Figure 6.17. This technique results in a significant improvement in both stroke and operating window for the actuator.
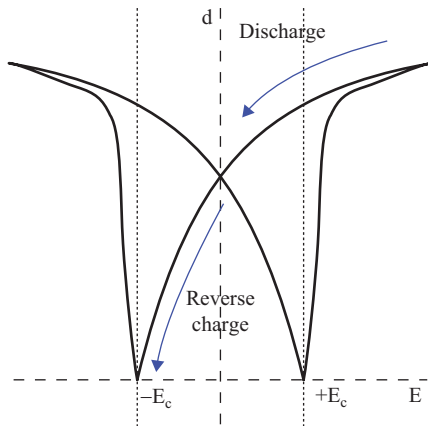
Figure 6.18. Dependence of strain on electric field in a ferroelectric material, showing the coercive field, E_C.
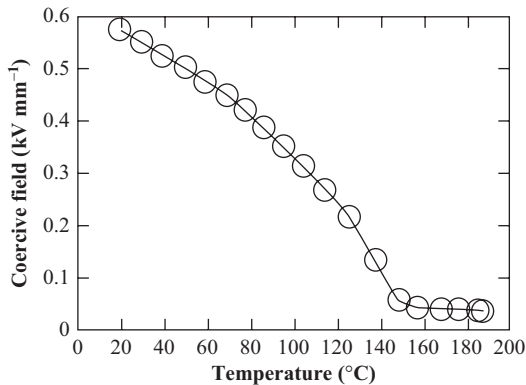
Figure 6.19. Temperature dependence of the coercive field (soft PZT: Fuji Ceramics C91).

6.4 MEASUREMENT TECHNIQUES

6.4.1 THE MEASUREMENT ENVIRONMENT

Harsh environments place considerable demands on measurement equipment, and it is generally necessary to implement some remote or non-contact method to keep the instrumentation out of the test environment. For

electrical measurements, this is usually no problem as far as the instrumentation is concerned as it can be located outside the controlled environment. However, care must be taken to ensure that conductors and their insulation are adequate for the environment and that long lead lengths or electrical loops do not compromise the integrity of electrical signals. For measurements of impedance it is usually necessary to compensate for the resistance and reactance of the leads. At high temperature, mechanical and chemical stability of electrodes and electrical contacts to the sample can be problematic, requiring the use of noble metal electrodes and conductors.

6.4.2 PIEZOELECTRIC MEASUREMENTS IN HARSH ENVIRONMENTS

Piezoelectric coupling involves the conversion from electrical to mechanical energy and vice versa, so measuring of piezoelectric properties directly would require measurement of mechanical as well as electrical properties within the harsh environment. For measuring piezoelectric displacement of actuators, contact probes such as linear variable differential transformer (LVDT) can be used at moderate temperatures, typically <100°C, although higher-temperature probes are available. For higher temperatures, dilatometer methods (up to 1,000°C or above) can be effective for low-frequency piezoelectric measurements (Weaver and Cain 2014). Contact methods such as these are generally only applicable for low frequency and calibration can be difficult.

Optical techniques potentially provide a more robust approach to measurement of mechanical displacement. One approach, described in Weaver et al. (2012), uses a fiber optic probe (supplied by MTI Instruments) to measure piezoelectric displacement in humid environments (Figure 6.20). This method has the advantage that it provides a robust optical fiber-based measurement with no sensing electronic components inside the humidity chamber.

Interferometry is a powerful technique for measuring piezoelectric properties, widely used in bulk and thin film measurements at room temperature (Huang and Leighton 2014; McCartney et al. 2014; Schmitz-Kempen et al. 2013). In principle, the displacement measurement can relate directly to the phase of the interference of light of a known wavelength, providing robust and traceable measurement. Application to high temperatures presents a number of challenges such as ensuring an optically well-characterized environment with optical access to the heated zone, changes in path length and
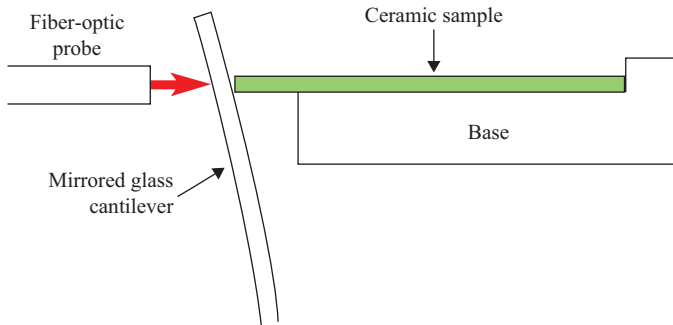
Figure 6.20. Schematic diagram showing the in situ measurement of piezoelectric response using a fiber-optic probe (Weaver et al. 2012).

uncertainty in refractive index due to thermal fluctuations, operation in vacuum (required to minimize thermal effects on path length), and provision of reflective surfaces at high temperature, as well as generic issues relating to electrical and mechanical integrity of the measurement environment. A European Metrology Research Programme (EMRP) project (http://projects. npl.co.uk/METCO) is currently under way to develop and compare methods for high-temperature piezoelectric measurements, with some initial results recently reported (Weaver et al. 2014).

6.4.3 MEASUREMENTS IN HUMID CONDITIONS

Measurements in humid conditions typically proceed by placing the samples in an environmental chamber with controlled temperature and humidity. A common accelerated test condition, widely used in the electronics industry, is 85 percent relative humidity and 85°C. However, particularly when devices are being tested, the high temperature can introduce other failure modes such as creep in adhesive joints, so lower test temperatures may be adopted. A stabilization period such as 30 minutes is usually allowed for equilibration of the sample with the environment before applying the electric field and starting the recording of the leakage current. This helps prevent any effects of condensing moisture when a cold sample is inserted into a hot, humid environment.

Because of the rather variable and stochastic nature of the growth of leakage current (Lipscomb et al. 2009), a large number of samples need to be tested simultaneously to obtain a representative spread of behaviors. Lipscomb et al. (2009) describe an automated multiplexer system for

simultaneous measurement of up to eight samples with 1 nA resolution, taking a complete set of readings every minute. The short-circuit current through a single multiplexer channel in series with the ammeter was 20mA. Currents measured in this study were up to a few tenths of a mA, so the impedance of the measuring system has a negligible effect on the results.

6.5   CONCLUSION

Ferroelectric and piezoelectric materials have many and varied applications in electronics, acoustics, and electromechanical systems. In most situations, they provide robust solid-state performance. As developers find more applications for these materials and push existing applications into new, more extreme environments with demands for higher power, smaller devices and more functionality, the boundaries of materials performance are being extended. To achieve successful new applications for these materials requires in-depth understanding of the ways in which these complex materials interact with their environment and the electrical and mechanical system in which they operate.

This chapter shows how the combination of high electric fields and humidity can compromise the electrical integrity of ferroelectric materials, and describes recent progress in the understanding of the mechanism of this interaction and potential mitigation methods. Ferroelectric materials and their piezoelectric properties are fundamentally tied to the effects of temperature, producing a complex interaction between thermal effects and piezoelectric properties. This affects reliability and performance of piezoelectric devices, so to achieve stability and high performance in applications such as valves requires an understanding of the coupling between piezoelectric and thermal properties, described in this chapter. This is a rapidly developing area with significant interest in using piezoelectric materials, as sensors, actuators, and ultrasound transducers at high temperatures for automotive, aerospace, energy, and medical applications.

To develop our scientific understanding of the performance of ferroelectric materials in harsh environments, as well as supplying reliable engineering data require the ability to measure complex coupled properties in extreme conditions. The chapter therefore concludes with a description of measurement challenges, techniques, and recent developments in this area.

ACKNOWLEDGMENTS

REFERENCES

Baniecki, J.D., J.S. Cross, M. Tsukada, and J. Watanabe. 2002. "H_2O Vapor-Induced Leakage Degradation of $Pb(Zr,Ti)O_3$ Thin-Film Capacitors with Pt and IrO_2 Electrodes." *Applied Physics Letters* 81, no. 20, pp. 3837–39. doi: http://dx.doi.org/10.1063/1.1519359

Bindig, R. and G. Helke. 2000. Application of piezoceramic multilayer actuators, experiences and solutions. *Proceedings on Actuator 2000*, pp. 53–7. Bremen, Germany.

Cain, M. 2014. "Dielectric Breakdown in Dielectrics and Ferroelectric Ceramics." In *Characterisation of Ferroelectric Bulk Materials and Thin Films*, pp. 243–66. Springer.

Chen, W.P., H.L.W. Chan, F.C.H. Yiu, K.M.W. Ng, and P.C.K. Liu. 2002. "Water-Induced Degradation in Lead Zirconate Titanate Piezoelectric Ceramics." *Applied Physics Letters* 80, no. 19, pp. 3587–89. doi: http://dx.doi.org/10.1063/1.1479205

Chen, W.P., X.P. Jiang, Y. Wang, Z. Peng, and H.L.W. Chan. 2003. "Water-Induced Degradation of Barium Titanate Ceramics Studied by Electrochemical Hydrogen Charging." *Journal of the American Ceramic Society* 86, no. 4, pp. 735–37. doi: http://dx.doi.org/10.1111/j.1151-2916.2003.tb03366.x

Dawber, M., K. Rabe, and J. Scott. October 2005. "Physics of Thin-Film Ferroelectric Oxides." *Reviews of Modern Physics* 77, no. 4, pp. 1083–130. doi: http://dx.doi.org/10.1103/revmodphys.77.1083

Huang, Z. and G. Leighton. 2014. "Interferometry for Piezoelectric Materials and Thin Films." In *Characterisation of Ferroelectric Bulk Materials and Thin Films*, ed. M.G. Cain, 15–36, Springer.

Jaffe, B., W. Cook, and H. Jaffe. 1971. *Piezoelectric Ceramics*. London: Academic Press.

Jiang, B., Y. Bai, J.-L. Cao, Y. Su, S.-Q. Shi, W. Chu, and L. Qiao. 2008. "Delayed Crack Propagation in Barium Titanate Single Crystals in Humid Air." *Journal of Applied Physics* 103, no. 11, p. 116102. doi: http://dx.doi.org/10.1063/1.2921136

Jones, B.J., A. Mahendran, A.W. Anson, A.J. Reynolds, R. Bulpett, and J. Franks. 2010. "Diamond-Like Carbon Coating of Alternative Metal Alloys for Medical and Surgical Applications." *Diamond and Related Materials* 19, no. 7–9, pp. 685–9. doi: http://dx.doi.org/10.1016/j.diamond.2010.02.012

Kostic, A., A. Rensch, and D. Sturm. 1995. Nickel Dendrites: A New Failure Mechanism in Ceramic Hermetic Packages. In *Reliability and Maintainability Symposium, 1995. Proceedings, Annual*, 39–43. Washington DC: IEEE.

Ling, H.C. and A.M. Jackson. March 1989. "Correlation of Silver Migration with Temperature-Humidity-Bias (THB) Failures in Multilayer Ceramic Capacitors." *IEEE Transactions on Components, Hybrids, and Manufacturing Technology* 12, no. 1, pp. 130–7. doi: http://dx.doi.org/10.1109/33.19027

Lipscomb, I.P., P.M. Weaver, J. Swingler, and J.W. McBride. 2008. "Micro-Computer Tomography—An Aid in the Investigation of Structural Changes in Lead Zirconate Titanate Ceramics after Temperature-Humidity Bias Testing." *Journal of Electroceramics* 23, no. 1, pp. 72–5. doi: http://dx.doi.org/10.1007/s10832-008-9537-8

Lipscomb, I.P., P.M. Weaver, J. Swingler, and J.W. McBride. 2009. "The Effect of Relative Humidity, Temperature and Electrical Field on Leakage Currents in Piezo-Ceramic Actuators under DC Bias." *Sensors and Actuators A: Physical* 151, pp. 179–86. doi: http://dx.doi.org/10.1016/j.sna.2009.01.017

Lupascu, D.C. 2004. *Fatigue in Ferroelectric Ceramics and Related Issues*, Vol. 61. Heidelberg: Springer Verlag.

Massey, L.K. 2003. *Permeability Properties of Plastics and Elastomers—A Guide to Packaging and Barrier Materials*. 2nd ed. William Andrew Publishing.

McCartney, L., L. Wright, M. Cain, J. Crain, G. Martyna, and D. Newns. 2014. "Methods for Determining Piezoelectric Properties of Thin Epitaxial Films: Theoretical Foundations." *Journal of Applied Physics* 116, no. 1, p. 014104. doi: http://dx.doi.org/10.1063/1.4885058

Mock, R. and K. Lubitz. 2008. "Piezoelectric Injection Systems." *Piezoelectricity, Springer Series in Materials Science* 114, pp. 299–310. doi: http://dx.doi.org/10.1007/978-3-540-68683-5_13

Nagaraj, B., S. Aggarwal, T. Song, T. Sawhney, and R. Ramesh. 1999. "Leakage Current Mechanisms in Lead-Based Thin-Film Ferroelectric Capacitors." *Physical Review B* 59, no. 24, pp. 16022–7. doi: http://dx.doi.org/10.1103/PhysRevB.59.16022

Nagata, K. and J. Thongrueng. 1998. "Effect of Temperature, Humidity and Load on Degradation of Multilayer Ceramic Actuator." *Journal of the Korean Physical Society* 32, pp. 1278–81.

Pertsch, P., S. Richter, D. Kopsch, N. Kramer, J. Pogodzik, and E. Henning. 2006. Reliability of piezoelectric multilayer actuators. In *Actuator 2006*, pp. 527–30. Bremen, Germany.

Pritchard, J., C. Bowen, and F. Lowrie. 2001. "Multilayer Actuators: Review." *British Ceramic Transactions* 100, no. 6, pp. 265–73. doi: http://dx.doi.org/10.1179/096797801681549

Rusek, K., J. Kruczek, K. Szot, D. Rytz, M. Gorny, and K. Roleder. 2008. "Non-Linear Properties of BaTiO$_3$ above T C." *Ferroelectrics* 375, no. 1, pp. 165–9.

Schmitz-Kempen, T., S. Tiedke, P. Mardilovich, S. Sivaramakrishnan, T. Lisec, F. Stoppel, S. Trolier-McKinstry, and P. Muralt. 2013. Compa-

rable measurements and modeling of piezoelectric thin films for MEMS application. In *2013 IEEE International Symposium on the Applications of Ferroelectric and Workshop on the Piezoresponse Force Microscopy (ISAF/ PFM)*, pp. 211–13.

Scott, J.F. 2006. "Nanoferroelectrics: Statics and Dynamics." *Journal of Physics - Condensed Matter* 18, no. 17, pp. R361–86. doi: http://dx.doi.org/10.1088/0953-8984/18/17/r02

Thongrueng, J., T. Tsuchiya, and K. Nagata. 1998. "Lifetime and Degradation Mechanism of Multilayer Ceramic Actuator." *Japanese Journal of Applied Physics* 37, no. Part 1, 9B, pp. 5306–10. doi: http://dx.doi.org/10.1143/jjap.37.5306

van den Ende, D., B. Bos, W. Groen, and L. Dortmans. 2009. "Lifetime of Piezoceramic Multilayer Actuators: Interplay of Material Properties and Actuator Design." *Journal of Electroceramics* 22, no. 1, pp. 163–70. doi: http://dx.doi.org/10.1007/s10832-007-9411-0

Waser, R., T. Baiatu, and K.H. Hardtl. 1990. "dc Electrical Degradation of Perovskite-Type Titanates: I. Ceramics." *Journal of the American Ceramic Society* 73, no. 6, pp. 1645–53. doi: http://dx.doi.org/10.1111/j.1151-2916.1990.tb09809.x

Weaver, P. and M. Cain. 2014. "Temperature Dependence of Ferroelectric and Piezoelectric Properties of PZT Ceramics." In *Characterisation of Ferroelectric Bulk Materials and Thin Films*, ed. M.G. Cain, pp. 115–46. Springer.

Weaver, P.M. April 2011. "A Sensorless Drive System for Controlling Temperature-Dependent Hysteresis in Piezoelectric Actuators." *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control* 58, no. 4, pp. 704–10. doi: http://dx.doi.org/10.1109/tuffc.2011.1863

Weaver, P.M., C. Baldauf, T.J. Stevenson, T. Quast, G. Bartl, T. Schmitz-Kempen, M. Cain, and M. Stewart. 2014. "High Temperature Piezoelectric Actuators – Materials and Measurement. In *Proceedings on Actuator 2014*, pp. 60–3, Bremen, Germany.

Weaver, P.M., M.G. Cain, and M. Stewart. 2010. "Temperature Dependence of High Field Electromechanical Coupling in Ferroelectric Ceramics." *Journal of Physics D: Applied Physics* 43, no. 16, p. 165404. doi: http://dx.doi.org/10.1088/0022-3727/43/16/165404

Weaver, P.M., M.G. Cain, M. Stewart, A. Anson, J. Franks, I.P. Lipscomb, J.W. McBride, D. Zheng, and J. Swingler. 2012. "The Effects of Porosity, Electrode and Barrier Materials on the Conductivity of Piezoelectric Ceramics in High Humidity and dc Electric Field." *Smart Materials and Structures* 21, no. 4, p. 045012. doi: http://dx.doi.org/10.1088/0964-1726/21/4/045012

Zheng, D., J. Swingler, and P. Weaver. 2010. "Current Leakage and Transients in Ferroelectric Ceramics under High Humidity Conditions." *Sensors and Actuators A: Physical* 158, no. 1, pp. 106–11. doi: http://dx.doi.org/10.1016/j.sna.2009.10.021

About the Contributing Authors

Dr. Jonathan Swingler received a Joint BSc (Hons) in Physics and Chemistry from Keele University in 1990 followed by a PhD at Loughborough University for his work on the degradation of electrical contacts. He subsequently moved to the University of Southampton where he pursued his research into the physics of degradation and reliability of electrical and electronic materials and devices. Currently Jonathan is concerned to develop reliability engineering science at Heriot-Watt University to enable the prediction of electrical component and system reliability under harsh operating conditions. He is an Associate Professor of Energy at Heriot-Watt University and Fellow of the Institute of Physics.

Dr. Jeff Jones is an Associate Professor at WMG with research interests in technical asset management and dependability. The focus of his research is the use of models, simulations, and artificial intelligence, and these have been applied to large multifaceted systems down to MEMS devices. He is the chair of IEC TC/56 WG2 on Dependability Techniques, and is principal UK expert on a number of drafting committee within IEC. He has recently been awarded the 1906 award by IEC for his work in standardization.

Dr. Michael D. Bryant, Accenture Endowed Professor of Mechanical Engineering, specializes in tribology, the design of mechatronic systems and system diagnostics of machinery. Dr. Bryant is a faculty member of the Dynamic Systems & Control and Manufacturing & Design programs of the Mechanical Engineering Department. Dr. Bryant earned his Ph.D. in engineering sciences and applied mathematics from Northwestern University in 1981. He has served on the Cockrell School of Engineering faculty since 1988 and previously taught at North Carolina State University. He has published more than 100 technical articles and reports. In 2005, he was appointed Editor in Chief of the ASME *Journal of Tribology* and Series Editor for *Dynamic Systems & Control and Mechatronics*, Springer Mechanical Engineering Series. He is a Fellow of ASME and member or IEEE.

Dr. Alec Feinberg is the founder of DfRSoft. He has a Ph.D. in Physics and is the principal author of the book, Design for Reliability. Alec has provided reliability engineering services in all areas of reliability including solar, thin film power electronics, defense, microelectronics, aerospace, wireless electronics, and automotive electrical systems. He has provided training classes in Design for Reliability, Shock and Vibration, Quality, Accelerated Testing, HALT, Reliability Growth, Electrostatic Discharge, Dielectric Breakdown, DFMEA and Thermodynamic Reliability Engineering. Alec has presented numerous technical papers and won the 2003 RAMS Alan O. Plait best tutorial award for the topic, "Thermodynamic Reliability Engineering."

Dr. Xiandong Ma is a Senior Lecturer in the Engineering Department at Lancaster University, UK. His previous career included Nanjing Automation Research Institute as a Senior Engineer, Lancaster University and then The University of Manchester as a Postdoctoral Researcher and ALSTOM Power Ltd., UK, as a Condition Monitoring Engineer. He has developed his broad range of research interests in condition monitoring of insulating material degradation, power station generators, and latest wind renewables. He has developed new generations of electromagnetic tomographic imaging system for metal production process monitoring. Dr. Ma has published over 80 refereed journal and international conference papers. He is a member of the IET and a Chartered Engineer.

Dr. Paul M. Weaver is principal research scientist at the UK's National Physical Laboratory. Paul's research includes the application of piezoelectric and ferroelectric materials for novel sensing, actuation, energy storage, and energy harvesting applications. A particular aspect of Paul's work is the science and measurement of behaviour under harsh environments such as elevated temperature or high humidity. Paul graduated with an MA degree in natural science from Cambridge University, and a PhD from Southampton University. He is a chartered engineer, member of the IET and is a visiting Reader at Southampton University.

INDEX

The Physics of Degradation in Engineered Materials and Devices

Fundamentals and Principles

Edited by Jonathan Swingler with major contribution by Alec Feinberg

Degradation is apparent in all things and is fundamental to both manufactured and natural objects. It is often described by the second law of thermodynamics, where entropy, a measure of disorder, tends to increase with time in a closed system. Things age!

This concise reference work brings together experts and key players engaged in the physics of degradation to present the background science, current thinking and developments in understanding, and gives a detailed account of emerging issues across a selection of engineering applications. The work has been put together to equip the upper level undergraduate student, postgraduate student, as well as the professional engineer and scientist, in the importance of physics of degradation. The aim of *The Physics of Degradation in Engineered Materials and Devices* is to bridge the gap between published textbooks on the fundamental science of degradation phenomena and published research on the engineering science of actual fabricated materials and devices.

A history of the observation and understanding of physics of degradation is presented and the fundamentals and principles of thermodynamics and entropy are extensively discussed. This is the focus of this book, with an extended chapter by Alec Feinberg on equilibrium thermodynamic damage and non-equilibrium thermodynamic damage. It concludes with two particular technologies to give examples of areas of application.

Dr. Jonathan Swingler received a Joint BSc (Hons) in physics and chemistry from Keele University in 1990, followed by a PhD at Loughborough University for his work on the degradation of electrical contacts. He subsequently moved to the University of Southampton where he pursued his research into the physics of degradation and reliability of electrical and electronic materials and devices. Currently Jonathan is working to develop reliability-engineering science at Heriot-Watt University, to enable the prediction of electrical component and system reliability under harsh operating conditions. He is an associate professor of energy at Heriot-Watt University and fellow of the Institute of Physics.

MOMENTUM PRESS
ENGINEERING