Modeling of Materials with Internal Variables
Using a Thermomechanical Approach

by

Xiaodong Zhang
Thesis submitted to the faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE
in
Mechanical Engineering

APPROVED:

[Signatures]
Dr. C.A. Rogers, Chairman
Dr. H.H. Robertshaw
Dr. C. Liang

December, 1992
Blacksburg, Virginia
Modeling of Materials with Internal Variables Using a Thermomechanical Approach

by

Xiaodong Zhang

Committee Chairman: Dr. C.A. Rogers

Mechanical Engineering

Abstract

In this thesis, the thermomechanical approach with internal variables has been thoroughly analyzed. This approach is based on the combination of thermodynamic principles and continuum mechanics. Therefore it reflects the physical essence of constitutive behavior of materials. Based on this approach, a one-dimensional constitutive model for the two-way shape memory effect and a one-dimensional constitutive model for piezoceramics have been developed, respectively.

In modeling the two-way shape memory effect, a residual stress $\sigma_{re}$ is introduced as a controlling parameter for the two-way shape memory effect. A further refinement of the transformation kinetics expression for two-way shape memory is derived. It is demonstrated that the material parameters required by this model can be calculated or measured using a standard materials testing apparatus. A numerical study is conducted and the effectiveness of this model is verified.

In the constitutive modeling of piezoceramics, a new internal state variable is introduced to relate the macroscopic behavior of a piezoceramic with its micro-properties. A phenomenological formulation of polarization reversal is proposed, and then a fully-coupled thermo-electro-mechanical model is developed. It is shown that the theory developed can describe the electromechanical behavior of piezoceramics well.
Acknowledgments

I would like to express my gratitude to all my committee members for all their advice and suggestions throughout my study. I am grateful to my committee chairman, Dr. Craig Rogers, for his support and encouragement during my study, especially at my most difficult time, which made this thesis possible. I would also like to thank Dr. Chen Liang for his helpful discussions.

I would like to thank all the staff and students in the Center for Intelligent Material Systems and Structures for their generous help during the past two years. I would especially like to thank Dr. Zaffir Ahmed Chaudhry for his constructive discussions and comments on the writing of the whole thesis. I would also like to thank Ms. Beth Howell for her proof-reading of my whole thesis and papers.

I would like to gratefully acknowledge the financial support provided by the Air Force Office of Scientific Research, grant No. AFOSR-91-0416.

Finally, I would like to thank my parents who encouraged and supported my high school and undergraduate studies.
Contents

1 Introduction ........................................ 1
   1.1 Introduction to Intelligent Material Systems .......... 3
   1.1.1 Actuators ...................................... 4
   1.1.2 Sensors .......................................... 5
   1.1.3 Controllers ...................................... 7
   1.2 Approach ........................................... 8
   1.3 Introduction to Shape Memory Alloys .................. 11
   1.4 Review of Piezoelectricity ........................... 14
      1.4.1 Historical Introduction ......................... 15
      1.4.2 Background on the Theory of Piezoelectricity ... 18
      1.4.3 General Considerations Regarding Piezoelectric Ceramics ... 30
      1.4.4 Phenomenological Modeling of Piezoelectricity .......... 37
      1.4.5 Applications .................................... 47
   1.5 Summary ............................................ 56

2 Theory of the Thermomechanical Approach ............... 58
   2.1 Introduction ....................................... 59
   2.2 Formulations of Thermomechanical Equations with Continuum Mechanics ....................... 61
   2.3 Axioms for Formulating Constitutive Equations .......... 65
2.4 Thermodynamic Restrictions ........................................... 76
2.5 Generalized Thermodynamic Forces and Fluxes .................. 80
2.6 Thermomechanical Approach with Internal Variables .......... 83

3 Modeling of the Two-Way Shape Memory Effect .................. 88
   3.1 Introduction .................................................................. 89
   3.2 The Mechanism and Training of the Two-Way Shape Memory Effect........ 91
       3.2.1 Mechanism of the Two-Way SME ............................... 91
       3.2.2 Training of the Two-Way SME ................................. 94
   3.3 Transformation Kinetics of Thermoelastic Martensitic Transformation 100
   3.4 Constitutive Model of Two-Way SME Alloys ....................... 106
       3.4.1 Constitutive Equations ......................................... 106
       3.4.2 Determination of the Material Parameters ................. 111
   3.5 Numerical Illustration .............................................. 114
   3.6 Concluding Remarks ................................................ 117

4 A Macroscopic Phenomenological Formulation for Coupled Electromechanical Effects in Piezoelectricity ............ 119
   4.1 Introduction ................................................................ 120
   4.2 Preliminary Considerations ......................................... 124
   4.3 Domain Dynamics ..................................................... 128
       4.3.1 Concepts of Domain Dynamics ................................. 128
       4.3.2 Basic Theories of Polarization Reversal Dynamics ........ 129
       4.3.3 Phenomenological Formulation of the Polarization Reversal Dynamics for Ceramics ................................. 135
   4.4 Thermodynamic Formulation for Coupled Electromechanical Effects in Piezoelectricity ............................... 139
4.5 A Special Case ..................................................... 144
4.6 Concluding Remarks ............................................. 145

5 Conclusions and Recommendations .................................. 147
  5.1 Conclusions .......................................................... 147
  5.2 Recommendations .................................................. 149

References ................................................................. 151
Vita ........................................................................ 164
List of Figures

1.1 One-Way and Two-Way Shape Memory Effect. .......................... 12
1.2 Interaction Processes Between the Electrical, Mechanical, and Thermal
1.3 Piezoelectric Phenomena .............................................. 22
1.4 Unit Cell of BaTiO₃ with a Perovskite Cubic Structure. ............. 31
1.5 Phase Diagram of Pb(Ti₁₋ₓZrₓ)O₃ System (after Jaffe et al., 1971) .... 32
1.6 180° and 90° Domain Configurations in Barium Titanate (after Moul-
    son and Herbert, 1990). .............................................. 34
1.7 Poling in a Two-dimensional Ceramic. .................................. 36
1.8 Variation of Elastic Gibbs Function with Polarization at Different Tem-
    peratures. ............................................................. 41
1.9 Reduced Curves of Polarization vs. Field at Different Temperatures
    (after Merz, 1953). ................................................... 43
1.10 Piezoelectric Ignition Devices. ........................................ 49
1.11 Cantilever Bimorphs. ................................................ 51
1.12 Sandwich Transducers. ............................................... 53
1.13 Schematic Representation of a Composite with 1-3 Connectivity. .... 55
3.1 Schematic Illustration of the Training and the Components of Strain. 95
3.2 Changes in Stress-Strain-Temperature Behavior During the Training Process (after Escher et al.) 97
3.3 Schematic Stress-Strain Curves at Different Deformation Temperatures 99
3.4 Typical Thermocycling Curves of Martensitic Fraction vs. Temperature 101
3.5 Influence of Applied Stress on Martensitic Transition Temperature 104
3.6 Typical Electrical Resistivity vs. Temperature Curves For a Deformed SMA 107
3.7 Young's Modulus of Nitinol vs. Temperature (after Cross et al., 1970) 113
3.8 Free Recovery of a One-Way Strain and a Two-Way Strain 115
3.9 Influence of the Residual Stress on the Free Recovery 118
4.1 Schematic Diagram of a Typical Hysteresis Loop for Piezoelectric Ceramics 121
4.2 Dielectric Hysteresis Loop for Different Biasing Stresses 123
4.3 Remanent Polarization $P_r$, Coercive Field $E_c$, and Ratio $P_r^3/E_c$ Shown as Functions of Temperature for PLZT Ceramics (after Gerthsen and Krüger, 1976) 125
4.4 Variation of Spontaneous Polarization of $BaTiO_3$ Ceramics at 70°C with Grain Size (after Shaikh and Vest, 1989) 134
4.5 Typical Hysteresis Loops Calculated from the Proposed Phenomenological Polarization Reversal Model 137
4.6 The Frequency Dependence of Coercive Field $E_c$ for PLZT Ceramics (after Gerthsen and Krüger, 1976) 140
4.7 Comparison of Experimental Data (after Anderson and Crawley, 1989) and Theoretical Prediction 146
List of Tables

1.1 Comparison of TiNi and CuZnAl Alloys .......................... 23
1.2 Types of Fundamental Piezoelectric Relations (after Ikeda, 1990) .......................... 27
1.3 Typical Values of the Properties of Some Piezoceramics .................. 38

2.1 Thermodynamic Force and Corresponding Thermodynamic Flux .................. 82

3.1 Mechanical Properties for Computer Simulation .......................... 116
Nomenclature

A  
Austenite phase

$A_0$  
Mean average of $A_s$ and $A_f$ (Chapter 3), Cross-sectional area of a thin ferroelectric surface (Chapter 4)

$A_e$  
Actual area covered by domains at time $t$

$A_{ex}$  
Extended area covered by domains at time $t$

$A_f, A_s$  
Austenite finish and start temperatures

$A_f^e, A_s^e$  
Mechanical austenite finish and start temperatures

$C_A, C_M$  
Stress influence coefficient

$C_\sigma, k_\sigma$  
Stress-state factors

$C_T, k_T$  
Temperature factors

$C_\omega, k_\omega$  
Frequency factors

$C_h, k_h$  
Geometric factors

$C_d, k_d$  
Grain-size factors

d  
Diameter of grain size

d_{ij}  
Piezoelectric coefficients

D  
Elastic modulus (Chapter 3), electric displacement (Chapter 4)

$D_A$  
Elastic modulus of austenite

$D_M$  
Elastic modulus of martensite

1Symbols which are not listed here are specified wherever they are used.
2Symbols without a specifying chapter were used throughout the thesis.
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E$</td>
<td>Applied electric field</td>
</tr>
<tr>
<td>$E_0$</td>
<td>Amplitude of sinusoidal field</td>
</tr>
<tr>
<td>$E_c$</td>
<td>Coercive field</td>
</tr>
<tr>
<td>$E_L$</td>
<td>Local electric field</td>
</tr>
<tr>
<td>$E_\infty$</td>
<td>Coefficient of coercive field</td>
</tr>
<tr>
<td>$f$</td>
<td>Deformation gradient</td>
</tr>
<tr>
<td>$h$</td>
<td>Thickness of specimen</td>
</tr>
<tr>
<td>$h_0$</td>
<td>Surface-layer thickness of specimen</td>
</tr>
<tr>
<td>$i$</td>
<td>Switching current</td>
</tr>
<tr>
<td>$i_{\text{max}}$</td>
<td>Maximum switching current</td>
</tr>
<tr>
<td>$k, k'$</td>
<td>Newiy introduced material parameters</td>
</tr>
<tr>
<td>$K'$</td>
<td>Cyclic strength coefficient</td>
</tr>
<tr>
<td>$L$</td>
<td>Deformation velocity</td>
</tr>
<tr>
<td>$M$</td>
<td>Martensite phase</td>
</tr>
<tr>
<td>$M_0$</td>
<td>Mean average of $M_f$ and $M_s$</td>
</tr>
<tr>
<td>$M_\text{d}$</td>
<td>Maximum temperature at which the SIM can be created</td>
</tr>
<tr>
<td>$M_f$, $M_s$</td>
<td>Martensite finish and start temperatures</td>
</tr>
<tr>
<td>$M_f^p$, $M_s^p$</td>
<td>Mechanical martensite finish and start temperatures</td>
</tr>
<tr>
<td>$n$</td>
<td>Transformation coefficient</td>
</tr>
<tr>
<td>$n'$</td>
<td>Cyclic strain-hardening exponent</td>
</tr>
<tr>
<td>$N$</td>
<td>Number of aligned dipoles</td>
</tr>
<tr>
<td>$N_1$</td>
<td>Number of domain nucleation sites</td>
</tr>
<tr>
<td>$P$</td>
<td>Plastic modulus (Chapter 3), instantaneous polarization at time $t$ (Chapter 4)</td>
</tr>
<tr>
<td>$F_{\text{cer}}$</td>
<td>Spontaneous polarization of ceramics</td>
</tr>
<tr>
<td>$P^{\text{int}}$</td>
<td>Introduced polarization as internal variable</td>
</tr>
</tbody>
</table>
$P^r$  Reversible polarization
$P_s$  Spontaneous polarization
$P_{sc}$  Spontaneous polarization of a single crystal in a certain direction
$Q$  Electrostriction constant
$q$  Internal heat source
$q_{sur}$  Heat flux from the surroundings
$S$  Entropy density
$t$  Time
$t_s$  Domain switching time
$t_{max}$  Domain switching time corresponding to $t_{max}$
$T$  Temperature
$T_M, T_A$  Initial state temperatures of SMA
$T_q$  Quenching temperature of steel
$U$  Internal energy density
$v$  Sideways wall velocity
$v_\infty$  Coefficient of sideways wall velocity
$X$  Original coordinate system
$z$  Current coordinate system
$Y$  Young's modulus
$\alpha$  Activation field for sideways wall motion
$\beta$  Correcting coefficient for calculating coercive field
$\nu$  Rate constant
$\epsilon, \epsilon$  Engineering strain
$\epsilon_e$  Elastic strain limit
$\epsilon_{pl}$  Residual plastic strain
$\bar{\epsilon}, \bar{\epsilon}$  Green strain
\( \varepsilon_L \) Maximum recoverable strain or recovery strain limit

\( \varepsilon_{res} \) Martensitic residual strain

\( \omega \) Angular frequency of a sinusoidal field

\( \Gamma \) Electromechanical tensor

\( \Theta \) Thermoelastic tensor

\( \xi, \Xi \) Martensite fraction (Chapter 3), relative polarization (Chapter 4)

\( \xi_A, \xi_M \) Initial martensite fractions of SMA

\( \Xi \) Electrical polarization tensor

\( \rho \) Mass density in current coordinate system

\( \rho_0 \) Mass density in original coordinate system

\( \sigma, \bar{\sigma} \) Engineering stress

\( \sigma, \bar{\sigma} \) The second Piola-Kirchhoff stress

\( \sigma_e \) Elastic stress limit

\( \sigma_{\text{ap}} \) External applied stress

\( \sigma_{\text{rc}} \) Residual stress of trained SMA

\( \Omega \) Transformation tensor (Chapter 3), polarization reversal tensor (Chapter 4)

\( \Phi \) Helmholtz free energy

\( \Pi \) Thermoelectrical tensor

\( \Upsilon \) Electrical tensor

---

**Subscripts and Superscripts**

\( 0 \) Initial condition
1w  One-way effect
2w  Two-way effect
A   Austenite
cer Ceramics
d   Grain size
e   Elastics
h   Geometric
int Internal variable
M   Martensite
pe  Pseudoelastic
pl  Plastics
r   Reversal
re  Residual
T   Training (Chapter 3), Temperature (Chapter 4)
ω   Frequency
σ   Stress
Chapter 1

Introduction

The science and technology issues related to advanced materials are playing a more and more important role as the driving force for technological innovations. The study of intelligent material systems and structures is one of the research areas that is emerging as advanced materials take a more important role in technological advancements. An intelligent material system is composed of four basic elements: a skeletal structure; a nervous system which is a network of embedded or attached sensors; a muscular system that is a network of muscle-like actuators such as shape memory alloy (SMA) fibers or piezoelectric chips; and a control system.

As far as the material issues are concerned, shape memory alloys and piezoelectric ceramics are the most widely researched materials in intelligent material systems. Shape memory alloys have been widely used as actuator materials in intelligent material systems. SMA fibers embedded in composite structures can be used to modify or control the static and dynamic structural response by actively changing the stiffness of the structure or adjusting the stress distributions of the structure (Rogers and Robertshaw, 1988). Piezoceramics have been explored as both actuator materials and sensors. The piezoceramic material changes its dimensions when an electric field is applied; hence, such a material can be bonded externally or embedded internally into
the composite structure to act as an actuator. Conversely, an electric charge may be accumulated on the surface of a piezoceramic material when a load is applied; therefore, such a material can also be used as a sensor.

In order to design an intelligent material system, one must understand the fundamental principles of sensing and actuating. Studies of the constitutive behavior of piezoceramics and the shape memory effect are, therefore, essential to the research of intelligent material systems and structures.

An irreversible thermodynamic approach to modeling materials with internal variables, i.e., the thermomechanical approach, may provide a suitable method of describing and predicting the constitutive behavior of shape memory alloys and piezoceramics. As we know, there are two approaches to modeling the constitutive relations of a material. One is the macroscopic phenomenological approach which is based on macroscopic experimental results. The other is the microscopic physical method which derives the constitutive relation from fundamental physical concepts. The phenomenological approach is often used in engineering practice, but can rarely provide the physical essence of the material behavior. Alternatively, the microscopic method can explain the experimental phenomena physically, but is far from quantitative. Therefore, a combination of the two approaches will perhaps give a better prediction of the material behavior. In the thermomechanical approach, a set of internal variables will be introduced to describe the constitutive behavior of the material by incorporating the internal variables into the thermodynamic principles of continuum media. It is the internal variable that relates the microscopic physical properties with its macroscopic behavior phenomenologically. Accordingly, the thermomechanical approach will be studied in detail in modeling the constitutive behavior of the
shape memory alloys and piezoceramics.

There are five chapters in this thesis. The first chapter presents the introduction and literature review; the second details the theory of the thermomechanical approach to provide a general background theory for modeling constitutive behavior of materials; the third describes an application of the thermomechanical approach in modeling the two-way shape memory effect; the fourth discusses the use of the thermomechanical approach in formulating the one-dimensional constitutive behavior of piezoceramics; the last chapter presents the conclusions and recommendations.

1.1 Introduction to Intelligent Material Systems

At the present time, there is no universally accepted definition of an intelligent material system. But the implications of an intelligent material system may be stated from different viewpoints. Technologically an intelligent material system refers to the integration of actuators, sensors, and controls with a material or structural component. This is essentially an anatomical description of an intelligent material system: actuators or motors acting like muscles; sensors behaving as a nervous system; communications and computational networks serving as the brain of a biological system; and a host material representing a skeletal structure of a biological system. Scientifically, an intelligent material system may be defined as a material system with intelligence and life features integrated in the microstructure of the material system to reduce mass and energy and produce adaptive functionality (Rogers, 1992).

As discussed above, an intelligent material system is an integration and interaction of actuators, sensors, host materials and control systems. In order to design a desired
intelligent material system, it is essential to understand the individual and the interactive behavior of each constituent part. The constituent components with 'smart' characteristics are sensors, actuators and controllers. The essence of 'smart' is based on the convertible interactions of material characteristics and external stimuli. Actuator materials have the ability to change the shape, stiffness, position, natural frequency, damping, and other mechanical characteristics of intelligent material systems in response to changes in temperature, electric field, or magnetic field. Contrarily, sensor materials have the ability to feedback stimuli such as thermal, electrical, or magnetic signals to the control systems in response to changes in the mechanical characteristics of intelligent material systems.

1.1.1 Actuators

The typical actuator materials are shape memory alloys, piezoelectric materials, magnetostrictive materials, thermally controllable materials, electrorheological fluids, and magnetorheological fluids. Among these, the most important actuator materials are shape memory alloys and piezoelectric ceramics. Shape memory alloys can exhibit recoverable strains of up to 8% by heating them above the phase transformation temperature. The heating is usually executed by applying an electric current. In the process of returning to its original shape, the SMA can produce a very large force which is used for actuation of SMA actuators. The response of SMA actuators is slow because actuation is dependent on heating and cooling. NiTi alloys are the most developed SMA actuator materials with high corrosion resistance and large recovery strains. Other shape memory alloys such as Cu-based alloys and ferrous alloys of Fe-Pt, Fe-Ni-C and Fe-Ni-Co-Ti type may also be used as actuator materials in the cases where cost is important and serving condition is not critical.
Although shape memory effects have been extensively researched for decades, it is only in recent years that shape memory alloys have been explored for actuator applications. Nitinol has been used in Japan in micromanipulators and robotic actuators to mimic the smooth motions of human muscles. Other applications include general-purpose linear actuators that incorporate a helical spring made of Nitinol, thermal actuators which automatically open and close the vents on storage buildings, engine mounts and suspension systems which control vibration. In addition, it has been suggested that vibration control and active damage control can be accomplished with hybrid adaptive structures in which the SMA actuators are embedded inside the composite material (Rogers, 1992).

Piezoelectric actuators can provide mechanical driving forces in different directions in response to an applied electric field. That piezoelectric materials can change shape is due to the reversibility of electrical dipoles within the materials and their crystallographic characteristics. It is reported that maximum strains of 200-300 microstrain can be produced. Piezoelectric actuators can be used in the cases of high precision, high speed (10 μs) or high generative force (400 kgf/cm²). They are finding applications in optical tracking devices, magnetic heads, deformable mirrors, micropositioners for robots, ultrasonic motors and speakers (Rogers, 1992; Uchino, 1986)

1.1.2 Sensors

Sensing is one of the critical parts of intelligent material systems. Sensors can be attached on the surface of a smart structure or embedded in it to provide accurate information to the controllers. Sensor materials include optical fibers, piezoelectric
materials, and tagging particles. Optical fiber material is widely used as sensors in various cases extrinsically or intrinsically. Extrinsic use of an optical fiber sensor involves the fact that it merely transmits light. Intrinsic optical sensors depend on changes in the light transmission characteristics of the optical fibers. (Rogers, 1992). Applications of optical fiber sensors include security systems, nondestructive materials evaluation, in-service structural health monitoring, damage detection, and composite cure monitoring.

Piezoelectric materials are also widely used as sensors in intelligent material systems. Piezoelectric sensors are based on the fact that these materials generate an electrical response in response to an applied force or pressure. They are usually made from piezoelectric polymers such as polyvinylidene fluoride (PVDF) due to the brittle nature of piezoelectric ceramics. PVDF can be formed into very thin films and attached to many surfaces to measure stresses in one direction or in a plane. The high sensitivity of PVDF to pressure changes has been tailored to make tactile sensors in robotics that can recognize the letters of the braille alphabet, different grades of sandpapers, and even to make skin-like sensors that can replicate the temperature and pressure sensing capabilities of the human skin (Jaeger and Rogers, 1988). In order to improve the thermal and mechanical performance of piezoelectric polymers, piezoelectric composites with 1-3 connectivity (thin PZT rods embedded in a continuous polymer matrix) have been successfully developed to be used as sensors in hydrophones and medical ultrasonic transducers. It is also suggested that polymers containing piezoelectric powders may be explored as sensor materials. Presently a PZT coating is being developed at the Center for Intelligent Material Systems and Structures that may be used in complex structures to monitor and detect damage. In addition, shape memory alloys have potential for sensing applications because of their
thermal response capabilities. An example of these is a prototype Nitinol-actuated automatic water sprinkler for putting out room fires. In this case, as a fire causes the temperature in a room to reach a pre-specified level, the Nitinol element contracts and then activate the sprinkler system (Robinson, 1987).

1.1.3 Controllers

For an intelligent material system with sensors and actuators, feedback signals from sensors must be received and processed by a controller which will provide new instructions to the actuators to guide the responses of the system. The controller is, therefore, crucial to an intelligent material system. A lot of control schemes including modern control approaches, adaptive control and neural networks have been proposed.

Studies of neural network based control systems have become prevalent due to their special features which are not available in conventional control systems (Chaudhry, 1992). The special feature of a neural network is in its ability to learn and modify the control scheme with time. A neural network is basically a massively parallel and interconnected network of basic computing elements which collectively learn to perform complex tasks with high efficiency. Neural-network based controllers are expected to have a most promising future although, at the present time, the suitable forms of neural network-based controllers still do not exist for a fully intelligent material system.
1.2 Approach

Classical thermodynamics deals with studies of the thermal behavior of bodies consisting of a finite number of parts. The principles of thermodynamics have found many applications in different branches of science and technology. It was used to understand the efficiency of heat engines, the electromotive force of galvanic cells, thermal junctions, chemical equilibrium and phase transformations. The basic principles are the first and second laws of thermodynamics. It is well known that the first law is concerned with the balance of heat and work and that the second law is an assertion about the rate of increase of entropy in a system or about the denial of the existance of certain perpetual motion machines. Although thermodynamics is the science of heat and temperature, its principles are often successfully applied to cases in which heat is not flowing or temperature is constant.

Classical continuum mechanics is based on the assumption that stress at a point (X) in a continuum media is dependent on the strain, strain rate, and strain history at the same point (X). It has been remarkably successful in accounting for the properties of deformable bodies.

The thermodynamics of continua is an extension of basic thermodynamic principles to continuum media. Namely, it is an incorporation of basic thermodynamics and classical continuum mechanics. To formulate the thermodynamics of continua, five new basic concepts such as temperature, specific internal energy, specific entropy, heat flux, and heat supply must be added to the concepts of classical continuum mechanics. The thermodynamics of continua has been developed since the 1960s. The underlying assumption of the thermodynamics of continua is that the material
undergoes a thermodynamic process which is defined by a pair of functions associated with each point \( X \): \( T(X,t), \sigma(X,t) \), where \( T \) represents the temperature, \( X \) denotes the original reference, \( \sigma \) is the current configuration and \( t \) is the time. Following this general assumption, a set of constitutive equations will be formulated. These constitutive equations will adhere to the fruitful axioms for formulating constitutive equations and the Clausius-Duhem inequality in terms of continuum mechanics. For different materials, the constitutive equations may have different forms. The detailed principles of formulating material constitutive equations will be addressed in the sequel.

Following these general principles, Coleman (1967) developed a thermodynamic theory with internal variables. The key issue in this theory is to postulate the existence of internal state variables which influence the free energy and whose rate of change is governed by differential equations. More specifically, the basic local mechanical and thermal variables such as stress tensor, the heat flux vector, the Helmholtz free energy and entropy are determined through constitutive equations in terms of a set of state variables such as the temperature \( T \), the deformation gradient or strain \( \varepsilon \), the temperature gradient \( \nabla T \), and a series of \( N \) internal or 'hidden' state variables \((\xi_1, ..., \xi_N)\). It is also assumed that the rate change of each internal variable \( \xi_i \) is governed by a general nonlinear function \( g_i \) of \( T, \varepsilon, \nabla T, \xi_1, ..., \xi_N \):

\[
\dot{\xi}_i = g_i(T, \varepsilon, \nabla T, \xi_1, ..., \xi_N) .
\] (1.1)

The general restrictions placed on the constitutive equations are by the second law of thermodynamics, in the form of Clausius-Duhem inequality. The conditions of the dynamic stability of the above equations are given in terms of the thermodynamic functions. Tanaka (1982, 1985, 1986) extended this thermodynamic theory to the
process of phase transformation, transformation superplasticity and the behavior of shape memory alloys. In Tanaka’s modeling approach, a scaler internal variable $\xi$ which characterizes the extent of the phase transition is introduced. When the phase transformation is among $N + 1$ phases, $\xi$ may be composed of the volume or mass fractions of any $N$ components. The specific expression for the phase transformation fraction is normally formulated in a phenomenological way. An example of a set of thermodynamic state variables used in modeling the thermomechanical behavior of materials in the process of phase transition is shown as follows:

$$\Lambda = (\epsilon, T, \nabla T, \pi, \pi, \xi)$$  \hspace{1cm} (1.2)

where $\pi$ and $\pi$ represent the internal variables which specify the internal crystallographic structure of the materials, and $\xi$ denotes the phase transition fraction of the materials and was phenomenologically formulated as an exponential function of stress state and critical temperatures. A different set of state variables may be selected for describing the same thermodynamic process. Based on Tanaka’s model of shape memory alloys, Liang and Rogers (1990) proposed several phenomenological formulations for the martensitic transformation fraction $\xi$ and material constants involved in Tanaka’s model to improve the accuracy and feasibility of Tanaka’s model of shape memory alloys.

In this thesis, the thermomechanical approach will be extended to formulate the theoretical models for the two-way shape memory effect and the thermo-electromechanical coupling effect of piezoelectricity. After mathematical formulations, the computer implementation will be executed and the numerical results will be compared with experimental data that has been published in the open literature.
1.3 Introduction to Shape Memory Alloys

The shape memory effect was first observed as early as 1932 by Ölander who noted rubberlike behavior in a Au-Cd alloy system. The martensitic transformation relating to shape memory effect was discovered in 1938 in CuZn and CuAl alloys (Isaitschev, et al., 1938). Later the concept of thermoelastic martensitic transformation provided a crystallographic and kinetic basis for the observation (Kurdyumov and Khandros, 1949). The first demonstration that the shape memory effect can be used to convert heat to useful work came in 1958 (Lieberman and Reed, 1958). It was in 1962 that Buehler and Wiley (1965) of the Naval Ordinance Laboratory explored the Nickel-Titanium alloy systems which exhibit remarkable shape memory effect. This alloy system was then named Nitinol. Nitinol has the highest efficiency (ratio of mechanical energy to heat energy) of any other shape memory alloy and also has other attractive features such as high strength and excellent corrosion resistance; therefore, it has become the most widely used and commercially available shape memory alloy.

The shape memory effect can be classified into one-way effect and two-way effect, as illustrated in Fig. 1. One-way shape memory effect can be described as follows: A material first undergoes a martensitic transformation. After deformation in the martensitic condition, the apparently permanent strain is recovered when the specimen is heated to cause the reverse martensitic transformation; it does not return to its deformed shape upon cooling. It is clear that the process of regaining the original shape is associated with thermoelastic characteristics of the reverse phase transformation of deformed martensite to high-temperature austenite. Two-way shape memory effect may be stated in a similar way: A material normally exhibiting the one-way effect is thermomechanically processed, then it has the capability to exhibit a spon-
Figure 1.1: One-Way and Two-Way Shape Memory Effect.
taneous shape change upon cooling through the martensitic transformation regime; upon heating, the inverse shape change occurs via the one-way mechanism (Wayman, 1980). It should be emphasized that the two-way shape memory alloys are obtained by training SMAs that exhibit one-way effect.

Alloys exhibiting the shape memory effect belong to two general categories: non-ferrous and ferrous. Nonferrous alloys include Au-Cd, Ni-Ti, Ni-Ai and copper-based alloy systems such as Cu-Zn, Cu-Zn-Al, Cu-Zn-Ga, Cu-Zn-Sn, Cu-Zn-Si, Cu-Al-Ni, Cu-Au-Zn, Cu-Sn and others. Ferrous shape memory alloys under development are Fe-Pt, Fe-Ni-C, Fe-Pd, and Fe-Ni-Co-Ti. The nonferrous alloys of Ni-Ti type known as Nitinol are the most common shape memory alloys, as indicated before.

The mechanical behavior of shape memory alloys is much different from other engineering materials. One remarkable characteristic of its mechanical behavior is the variation of the Young’s modulus. For most metals, the Young’s modulus is insensitive to temperature. However, the Young’s modulus of shape memory alloys changes dramatically within the martensitic transformation temperature regime. It is shown that the Young’s modulus of Nitinol increases by 3 to 4 times from a temperature below $M_f$ to a temperature above $A_f$. Another unusual mechanical behavior of shape memory alloys is the large recoverable plastic deformation. The material can be plastically deformed in its low-temperature martensitic phase and then restored to its original configuration by heating it above the characteristic transition temperature. For Nitinol, plastic strains of typically 6-8% may be completely recovered by heating the material to its austenite temperature range. If the material is restrained from regaining its original shape, stresses with a magnitude of 100,000 psi can be induced within the material.
Due to their unique mechanical characteristics, applications for shape memory alloys are expanding. In the 1980s, many application patents using SME were submitted and the total number of registered patents has reached more than 4000 (Miyazaki and Otsuka, 1989). These applications include most of the industrial fields such as electrical engineering, machinery, transportation, chemical engineering, energy, and medicine. However, it should be noted that the application products commercially available in the market amount to less than one percent of the total number of applied patents. One possible reason is the current high price of Nitinol, especially the high manufacturing and processing cost. This problem may be solved with increases in the productivity. In addition, if the characteristics of Cu-based and ferrous shape memory alloys are improved to a level suitable for commercial use, the cost will be decreased further. A comparison of Nitinol and CuZnAl shape memory alloys is shown in Table 1. Another possible reason that so few products are available commercially is the stability of the shape memory performance of Nitinol. In order to meet the demand for more precise and repeating functions, further improvement of the quality of Nitinol is necessary. Therefore, much research is still needed to increase the duration and repeatability of Nitinol, improve the manufacturing processes, develop Cu-based and ferrous shape memory alloys, purify the alloys, determine the effect of adding a third element in Nitinol, and investigate the sensing characteristics of Nitinol.

1.4 Review of Piezoelectricity

Piezoelectricity is a reciprocal interaction between electrical and mechanical systems. The direct piezoelectric effect is that certain crystalline materials may develop an electric charge proportional to an applied stress. Contrarily, the converse piezoelec-
The pyroelectric effect is that some materials possess temperature-dependent spontaneous electric moments. When the material is heated, an electric charge may be accumulated on its surface. The phenomenon of pyroelectricity has been known since the eighteenth century. In 1824, Brewster found the effect with different kinds of crystals and coined the term pyroelectricity (the prefixes 'pyro-' and 'piezo-' are derived from the Greek words for 'fire' and 'press', respectively). Later, Lord Kelvin proposed that pyroelectricity is due to a permanent polarization and that the pyroelectric effect is a simply evidence of the temperature coefficient of this polarization (Ikeda, 1990). Analogous to pyroelectricity, piezoelectricity was discovered in 1880 by Pierre Curie and his brother, Jacques Curie, through studying the relation between pyroelectricity and crystal symmetry. This discovery was made by placing a weight on the surface of hemihedral crystals of Zinc blende (ZnS), sodium chlorate (NaClO₃), boracite, tourmaline and quartz to produce a charge which is proportional to the applied weight. The brothers also established the symmetry conditions necessary for its existence in crystals and proposed the quantitative laws of piezoelectricity (Seymour and Kaufman, 1990). In the same year of the discovery of the direct effect, Lippmann predicted the existence of the converse effect from thermodynamic considerations, which was
verified experimentally by the Curie brothers before the end of 1881. The reason that
the converse piezoelectric effect was explored after its counterpart is because of the
experimental difficulty in measuring the minute deformation of the crystals at that
time. Hankel proposed the term ‘piezoelectricity’ for this phenomenon (Ikeda, 1990).

Associated with piezoelectricity, ferroelectricity is also an important phenomenon on
which the piezoelectric effect of polycrystalline materials depends. Ferroelectricity is
the presence of a spontaneous electric moment in a crystal which can be changed in its
crystallographic orientation by applying an electric field (Jaffe et al., 1971). Namely,
a ferroelectric crystal is a pyroelectric crystal with reversible polarization. The term
‘ferroelectricity’ is coined analogous to ferromagnetism. Most ferroelectric crystals
are strongly pyroelectric and piezoelectric. Accordingly, a study of ferroelectricity is
beneficial to the understanding of piezoelectricity. Rochelle salt was the first ferro-
electric crystal which was discovered in 1921, and new ferroelectric materials have
been found in succession. Up to now, more than 100 ferroelectric materials have been
found.

The first ceramic form of piezoelectric materials is $BaTiO_3$-based ceramics which were
developed after World War II. This discovery could be classified as three basic steps
(Jaffe et al., 1971). The first step was the discovery of the high dielectric constant.
The second step, which occurred in the mid 1940s, was the understanding that the
high dielectric constant is due to ferroelectricity. The third step, the most important
procedure, was the poling process that a sufficient high voltage is applied to reverse
the electric moments of spontaneously polarized regions in the ceramic. Modifications of $BaTiO_3$ were conducted to improve the temperature stability or to obtain
better voltage output by substituting $Pb$, $Sr$, or $Ca$ for $Ba$, and $Zr$ or $Sn$ for $Ti$. 

16
Along with the improvement of the properties of $\text{BaTiO}_3$ ceramics, an active search for other piezoelectric ceramics was ongoing; examples of these are lead niobate and lead titanate.

The discovery of a very strong and stable piezoelectric effect in lead titanate zirconate in 1954 is the most significant event in the history of piezoelectric materials. Lead titanate zirconate with various additives has since become the dominant piezoelectric ceramic. The piezoelectric coefficients of this material were higher than those of $\text{BaTiO}_3$ ceramics and dielectric and elastic losses were lower at high drive levels. In the last two decades, significant progress has been made in the research of this material. One development is the understanding of fracture processes in piezoelectric ceramics; another is the development of composite piezoelectric materials. In addition, there are significant advances in powder processing, powder characterization, and sintering techniques.

Polymer material is an emerging piezoelectric material which has attracted many researchers to the study of PVDF and related materials. PVDF was patented in 1948 and was commercially available in 1965. It has a melting point of 170°C and a specific gravity of 1.78. The strong piezoelectric properties of the crystalline PVDF films were demonstrated in 1969. PVDF is a polymer of $(-\text{CH}_2-\text{CH}_2-)_n$ that has crystallinity of 40-50%. There are two types of PVDF, namely, type I (or $\alpha$) and type II (or $\beta$). The commercial PVDF is normally as type II which is poor in piezoelectric effect. Type I is polar and piezoelectric. The improved poling process involves holding specimen at 90 ~ 130°C for 15 min ~ 2 hr under a dc field of 500 ~ 1000$kVcm^{-1}$ and then cooling it to room temperature at the same field level. Recent progress has been made in developing piezoelectric polymer materials out of copolymers. The
copolymers VDF/TrFE is composed of mostly the piezoelectric type I phase with high crystallinity of about 90%. The copolymers have a higher piezoelectric coupling coefficient.

1.4.2 Background on the Theory of Piezoelectricity

Pyroelectricity is a linear interaction between electrical and thermal systems, whereas piezoelectricity is between electrical and mechanical systems. A diagram of the interaction processes of electrical, mechanical and thermal systems is shown in Fig. 1.2 (Nye, 1957; Ikeda, 1990). The interaction processes shown in this diagram are linear interactions between any two systems. Quantities at the vertices of the outer triangle represent the intensive variables, and those at the inner triangle denote the extensive variables. A connection line in this diagram indicates the coupling between the respective variables, where an unprimed notation shows a direct effect and a primed notation represents converse effect. Since pyroelectricity and piezoelectricity are thermal-electrical-mechanical related, therefore, the study of piezoelectricity may provide a general understanding of linear interaction processes.

Concepts of piezoelectricity

The phenomenon of piezoelectricity is shown in Fig. 1.3 (Moulson and Herbert, 1990). In this case, a piezoelectric plate is polarized in the direction $\mathbf{P}$ and electrodes are attached to its two flat surfaces. If a compressive load is applied, a transient current will be produced and flow in the external circuit. Alternatively, a tensile load will induce an electric current in the opposite direction (Fig. 1.3(a)). Conversely, if an electric field is applied, a strain will be produced in the plate. The induced strain
Figure 1.2: Interaction Processes Between the Electrical, Mechanical, and Thermal Systems (after Ikeda, 1990).
changes its sign with the electric field (Fig. 1.3(b)). The direct piezoelectric effect is that the induced polarization is proportional to the applied stress, namely:

\[ P_i = d_{ikl} \sigma_{kl} , \]  

(1.3)

where \( P \) is the polarization, \( d_{ikl} \) is the piezoelectric moduli and \( \sigma_{kl} \) the stress tensor. The converse effect is that the induced strain is proportional to the applied field, i.e.:

\[ \epsilon_{ij} = d_{kij} E_k \]  

(1.4)

where \( \epsilon_{ij} \) represents the strain tensor. Since strain tensor \( \epsilon_{ij} \) is symmetric, it is implied that tensor \( d_{ijk} \) is symmetric in the last two indices. Parton et al. (1988) showed simply that a centrosymmetrical crystal cannot be piezoelectric. If all three Cartesian coordinate axes in a centrosymmetrical crystal are inverted, then one may obtain:

\[ d'_{nlm} = \alpha_{nk} \alpha_{li} \alpha_{mj} d_{kij} = -d_{nlm} \]

where

\[ \alpha_{nk} = -\delta_{nk} = \begin{cases} -1 & n = k \\ 0 & n \neq k \end{cases} , \]

the primed components refer to the reversed coordinate axes. On the other hand, \( d_{kij} \) cannot change under this transformation due to its centrosymmetric characteristic, i.e., \( d'_{nlm} = d_{nlm} \). Accordingly, all the components of the tensor \( d_{nlm} \) are zero, namely, there is no piezoelectricity in a centrosymmetrical crystal. From a viewpoint of crystallography, if a stress is applied to a crystal with centrosymmetry, there is no net change of dipole moments due to the cancellation of the symmetric ion displacement in the crystal. There are 32 crystal classes of single-crystal materials, eleven of them are non-polar and with a center of symmetry. The other 21 crystal classes are non-centrosymmetric, and 20 of them have the piezoelectric effect. The one exception is that a crystal in the cubic system 432 possesses the crystallographic symmetry which
cancels the net piezoelectric effect.

A few important concepts such as pyroelectricity, ferroelectricity and electrostriction are closely related to piezoelectricity. Among 20 piezoelectric crystallographic classes, 10 have a unique polar axis. The crystals of these 10 classes are termed polar crystal since they possess electric moments without an applied stress. These polar crystals can produce a charge on their surface when uniformly heated due to a change in the magnitude of the dipole moments with temperature. This phenomenon is called pyroelectricity. Accordingly, the 10 polar crystal classes are termed the pyroelectric classes. In piezoelectric but nonpyroelectric crystals such as quartz, dipoles are arranged in several compensating directions to cancel the net crystal dipole moments. Ferroelectricity refers to the reversibility of the direction of electric dipoles in a polar crystal under an applied electric field. The preliminary condition for ferroelectricity is that the crystal belongs to pyroelectric classes because it requires polarity in a material. However, the pyroelectric characteristic of a crystal does not guarantee its ferroelectricity. In some pyroelectric crystals, the dipoles cannot be reversed by an electric field either because it requires a field larger than breakdown field or due to an asymmetric and irreversible arrangement of the atoms. For example, tourmaline and hexagonal CdS are pyroelectric but not ferroelectric. From the above discussion, it is clear that ferroelectric crystals are both pyroelectric and piezoelectric. Electrostriction is a higher-order electromechanical coupling. In electrostriction, the sign of the deformation is independent of the polarity of the field, and the magnitude of the deformation is proportional to even powers of the field. All materials exhibit weak electrostriction. In ferroelectric materials at the temperature just above their Curie point, large electrostriction may be presented.
Figure 1.3: Piezoelectric Phenomena: (a) Direct Piezoelectric Effect, (b) Converse Piezoelectric Effect.
Table 1.1: Comparison of TiNi and CuZnAl Alloys

<table>
<thead>
<tr>
<th>Item</th>
<th>TiNi Alloy</th>
<th>CuZnAl Alloy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recovery Strain</td>
<td>max 8%</td>
<td>max 4%</td>
</tr>
<tr>
<td>Recovery stress</td>
<td>max 400 MPa</td>
<td>max 200 MPa</td>
</tr>
<tr>
<td>Repetition Life</td>
<td>$10^5 (\varepsilon = 0.02)$</td>
<td>$10^2 (\varepsilon = 0.02)$</td>
</tr>
<tr>
<td></td>
<td>$10^7 (\varepsilon = 0.005)$</td>
<td>$10^4 (\varepsilon = 0.005)$</td>
</tr>
<tr>
<td>Corrosion Resistance</td>
<td>Good</td>
<td>Problematic, especially stress corrosion cracking</td>
</tr>
<tr>
<td>Workability</td>
<td>Poor</td>
<td>Fair</td>
</tr>
<tr>
<td>SME Processing</td>
<td>Comparatively easy</td>
<td>Fairly difficult</td>
</tr>
</tbody>
</table>
Basic constitutive relations of piezoelectricity

Piezoelectric properties are usually described in terms of the stress \( T \), strain \( S \), electric field \( E \) and dielectric displacement \( D \) or polarization \( P \). By thermodynamic argument, it is shown that the piezoelectric coefficients are the same for direct piezoelectric effect and converse effect (Mouison and Herbert, 1990). The most common piezoelectric coefficients \( d \), \( g \), \( e \) and \( h \) are defined as follows:

\[
\begin{align*}
    d &= \left( \frac{\partial D}{\partial T} \right)_E = \left( \frac{\partial S}{\partial E} \right)_T \\
    g &= \left( \frac{\partial S}{\partial T} \right)_D = \left( \frac{\partial D}{\partial E} \right)_T \\
    e &= \left( \frac{\partial E}{\partial T} \right)_S = \left( \frac{\partial D}{\partial S} \right)_E \\
    h &= \left( \frac{\partial E}{\partial D} \right)_S = \left( -\frac{\partial S}{\partial S} \right)_D
\end{align*}
\] (1.5)

where the subscripts represent the quantities that are kept constant when the coefficients are evaluated. It should be noted that the temperature is a constant for all the above relations.

Ikeda (1990) developed a general description of constitutive relations in a linearly coupled system. Assuming that the variables for two different phenomena in a system are linearly coupled, then the free energy \( F \) in this system is uniquely determined by the homogeneous quadratic expression:

\[
F(x, y) = \frac{1}{2}a_{11}x^2 + a_{12}xy + \frac{1}{2}a_{22}y^2,
\] (1.6)

where \( x \) and \( y \) represent the two state variables for the two phenomena in the coupled system. The state for \( x = y = 0 \) is an equilibrium or \( F(0,0) = 0 \). The interaction between the two state variables depends on the equilibrium coefficient \( a_{12} = [\partial^2 F/\partial x \partial y]_{x,y=0} \). If the free energy function \( F(x, y) \) is a continuous function of \( x \) and \( y \), then the exact differential of \( F(x, y) \) is given by:

\[
dF = X \, dx + Y \, dy,
\] (1.7)
where \( X \) and \( Y \) are the generalized forces associated with the state variables \( x \) and \( y \). Accordingly, the expressions for \( X \) and \( Y \) may be obtained:

\[
X = \frac{\partial F}{\partial x} \quad Y = \frac{\partial F}{\partial y}.
\] (1.8)

From these differentials, one may obtain the linearly coupled constitutive equations as follows:

\[
\begin{align*}
X &= a_{11} x + a_{12} y \\
Y &= a_{12} x + a_{22} y.
\end{align*}
\] (1.9)

It is suggested that the above linearly coupled constitutive equations can be derived alternately by the Taylor expansion of \( X \) and \( Y \). Therefore, the coefficient of \( x \) in the second equation can be written as \( a_{21} \) due to the fact that \( a_{12} = a_{21} \) can be easily proven by the above thermodynamic argument.

Coefficients \( a_{11} \) and \( a_{22} \) are, respectively, the principal constants of the \( x \) and \( y \) systems. If there is no coupling, i.e., \( a_{12} = 0 \), then each principal constant can be determined in its own system. In the coupled system, the principal constants must be determined under the condition where the opposite system is specified. In the above constitutive equations, \( a_{11} \) and \( a_{22} \) can only be obtained for constant \( y \) and constant \( x \), respectively. Accordingly, the principal constants are usually denoted as \( a_{11}^y \) and \( a_{22}^x \), where superscripts \( x \) and \( y \) represent that the quantities are kept constant when the principal constants are measured.

The constitutive relations in Eq. (1.9) are in terms of independent variables \( x \) and \( y \). If this form of constitutive relation is referred to as \((x, y)\)-type, then the choice of a different set of variables provides other types of constitutive relations: \((x, Y)\), \((X, y)\) and \((X, Y)\) type. If one set of constitutive relations is known, then other sets of constitutive relations can be easily found by rearranging the constitutive equations.
already known. As an example, the \((X, y)\)-type of the constitutive relations can be formulated as follows by arranging the Eq. (1.9):

\[
\begin{align*}
X &= \frac{1}{a_{11}^*} X - \frac{a_{12}^*}{a_{11}^*} Y \\
Y &= \frac{a_{22}^*}{a_{11}^*} X + \left( a_{22}^* - \frac{a_{22}^*}{a_{11}^*} \right) Y.
\end{align*}
\] (1.10)

For piezoelectricity, if the four variables of strain \(S\), stress \(T\), polarization \(P\) and electric field \(E\) are chosen for constructing the constitutive relations, then by using the general argument above, one may obtain expressions for Helmholtz free energy, Gibbs free energy, elastic Gibbs energy, electric Gibbs energy, and the corresponding forms of the constitutive relations for piezoelectricity, as shown in Table 1.2 (a). If the electric displacement \(D\) is used instead of polarization in the derivation, one may obtain the different forms of corresponding constitutive relations for piezoelectricity, as shown in Table 1.2 (b).

Among all the different forms of piezoelectric relations, the \(d\)-form of the relation shown in Table 1.2 (b) is the most useful piezoelectric relation for piezoelectric ceramics. For the direct piezoelectric effect, the \(d\)-form relation may be expanded to:

\[
\begin{align*}
D_1 &= \varepsilon_1^T E_1 + d_{15} T_5 \\
D_2 &= \varepsilon_1^T E_2 + d_{15} T_4 \\
D_3 &= \varepsilon_3^T E_3 + D_{31}(T_1 + T_2) + d_{33} T_3,
\end{align*}
\] (1.11)

and for the converse piezoelectric effect, the \(d\)-form relation may be expressed as:

\[
\begin{align*}
S_1 &= s_{11}^E T_1 + s_{12}^E T_2 + s_{13}^E T_3 + d_{31} E_3 \\
S_2 &= s_{12}^E T_2 + s_{12}^E T_1 + s_{32}^E T_3 + d_{31} E_3 \\
S_3 &= s_{13}^E (T_1 + T_2) + s_{33}^E T_3 + d_{33} E_3 \\
S_4 &= s_{44}^E T_4 + d_{15} E_2 \\
S_6 &= 2(s_{11}^E - s_{12}^E) T_6.
\end{align*}
\] (1.12)

In the above piezoelectric relations, the poling direction is in the 3-direction, and the 1-2 plane is an isotropic plane. The subscripts 4, 5 and 6 represent the shear planes perpendicular to the 1, 2 and 3 directions, respectively. Accordingly, \(d_{31}\) denotes the coefficient relating the field along the polar axis to the strain perpendicular to it,
Table 1.2: Types of Fundamental Piezoelectric Relations

(a) Polarization scheme

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Piezoelectric relation</th>
<th>Thermodynamic function</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S, P$</td>
<td>$T = e^S S - aP$</td>
<td>Helmholz free energy</td>
</tr>
<tr>
<td></td>
<td>$E = -aS + \chi^S P$</td>
<td>$F = \frac{1}{2}e^S S^2 - aSP + \frac{1}{2}\chi^S P^2$</td>
</tr>
<tr>
<td>$T, E$</td>
<td>$S = s^T T + dE$</td>
<td>Gibbs free energy</td>
</tr>
<tr>
<td></td>
<td>$P = dT + \kappa^T E$</td>
<td>$G = -\frac{1}{2}s^T T^2 - dTE - \frac{1}{2}\kappa^T E^2$</td>
</tr>
<tr>
<td>$T, P$</td>
<td>$S = s^T T + bP$</td>
<td>Elastic Gibbs energy</td>
</tr>
<tr>
<td></td>
<td>$E = -bT + \chi^T P$</td>
<td>$G_1 = -\frac{1}{2}s^T T^2 - bTP + \frac{1}{2}\chi^T P^2$</td>
</tr>
<tr>
<td>$S, E$</td>
<td>$T = e^S S - eE$</td>
<td>Electric Gibbs energy</td>
</tr>
<tr>
<td></td>
<td>$P = -eS + \kappa^S E$</td>
<td>$G_2 = \frac{1}{2}e^S S^2 - eSE - \frac{1}{2}\kappa^S E^2$</td>
</tr>
</tbody>
</table>

$G = F - TS - EP, \ G_1 = F - TS, \ G_2 = F - EP$

(b) Electric flux density scheme (electric displacement scheme)

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Type</th>
<th>Piezoelectric relation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S, D$</td>
<td>Extensive</td>
<td>$T = e^D S - hD$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$E = -hS + \beta^D$</td>
</tr>
<tr>
<td>$T, E$</td>
<td>Intensive</td>
<td>$S = s^E T + dE$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$D = dT + e^E$</td>
</tr>
<tr>
<td>$T, D$</td>
<td>Mixed</td>
<td>$S = s^D T + gD$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$E = -gT + \beta^D$</td>
</tr>
<tr>
<td>$S, E$</td>
<td>Mixed</td>
<td>$T = e^S S - eE$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$D = eS + \epsilon^E$</td>
</tr>
</tbody>
</table>

27
\( d_{33} \) is the corresponding coefficient that both strain and polar axis are in the same direction, and \( d_{15} \) is the only coefficient for shear strain in which an electric field is applied perpendicular to the polar axis.

**Parameters of piezoelectric ceramics**

As discussed before, the most common piezoelectric coefficients are \( d, g, e \) and \( h \). The relations of these four piezoelectric tensors may be drawn from their definitions:

\[
\begin{align*}
\frac{d}{g} &= \frac{(\frac{\partial s}{\partial E})_T}{(\frac{\partial D}{\partial E})_T} = (\frac{\partial D}{\partial E})_T = \varepsilon^T \\
\frac{e}{h} &= \frac{-\frac{\partial s}{\partial T}}{-\frac{\partial D}{\partial T}} = (\frac{\partial D}{\partial E})_S = \varepsilon^S
\end{align*}
\]

(1.13)

Specifically, one may write the above relations into component forms as follows:

\[
\frac{d_{ij}}{g_{ij}} = \varepsilon_{ii}^T, \quad \frac{e_{ij}}{h_{ij}} = \varepsilon_{ii}^S.
\]

(1.14)

The elastic parameters in piezoelectric ceramics are stiffness coefficients \( (c_{ij}) \) and compliance coefficients \( (s_{ij}) \). By inverting the linear d-form piezoelectric relation in Eq. (1.12), one may obtain the following relations which are valid for either open- or short-circuit conditions:

\[
\begin{align*}
c_{11} &= (s_{11}s_{33} - s_{13}^2)/f(s) \\
c_{12} &= -(s_{12}s_{33} - s_{23}^2)/f(s) \\
c_{13} &= -s_{13}(s_{11} - s_{12})/f(s) \\
c_{33} &= (s_{11}^2 - s_{12}^2)/f(s) \\
c_{44} &= 1/s_{44} \\
c_{ij} &= \frac{1}{s_{ij}}
\end{align*}
\]

(1.15)

where \( f(s) = (s_{11} - s_{12})(s_{33}(s_{11} + s_{12}) - 2s_{13}^2) \). It should be noted that the elastic constants are different for open-circuit and short-circuit conditions which are denoted by superscripts \( D \) and \( E \), respectively. The effective stiffness constants with electrodes short circuited is normally lower than those with the electrodes open circuited because mechanical stressing of a ceramic produces an electric charge which is opposite to the
resultant strain. The Poisson ratios \( \nu \) are also important elastic parameters for piezoelectric ceramics. They are defined in terms of open-circuit compliances:

\[
\nu_{12} = -\frac{s_{12}^E}{s_{11}^E}, \quad \nu_{13} = -\frac{s_{13}^E}{s_{33}^E}.
\]

(1.16)

The electromechanical coupling coefficient \( k \) is an important parameter in measuring the strength of the piezoelectric effect. When an electric field is applied, it measures the fraction of the electrical energy converted to mechanical energy (or vice versa when a stress is applied). It is defined as follows:

\[
k^2 = \frac{\text{electrical energy converted to mechanical energy}}{\text{input electrical energy}}
\]

or

\[
k^2 = \frac{\text{mechanical energy converted to electric energy}}{\text{input mechanical energy}}.
\]

(1.17)

Since the energy conversion of any two different forms is always incomplete, accordingly, \( k^2 < 1 \) or \( k < 1 \). Typical values of \( k \) are 0.4 for barium titanate ceramic, 0.5-0.7 for PZT (Jaffe et al., 1971).

The electromechanical coupling factors are material constants for piezoelectric materials. The magnitude of the coupling factors depends on the degree of poling. The most common coupling factors are those derived from the d-form constitutive equations and \( k_t \):

\[
\begin{align*}
k_{15} &= d_{15}/\sqrt{s_{44}^Es_{11}^T} \\
k_{33} &= d_{33}/\sqrt{s_{33}^Es_{33}^T} \\
k_{31} &= d_{31}/\sqrt{s_{11}^Es_{33}^T} \\
k_p &= d_{31}/\sqrt{2/(s_{11}^E + s_{15}^E)e_{33}^T} = k_{31}\sqrt{2/(1 - \nu_{12})} \\
k_t &= e_{33}/\sqrt{e_{33}^Te_{33}^D},
\end{align*}
\]

(1.18)

where \( \nu_{12} \) is the Poisson’s ratio as defined before. \( k_p \) is called the planar coupling factor which refers to the simultaneous application of equal stress along the 1 and 2 direction or of uniform radial tension or compression in the plane normal to the poling
axis. \( k_t \) refers to the longitudinal thickness vibration of laterally clamped plates.

The coupling constants may relate the open- and short-circuit elastic constants by manipulating the d-form constitutive equations with the above equations:

\[
\begin{align*}
    s_{44}^D &= s_{44}^E (1 - k_{15}^2) \\
    s_{33}^D &= s_{33}^E (1 - k_{33}^2) \\
    s_{11}^D &= s_{11}^E (1 - k_{31}^2) \\
    s_{12}^D &= s_{12}^E - k_{31}^2 s_{11}^E \\
    c_{33}^E &= c_{33}^D (1 - k_t^2).
\end{align*}
\]  

(1.19)

1.4.3 General Considerations Regarding Piezoelectric Ceramics

Microstructures of Piezoelectric Ceramics

The first developed commercial piezoelectric ceramic is \( \text{BaTiO}_3 \) with a perovskite structure. The solid solution system \( \text{Pb(Ti, Zr)}O_3 \) or PZT also has the perovskite microstructure \( (\text{ABO}_3) \). Above the Curie point, the unit cell of the perovskite structure is cubic, as shown in Fig. 1.4. In this structure, a large cation \( (A) \) is at the corners, a small cation \( (B) \) is in the body center, and oxygens \( (O) \) are in the centers of the faces. The structure can also be considered as a network of corner-linked oxygen octahedral, with the cation \( B \) filling the octahedral holes and the cation \( A \) occupying the dodecahedral holes (Jaffe et al., 1971). In most cases, the perovskite structure is slightly distorted. Also, the anion need not be oxygen.

The PZT phase diagram is shown in Fig. 1.5. A significant feature in this diagram is the morphotropic phase boundary (MPB) in which a structure changes abruptly with composition at the constant temperature in a solid solution range. In the PZT system the MPB occurs at the composition where the ratio of \( \text{PbTiO}_3 \) to \( \text{PbZrO}_3 \) is
Figure 1.4: Unit Cell of $BaTiO_3$ with a Perovskite Cubic Structure.
Figure 1.5: Phase Diagram of $Pb(Ti_{1-x}Zr_x)O_3$ System (after Jaffe et al., 1971).
1. It is shown that the PZT with composition near the MPB has electromechanical coupling and the relative permittivity peak.

**Effects of Domains on Properties of Piezoceramics**

In a single crystal or a crystallite, the dipoles are usually not polarized uniformly in one direction. Single crystals or crystallites are composed of numerous domains which contain large numbers of dipoles all aligned in the same direction. The domains are separated by domain walls which are the loci of the points where the dipole orientation suddenly changes (Fatuzzo and Merz, 1967). Two kinds of domain walls are formed in one crystallite: 90° and 180° in tetragonal symmetry or 71°/109° and 180° in rhombohedral symmetry. The walls which separate antiparallel domains are called 180° domain walls and those which separate dipoles at right angles to each other are called 90° walls. Figure 1.6 is a graphical representation of 180° and 90° domains in barium titanate.

The static configurations have been studied using several techniques such as induced strain, etching, powder, TEM and SEM. Domain reorientation can be caused by either an electric field or mechanical stress, but domain processes are influenced differently. In one crystallite, both 90° and 180° domain reorientations in tetragonal symmetry or both 71°/109° and 180° domain reorientations in rhombohedral symmetry can occur under a high dc electric field. The domains with their polar direction along the field direction grow at the expense of those opposite to the field direction. It is shown that the magnitude of the 180° domain reorientation is larger than that of the 90° domain reorientation because the response of the 90° domains will be restricted by the strain situation within and between the crystals. Mechanical stress can only
Figure 1.6: 180° and 90° Domain Configurations in Barium Titanate (after Moulson and Herbert, 1990).
cause 90° domain reorientation in tetragonal symmetry or 71°/109° domain process in rhombohedral symmetry because only these non-180° domain processes incur the mechanical strain or dimensional change of the specimen. Specifically, compression tends to orient the polar direction perpendicular to the stress whereas tension tends to orient the polar direction along the stress, as shown in Fig. 1.7.

The poling and depoling processes are also shown in Fig. 1.7. Poling refers to the process which transforms the ferroelectric ceramics into polar materials by applying a static field. Poling is normally executed by applying a field of 1-4 MVm\(^{-1}\) that lasts for several minutes at temperatures above 100°C. The higher the temperature and the electric field, the easier the domain alignment. But the temperature is limited by the leakage current which leads to thermal breakdown, and the field is restricted by the breakdown strength. The specimen is normally immersed in an insulating oil to prevent surface breakdown. Depoling can be induced by applying a field in the direction opposite to the poling field, or by applying a high AC field and gradually reducing it to zero. It should be noted that overheating may occur due to the high dielectric loss at high fields. In addition, depoling may take place when a compressive stress is applied (10-100 MPa). Complete depoling can be obtained by heating the specimen to well above the Curie point and cooling it to room temperature without an electric field.

Domain dynamic characteristics are the causes of the hysteresis and energy loss. It is well known that dielectric hysteresis is caused by the delayed response of the domain switching. Details of the issues relating to hysteresis will be addressed in Chapter 4. Alternating field may cause the domain walls to oscillate, which may lead to the conversion of some electrical energy into heat (part of the dielectric loss).
Figure 1.7: Poling in a Two-Dimensional Ceramic: (a) Unoriented Material; (b) Oriented by 180° Domain Changes; (c) Oriented by Both 180° and 90° Domain Changes; (d) Disoriented by Stress (after Moulson and Herbert, 1990).
Aging refers to the changes in piezoelectric properties with time in the absence of either mechanical or electrical stresses or temperature changes. Most of the piezoelectric properties except the frequency constant decrease by a few percent per decade. The cause of aging is the diminution of domain wall mobility which is due to the gradual build-up of inhibiting factors such as the accumulation of defects in the domain walls, the redistribution of internal strains due to crystal anisotropy, and the alignment of dipoles formed from lattice defects and impurities.

**Important Commercial Piezoceramics**

The important commercial piezoceramics include barium titanate $\text{BaTiO}_3$, solid solution lead zirconium titanate $\text{Pb(Zr,Ti)}O_3$ (PZT), lead niobate, lithium niobate, lithium tantalate, and the recently developed piezoceramic-polymer composites. For different applications, different piezoceramics may be chosen or their compositions may be varied. Table 1.3 is a summary of the useful properties of the typical piezoceramics.

### 1.4.4 Phenomenological Modeling of Piezoelectricity

The phenomenological formulation of piezoelectricity has been intensively researched since the 1940s. Many different theories have been developed, but most of the existing phenomenological theories on piezoelectricity are restricted to single crystals. Only in recent years have a few phenomenological theories been proposed for modeling the electromechanical behavior of piezoceramics. In this section, a detailed review of the phenomenological modeling of piezoelectricity will be presented.
Table 1.3: Typical Values of the Properties of Some Piezoceramics

<table>
<thead>
<tr>
<th>Properties</th>
<th>Lead zirconate titanates</th>
<th>Niobates $Pb(NbO_3)_2$</th>
<th>Barium titanate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4</td>
<td>5A</td>
<td>6B</td>
</tr>
<tr>
<td>Coupling factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{33}$</td>
<td>0.70</td>
<td>0.705</td>
<td>0.375</td>
</tr>
<tr>
<td>$k_t$</td>
<td>0.51</td>
<td>0.486</td>
<td>0.30</td>
</tr>
<tr>
<td>$k_p$</td>
<td>-0.58</td>
<td>-0.60</td>
<td>-0.25</td>
</tr>
<tr>
<td>Elastic compliances ($10^{-12} m^2 N^{-1}$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$s_{33}^D$</td>
<td>7.9</td>
<td>9.46</td>
<td>8.05</td>
</tr>
<tr>
<td>$s_{33}^{E}$</td>
<td>15.5</td>
<td>18.8</td>
<td>9.35</td>
</tr>
<tr>
<td>$s_{11}^{E}$</td>
<td>12.3</td>
<td>16.4</td>
<td>9.0</td>
</tr>
<tr>
<td>Dielectric constants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\varepsilon_{33}^{T}/\varepsilon_0$</td>
<td>1300</td>
<td>1700</td>
<td>460</td>
</tr>
<tr>
<td>$\varepsilon_{33}^{E}/\varepsilon_0$</td>
<td>635</td>
<td>830</td>
<td>385</td>
</tr>
<tr>
<td>Piezoelectric constants ($10^{-12} C N^{-1}$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d_{33}$</td>
<td>289</td>
<td>374</td>
<td>71</td>
</tr>
<tr>
<td>$d_{31}$</td>
<td>-123</td>
<td>-171</td>
<td>-27</td>
</tr>
<tr>
<td>Quality factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Q_E$</td>
<td>250</td>
<td>50</td>
<td>110</td>
</tr>
<tr>
<td>$Q_M$</td>
<td>560</td>
<td>75</td>
<td>1300</td>
</tr>
<tr>
<td>Limiting data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curie point ($^\circ C$)</td>
<td>328</td>
<td>365</td>
<td>350</td>
</tr>
<tr>
<td>Rated dynamic tensile strength ($10^7 N m^{-2}$)</td>
<td>4</td>
<td>2.7</td>
<td>2.4</td>
</tr>
<tr>
<td>Stability data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta s$/time decade (%)</td>
<td>1.5</td>
<td>0.2</td>
<td>&lt; 0.1</td>
</tr>
<tr>
<td>$\Delta \varepsilon_{33}^{T}$/time decade</td>
<td>-5.8</td>
<td>-1</td>
<td>-0.6</td>
</tr>
</tbody>
</table>
Devonshire’s Theory of Ferroelectricity

A phenomenological theory of ferroelectricity has been proposed by Devonshire (1949, 1951, 1954) based on the study of BaTiO$_3$ ferroelectrics. It is assumed that an expression for the free energy of the material is a function of temperature, stress and polarization. The general free energy expression in a three-dimensional coordinate for an unstressed material can be written as a series in terms of powers of polarization $P$:

$$
G = \frac{1}{2} \chi^X (P_1^2 + P_2^2 + P_3^2) + \frac{1}{4} \xi^X (P_1^4 + P_2^4 + P_3^4) \\
+ \frac{1}{6} \zeta^X (P_1^6 + P_2^6 + P_3^6) + \frac{1}{2} \lambda (P_1^2 P_2^2 + P_2^2 P_3^2 + P_3^2 P_1^2 + P_1^2 P_2^2) + \ldots \ , \quad (1.20)
$$

where the superscript $X$ represents the quantities evaluated in a stress-free state. The coefficients $\chi^X, \xi^X$ and $\zeta^X$ are functions of temperature. The last term in the above equation refers to an interaction between the components of polarization along the different axes. The coefficient $\lambda$ may either be positive or negative; and positive $\lambda$ denotes that the polarization along one axis tends to prohibit polarization along another axis at right angles. The above equation has a series of minima for zero field satisfying the following four different conditions, which correspond to the cubic, tetragonal, orthorhombic and rhombohedral crystal structures:

$$
P_1 = P_2 = P_3 = 0 \quad (a) \\
P_1 = P_2 = 0, P_3 \neq 0 \quad (b) \\
P_2 = 0, P_1 = P_3 \neq 0 \quad (c) \\
P_1 = P_2 = P_3 \neq 0 \quad (d) \ . \quad (1.21)
$$

To begin with, one may consider the case where only one direction of polarization, say $P$, is involved in the transition; the expression for the free energy may be simplified as:

$$
G = \frac{1}{2} \chi^X P^2 + \frac{1}{4} \xi^X P^4 + \frac{1}{6} \zeta^X P^6 \quad (1.22)
$$
where the coefficient $\chi^X$ is the dielectric stiffness or reciprocal susceptibility for zero polarization. The meaning of this coefficient can be explored by differentiating the above equation with respect to $P$:

\[
\begin{align*}
E &= \frac{\partial G}{\partial P} = \chi^X P + \xi^X P^3 + \zeta^X P^5 \\
\frac{\partial E}{\partial P} &= \frac{\partial^2 G}{\partial P^2} = \chi^X + 3\xi^X P^2 + 5\zeta^X P^4 \\
(\frac{\partial E}{\partial P})_{P=0} &= \chi^X.
\end{align*}
\] (1.23)

The nature of the ferroelectric transition depends on the magnitudes and signs of the coefficients $\chi^X, \xi^X$ and $\zeta^X$. Devonshire assumed that $\chi^X$ decreases linearly with temperature, and that $\xi^X$ and $\zeta^X$ are constant. If $\xi^X$ and $\zeta^X$ are positive, and $\chi^X$ changes continuously from positive to negative, one may obtain a series of $G$ vs $P$ curves for different temperatures, as shown in Fig. 1.8(a). The temperature for which $\chi^X = 0$ is the critical temperature of the ferroelectric-paraelectric transition. Below the critical temperature, the material is in the ferroelectric phase, and the single minimum splits into a pair. The separation $2P_c$ between the two minima increases as the temperature decreases. This case represents a second-order transition because there is no discontinuity in $P$. Alternatively, if $\xi^X$ is negative and $\zeta^X$ positive, there exists a certain temperature range in which there are three minima, as shown in Fig. 1.8(b). The phase transition occurs at the temperature at which they have equal depth. This case denotes a first-order transition because of its discontinuous nature.

Merz (1953) extended Devonshire’s theory into a very simple form. The coefficients in Devonshire’s relation may be rewritten as:

\[
\begin{align*}
\chi^X &= 2A(T - T_0) \\
\xi^X &= -4B \\
\zeta^X &= 6C,
\end{align*}
\] (1.24)

where $A$, $B$ and $C$ are newly introduced coefficients, and $T$ is the temperature. In
Figure 1.8: Variation of Elastic Gibbs Function with Polarization at Different Temperatures: (a) A Second-Order Transition, (b) A First-Order Transition (after Devonshire, 1954).
addition, a series of dimensionless coefficients are defined as follows:

\[
\begin{align*}
p &= \left( \frac{3C}{B} \right)^{\frac{1}{2}} P \\
e &= \frac{1}{B} \left( \frac{3C}{B} \right)^{\frac{3}{2}} E \\
t &= \frac{A}{B} \left( \frac{3C}{B} \right)(T - T_0).
\end{align*}
\]  

(1.25)

Accordingly, Eq. (1.22) becomes

\[
e = 2pt - 4p^3 + 2p^5.
\]  

(1.26)

By plotting the curves of \(e\) against \(p\) for a series of \(t\), as shown in Fig. 1.10, one may explain the dielectric hysteresis loops. The regions in which \(\partial e/\partial p\) are negative represent unstable states, and hence the crystal switches over to a state with a positive \(\partial e/\partial p\) value at the same value of \(e\), as shown by the vertical lines in Fig. 1.9.

Devonshire's theory provides a very effective approach to describing the ferroelectric behavior of \(BaTiO_3\) and other related single crystals. This theory has been studied intensively in the last four decades, and many fruitful results have been drawn from it. The approach of this theory has even been extended by Falk (1980) to model shape memory effects. However, this theory is basically microscopic in nature; hence it is not easy to use it in the modeling of piezoceramics.

**Chen's Semi-Microscopic Theory of Ferroelectricity**

From a physical viewpoint, a semi-microscopic theory of ferroelectricity has been developed to model both the dynamic and static behavior of ferroelectric materials (Chen and Peercy, 1979; Chen and Montgomery, 1980; Chen, 1980, 1983). One important assumption in this theory is an introduced vector \(N\) which represents the effective number of aligned dipoles in the polarization direction. The parameter \(N\)
Figure 1.9: Reduced Curves of Polarization vs. Field at Different Temperatures (after Merz, 1953).
may be defined as follows:

\[ N = \left( \sum \mu \cdot j \right) j , \]  

(1.27)

where \( \mu \) denotes the electric dipole moment in which its magnitude and direction depend on the mechanical strain \( S \), the absolute temperature \( T \) and the external electric field \( E \). The direction of the unit vector \( j \) defines the polarization direction.

For the quasi-static domain switching processes, it is presumed that the mechanical stress \( T \) and electric displacement \( D \) depend on the strain \( S \), the absolute temperature \( \theta \), the electric field \( E \) and the magnitude of the effective number \( N \). Accordingly, a set of three-dimensional constitutive relations may be formulated as follows:

\[
\begin{align*}
T &= \hat{T}(S, \theta, E, N) \\
D &= \hat{D}(S, \theta, E, N),
\end{align*}
\]  

(1.28)

with an assumption

\[ \dot{N} = f(S, \theta, E, N) . \]

For the dynamic electromechanical case, it is further presumed that the dipole moment \( \mu \) can be decomposed into its transient response \( \mu_t \) and its instantaneous response \( \mu_i \) so that

\[ \mu = \mu_t + \mu_i . \]  

(1.29)

In general the vectors \( \mu, \mu_t \) and \( \mu_i \) at each material point and time \( t \) need not be parallel with each other. In this case, a set of general dynamic electromechanical constitutive equations may be written:

\[
\begin{align*}
T &= \hat{T}(S, \theta, E, \mu_t, N) \\
D &= \hat{D}(S, \theta, E, \mu_t, N),
\end{align*}
\]  

(1.30)

together with the rate laws

\[
\begin{align*}
\dot{N} &= f(S, \theta, E, \mu_t, N) \\
\dot{\mu}_t &= g(S, \theta, E, \mu_t, N) .
\end{align*}
\]
Chen's constitutive relations including the rate laws of the effective dipole number $N$ and $\mu_i$ are entirely phenomenological. It is based on the assumptions that the mechanical properties of the ceramic are thermoelastic, and that the piezoelectric and dielectric properties of the material are due to transient and instantaneous responses of the dipole moments. Because the concept of the dipole moment is purely physical, it is not easy to relate this notion to real polycrystal materials. Therefore the applications of Chen's theory to the engineering problems are limited although Chen and his co-workers have used this theory to explain the hysteresis and butterfly loops in piezoceramics.

**Bassiouny's Macroscopic Phenomenological Theory of Piezoceramics**

Based on the thermodynamic principles of internal variables, a thermodynamic phenomenological theory has been proposed by Bassiouny and co-workers (Bassiouny, Ghaleb and Maugin, 1988, 1989) for modeling the electromechanical hysteresis effects. This theory may be applied, in principle, to piezoceramics because of its macroscopic characteristic. Two types of state variables are employed in this theory. One is a type of state variable for a reversible electromechanical process which includes the temperature $\theta$, the total strain $S$, and the reversible polarization; the other is a type of state variable for describing the irreversible characteristic of an electromechanical process which involves, in this case, a newly introduced polarization vector $P^{\text{int}}$. Assuming that the electric polarization $P$ is the sum of a reversible part $P^r$ and a residual part $P^R$, one may obtain the Helmholtz free energy expression in terms of the state variables:

$$\Phi = \Phi(S, P^r, P^{\text{int}}, \theta),$$

(1.31)
where
\[
P^r = P - P^R,
\]
\(\theta\) is the absolute temperature measured from a reference temperature \(\theta_0\), and
\(P^{int} = \int_0^t \|P^R\| \, dt'\), i.e. \(P^{int}\) denotes the cumulative residual polarization.

In the formulation, the Helmholtz free energy is decomposed into the following forms:
\[
\Phi = \Phi_{pe}(\theta, P^r, S) + \Phi_{int}(P^{int}, \theta)
\]  
(1.32)

where \(\Phi_{pe}\) denotes the thermo-electro-elastic energy, and \(\Phi_{int}\) represents the energy stocked through the electric internal variables. Similarly, the local electric field \(E^L\) may also be decomposed into the relaxation part \(E^{relax}\) and the reversible part \(E^{Lr}\), namely
\[
E^L = E^{Lr} + E^{relax}.
\]  
(1.33)

The local electric field \(E^L\) is simply related to the electromotive intensity \(\mathcal{E}\) by:
\[
\mathcal{E} + E^L = 0.
\]  
(1.34)

The necessary complementary constitutive equations in this theory are formulated through the Clausius-Duhem inequality and decomposition of its dissipation inequality. Analogous to the rate-independent magnetic hysteresis effects (Maugin, Sabir and Chambon, 1987) and the plasticity flow rule, an electric loading function \(f_1\) is proposed so that
\[
\begin{align*}
p^{int}_R &= \lambda \partial f_1 / \partial E^{int}_K, \\
p^{int}_K &= -\lambda \partial f_1 / \partial E^L_K,
\end{align*}
\]  
(1.35)

where \(E^{int}_K\) represents a local electric field related to \(P^{int}\), and \(\lambda\) is called the plastic multiplier which is determined analogously from Drucker's inequality. Accordingly, a fully-coupled 3-D constitutive equation governing the hysteresis effect can be derived
as:

\[ \dot{P}_R^R = \pm 4E_c^2 h^{-1} \| \bar{E} \| A_{KP} \dot{\varepsilon}_P < \pm \dot{\varepsilon}_M A_{MN} \dot{\varepsilon}_N >, \]  

(1.36)

with

\[ \dot{\varepsilon} = \varepsilon + E^{int}, \quad \dot{E} = E^L - E^{int}, \]

and

\[ \| \bar{E} \| = (\bar{E}_K A_{KL} E_L)^{1/2}, \quad h = 4E_c^2 \partial^2 \Phi_{int}/\partial (P^{int})^2, \]

where \( A_{KL} \) represents a positive-definite symmetric tensor which accounts for the electric anisotropy, \( E_c \) represents the coercive field, and the symbol \( < \cdot > \) denotes the positive part of its enclosure.

The above formulation provides a fully-coupled 3-D approach to modeling the electromechanical hysteresis for piezoceramics. It presents a good modeling approach and may shed light on the physical essence of the electromechanical behavior of piezoceramics because it is based on the thermodynamic principles. At this stage, this theory still seems to be a mathematical issue because too many assumptions and approximations are employed. For example, the creation of the electric loading function \( f_1 \) is analogous to the plasticity flow rule, and the introduction of the plastic multiplier \( \dot{\lambda} \) is analogous to Drucker’s inequality; the coercive field \( E_c \) is assumed to be constant but eventually depends on the applied stress and temperature. Also, the feasibility of this theory lacks experimental verification.

1.4.5 Applications

The first application of piezoelectricity is the “Langevin Sandwich” which developed out of the electromechanical properties of quartz during World War I. The Langevin sandwich is a mosaic of quartz glued between steel plates which produces ultrasonic
waves to detect submarines and underwater obstacles (Seymour and Kauffman, 1990). The first commercially available piezoceramic device was the barium titanate phonograph pickups which were marketed by the Sonotone Corporation in 1947. Since then numerous applications have been developed. The applications of piezoceramics may be classified into the following four categories (Moulson and Herbert, 1990):

1. the generation of charge at high voltages;
2. the actuation and detection of mechanical vibrations;
3. the control of frequency;
4. the generation of acoustic and ultrasonic vibrations.

The first category is based on the direct piezoelectric effect and requires a high $d$ coefficient with resistance to either electrical or mechanical damage; the second needs high piezoelectric $g$ coefficients combined with low permittivity; the third demands piezoelectric properties stable with time and temperature; and the fourth requires that the material has low losses under high applied fields. In general, the applications are dependent upon finding the most suitable material compositions and configuration. The following are some typical examples which may illustrate the important design ideas.

**Ignition Devices**

Piezoelectric ignition-generating devices involve use of the direct piezoelectric effect in which a high voltage is produced by means of a compressive stress. Piezoelectric ignition devices for camera flashlamps were first developed by General Electric and
Figure 1.10: Piezoelectric Ignition Devices: (a) Piezoelectric Ignition Capsule for Camera (after Berlincourt, 1976); (b) Piezoelectric Spark Generator (after Moulson and Herbert, 1990).
became commercially available in the early 1970s. Fig. 10 (a) is a schematic illustration of the piezoelectric ignition capsule for a camera flashlamp system. A small ceramic cube (0.1 inch) is used to provide ignition. The tiny cube is struck by a mini-striker connected to the shutter of the camera (Berlincourt, 1975). The selection of the material depends on its depoling resistance under compression. The piezoelectric element should function well without fracture or depolarization during the projected lifetime.

Another piezoelectric ignition device is the gas ignitor. A schematic diagram for a piezoelectric spark generator is shown in Fig. 10 (b). Two poled cylinders were linked together back to back to produce the charge available for the spark. The force should be applied quickly to prevent the voltage generated from leaking away. By electromechanical analysis of the gas ignitor, the total energy $W_T$ that can be dissipated in the spark for a single cylinder may be obtained:

$$W_T = \frac{1}{2} k_{33}^2 \varepsilon_{33} \frac{F^2 L}{A} = \frac{1}{2} d_{33} g_{33} \frac{F^2 L}{A},$$

where $F$ denotes the applied force, $A$ and $L$ represent the cross-sectional area and the length of the poled cylinder, respectively. From the above equation, it is clear that in order to have a better ignitor, the piezoceramic cylinder should have a larger coupling coefficient $k_{33}$ or higher piezoelectric coefficients $d_{33}$ and $g_{33}$.

**Phonograph Pickups**

Single crystal Rochelle salt initially dominated the industrial area for phonograph pickups in the period of 1935-1950 (Jaffe et al., 1971). Then ceramics came out to be superior in their mechanical and chemical stability. Ceramic phonograph pickups
Figure 1.11: Cantilever Bimorphs: (a) Series Connection, (b) Parallel Connection (after Jaffe et al., 1971).
are normally a bimorph connected in series or parallel, as shown in Fig. 1.11. The bending of the bimorph will result in strong compression in one half of the beam and compression in the other half. If the poling directions of the two halves are opposite from the center outwards or from the surface inwards, the voltage induced on the outer electrodes will be additive and hence the series connection should be used (Fig. 11(a)). Alternatively, if the two plates have the same poling direction, the output voltage can be received by parallel connection in which the outer electrodes are connected together (Fig. 11(b)).

The output voltage may be calculated theoretically by analyzing the stress distribution of the cantilever beam. The coupling coefficient $k_{31}$ is the most involved constant because the generated voltage is normal to the stress. The induced voltage $U_T$ may be expressed (Moulson and Herbert, 1990) by:

$$ U_T = \frac{3}{8} \frac{H^2}{L} h_{31} \varepsilon, $$

(1.38)

where $h_{31} = -(\partial E_3/\partial x_1)^D$; $\varepsilon$ is the deflection of the free end from the equilibrium position; $H$ and $L$ are the thickness and length of the bimorph, respectively.

**Underwater Sound and Generation of Sonic Energy**

Piezoceramics have been widely used in the generation and reception of sound waves in water and other environmental conditions. These applications include ultrasonic cleaning transducers, high-fidelity ‘tweeters’, hydrophones, and other underwater detection transducers.

PZT sandwich transducers have been widely used for ultrasonic cleaning. The sand-
Figure 1.12: Sandwich Transducers: (a) Basic Structure of Pre-stressed Sandwich Transducer; (b) Multi-Element Sandwich Transducer (after Gallego-Juárez, 1989).
wich transducer is a half-wave resonant length-expander structure which consists of a
disc or pair of discs of piezoceramics enclosed between two metal blocks. Figure 12 (a)
is a schematic illustration of a basic sandwich structure. The two PZT discs are poled
in opposite directions and separated by an electrode. They are driven to resonance
in the fundamental thickness mode. The pre-stressing of the assembly along the axis
direction could increase the coupling between the piezoelectric elements. Figure 12
(b) is a multi-element sandwich transducer in which an even number of PZT discs is
incorporated (Gallego-Juárez, 1989).

High-fidelity 'tweeters' are widely used to generate sonic energy. The key element
in these structures is a circular bimorph in which an alternating voltage drives the
disc to move and, in turn, forces a lightweight cone to vibrate. A circular bimorph
provides a high compliance suited to radiate power into air, and offers a large
electromechanical coefficient.

A hydrophone is an underwater transducer used to detect underwater sound. Its
sensitivity depends on the voltage induced by a hydrostatic pressure wave. The
product of hydrostatic strain coefficient $d_h$ and the hydrostatic voltage coefficient $g_h$,
or $d_h g_h$ is very useful in selecting the hydrophone materials because the coefficient $g_h$
relates the induced electric field in response to the applied hydrostatic stress, and $d_h$
relates the polarization under the applied hydrostatic stress. PZT is widely used as
the transducer material because of its high electromechanical coefficient. However,
PZT is not a suitable hydrophone material because of its insignificant magnitude of
$d_h g_h$. The product $d_h g_h$ is given by (Pohanka and Smith, 1988):

$$
{d_h g_h} = \frac{(d_{33} + 2d_{31})}{\varepsilon_0 K},
$$

(1.39)

54
Figure 1.13: Schematic Representation of a Composite with 1-3 Connectivity.
where \( \varepsilon_0 \) denotes the permittivity of free space and \( K \) represents the relative permittivity of the material. For PZT, \( d_{33} \) and \( 2d_{31} \) are opposite in sign and almost cancel one another, and the relative permittivity \( K \) reaches as high as 1800.

It has been shown in recent years that piezoelectric ceramic-polymer composites with different connectivities can provide a suitable hydrophone materials. Figure 1.13 shows a composite with PZT fibers embedded in a continuous polymer matrix. When a compressive stress is applied, the polymer matrix will shed the load to the PZT rods because the polymer phase is much more compliant than PZT. Namely, with the volume fraction of PZT decreased, the stress on the rods increases proportionally to keep the charge per unit area on the surface of the composite constant. It is clear that there is no stress transfer from polymer to PZT because of their series connection in the \( X_1 \) and \( X_2 \) direction. Accordingly the \( \bar{d}_{31} \) of the composite will decrease with decreasing volume fraction. If the decrease of \( \bar{d}_{31} \) does not affect \( \bar{d}_{33} \) substantially, then \( \bar{d}_A \) will be increased. In addition, \( \bar{d}_A \) will be enhanced because the dielectric constant \( K \) of the composites in the \( X_3 \) direction decreases with the volume fraction of PZT. The design principle for hydrophone composite is to enhance the product \( \bar{d}_h \bar{g}_h \) by keeping \( \bar{d}_{33} \) as large as possible and reducing \( \bar{d}_{31} \) and the dielectric constant.

### 1.5 Summary

This chapter provides general information about the intelligent material system, the thermomechanical approach and shape memory alloys. The necessity to model the two-way shape memory effect and the electromechanical coupling effects of piezoceramics is clearly presented. The thermomechanical approach, which is a combination of thermodynamic principles and continuum mechanics has been briefly introduced.
Piezoelectric materials and their applications are thoroughly reviewed to provide the background theory for modeling piezoceramics which will be addressed in detail in the sequel. For the background theory on modeling shape memory alloys, one may refer to Liang (1990) and Dye (1990).
Chapter 2

Theory of the Thermomechanical Approach

As discussed before, the main objective of this thesis is to formulate the constitutive equations for shape memory alloys and piezoceramics by using a thermomechanical approach. What is the thermomechanical approach? How is it developed? To answer these questions, it is necessary to study the thermomechanical approach systematically. In this chapter, the derivation of the thermomechanical approach will be analyzed in detail. To begin with, the basic concepts associated with the thermomechanical approach will be introduced. Then the second law of thermodynamics, i.e., the Clausius-Duhem inequality, will be derived in terms of continuum mechanics concepts. Third, by studying the axioms of formulating constitutive equations and their restrictions on the materials, the constitutive relation will be derived from the basic assumptions to its generalized and specified constitutive form. Fourth, the restrictions from the Clausius-Duhem inequality for the typical materials will be introduced, and the forms of the constitutive equations will be shaped further. Finally, we will conclude this chapter with an analysis of the thermomechanical approach by discussing Tanaka's work.
2.1 Introduction

The basic principles of the thermomechanical approach, as indicated before, have been well developed since the 1960s. In this section, the fundamental concepts relating to the thermomechanical approach will be briefly introduced based on Eringen (1966, 1975, 1980), and Coleman’s (1967) work.

A continuous media called body $B$ is a collection of material points which have continuous mass density at all times. In a Cartesian coordinate system, the original coordinate of a generic point $P$ in $B$ at time $t = 0$ is represented by the vector $X$ or its three coordinate components $X_K$ ($K = 1, 2, 3$); the deformed current configuration of the material point $P$ is denoted by the vector $x$. Accordingly, the motion of $P$ may be expressed as:

$$x = x(X, t) \quad \text{or} \quad X = X(x, t), \quad (2.1)$$

where a unique inverse relation is assumed between the two equations.

In thermodynamics, the internal energy density $u$ and temperature $T$ of a material point $X(x, t)$ depend on a class of $n + 1$ independent variables $s$ and $\nu_\alpha$ ($\alpha = 1, 2, \ldots, n$). Such a class of $n + 1$ independent variables $s$ and $\nu_\alpha$ is called a thermodynamic state. The parameters $s$ and $\nu_\alpha$ are functions of motion and time. The choice of $\nu_\alpha$ depends on the properties of the systems under consideration. Normally $\nu_\alpha$ is selected to define the thermodynamic characteristics of a physical system. Once $\nu_\alpha$ is selected, the thermodynamic state of an ideal system will be determined. Accordingly, a thermodynamic state of an ideal system may be represented as (Eringen, 1980):

$$\begin{cases} u = u(s, \nu_\alpha, X) \\ T = T(s, \nu_\alpha, X) \end{cases} \quad (2.2)$$
If Eq. (2.2) is soluble for $s$, then temperature $T$ may be selected as a new state variable to replace $s$, and in this case, the equation of state may be alternatively written as:

\[
\begin{align*}
\{ u &= u(T, \nu, X) \\
\quad s &= s(T, \nu, X) .
\end{align*}
\] (2.3)

Correspondingly, a thermodynamic process may then be defined as the changes that occur in $u$ and $s$ due to changes in $T$ and $\nu$. A thermodynamic process is called homogeneous if the functions $u$ and $s$ are independent of $X$.

Associated with a thermodynamic process, several concepts are important to describe the thermodynamic behavior of a material (Eringen, 1980):

- The entropy influx $S$ and entropy source $b$ are composed of different types of effects, and may be expressed as:

\[
S = \frac{q}{T} + S_1 \quad b = \frac{h}{T} + b_1 ,
\] (2.4)

where $q/T$ represents the entropy influx due to heat input, $h/T$ denotes the entropy source supplied by the energy source, and the remaining terms $S_1$ and $b_1$ refer to the entropy flux and source due to all other effects, respectively. In the above equations, if $S_1 = 0$, and $b_1 = 0$, then a thermodynamic process will be called a simple thermomechanical process.

- A process is called thermodynamically admissible if and only if it obeys the local Clausius-Duhem inequality and possesses a nonnegative finite temperature.

- A thermodynamic process is said to be mechanically admissible if it obeys the conservation of mass, balance of momenta, and conservation of energy. It is called constitutively admissible if it is restricted by the constitutive axioms which we will address later.
• Thermostatic temperature $\tilde{T}$ and thermostatic tensions $\tau_\alpha$ may be defined by:

$$\tilde{T} = \frac{\partial u}{\partial s} |_{\nu, X}, \quad \tau_\alpha = \frac{\partial u}{\partial \nu_\alpha} |_{s, \nu_\alpha, X} \quad \beta \neq \alpha.$$  

(2.5)

It should be emphasized that the temperature $T$ and the thermostatic temperature $\tilde{T}$ are different. But they are the same in a thermodynamically admissible simple thermomechanical process characterized by a set of state variables $(s, \nu_\alpha)$ with $\nu_\alpha$ being independent of $s$.

2.2 Formulations of Thermomechanical Equations with Continuum Mechanics

The thermomechanical behavior of a material is governed by a set of field equations which consist of conservation of mass, balance of momentum, balance of moment of momentum, conservation of energy, Clausius-Duhem inequality, and the constitutive relations (stress-strain relation). A thermodynamically and mechanically admissible process must obey all the field equations.

Basic Axioms of Continuum Mechanics

The basic axioms of continuum mechanics include conservation of mass, balance of momentum, and balance of moment of momentum. The law of conservation of mass states that the total mass of a body is unchanged during its motion; in other words, the initial total mass of the body is the same as the total mass of the body at any other time. The law of balance of momentum dictates that the time rate of change of the momentum is equal to the resultant force $F$ acting on the body. Balance of moment of momentum defines that the time rate of change of moment of momentum is equal to the resultant moment of all forces and couples acting on the body. Assuming
that at time $t$ the body $B$ occupies a volume $V$ and a continuous surface $A$ without singularities or discontinuities, then the laws of conservation of mass, balance of momentum, and balance of moment of momentum may be formulated, respectively, as follows:

$$\frac{D}{Dt} \int_V \rho dV = 0, \quad (2.6)$$

$$\frac{D}{Dt} \int_V \rho v_i dV = \oint_A \sigma_{ji} n_j dA + \int_V \rho f_i dV, \quad (2.7)$$

$$\frac{D}{Dt} \int_V \rho e_{ijk} x_j v_k dV = \oint_A \epsilon_{ijk} x_j \sigma_{mk} n_m dA + \int_V \rho e_{ijk} x_j f_k dV, \quad (2.8)$$

where $\sigma_{mk}$, $f_k$ and $v_k$ are the stress tensor, the body force per unit mass, and the velocity vector, respectively. $n_m$ is the unit vector to the surface vector $dA$ on $A$, and $\epsilon_{ijk}$ is the alternating tensor. The above integral-form equations may be written in the local form:

$$\frac{D}{Dt} + \rho v_i \kappa_i = 0 \quad \text{or} \quad \rho_0/\rho = J = \det(x_{K,K}), \quad (2.9)$$

$$\sigma_{ji,j} + \rho f_i = \rho \frac{Dv_i}{Dt}, \quad (2.10)$$

$$\sigma_{ij} = \sigma_{ji}, \quad (2.11)$$

where $\frac{D}{Dt}$ indicates material differentiation and a subscript following a comma denotes partial differentiation, namely:

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + v_i \frac{\partial}{\partial x_i} \quad v_{i,j} = \frac{\partial v_i}{\partial x_j} \quad x_{K,K} = \frac{\partial x_k}{X_K}. \quad (2.12)$$

Conservation of Energy

The law of conservation of energy states that the time rate of change of the total energy of a system is equal to the sum of the work done by external forces and the heat exchange in the system. Mathematically it may be expressed as:

$$\frac{D}{Dt}(\mathcal{E} + \mathcal{E}) = \mathcal{W} + \mathcal{H}, \quad (2.12)$$
where $\mathcal{K}$, $\mathcal{E}$, $\mathcal{W}$ and $\mathcal{H}$ represent the kinetic energy, the internal energy, the rate of work done by external forces and the heat input to the system, respectively. Specifically, the internal energy $\mathcal{E}$ may be expressed in terms of internal energy density $u$:

$$\mathcal{E} = \int_V \rho u dV ;$$

the total kinetic energy of the body is given by:

$$\mathcal{K} = \frac{1}{2} \int_V \rho v \cdot v dV ;$$

and the rate of work done by the external forces is written as:

$$\mathcal{W} = \int_V \rho f \cdot v dV + \oint_A \sigma(n) \cdot v dA ,$$

where $\sigma(n)$ refers to the stress vector, and $f$ is the body force per unit mass. Let $q$ denote the heat vector per unit area acting on the body surface directed outward, and let $h$ represent the distributed heat sources per unit mass of the body, then the heat input may be written as:

$$\mathcal{H} = \oint_A q \cdot n dA + \int_V \rho h dV .$$

Accordingly, Eq. (2.12) may be rewritten as:

$$\frac{D}{Dt} \int_V \rho (u + \frac{1}{2} v_i v_i ) dv = \oint_A (\sigma_{ij} v_i + q_j) n_j dA + \int_V \rho (f_i v_i + h) dV . \quad (2.13)$$

When the above equation is assumed to be valid for all the body $B$, then it may be localized as:

$$\rho \dot{u} = \sigma_{ij} v_{j,i} + q_{i,i} + \rho h . \quad (2.14)$$

**Principle of Entropy**

The principle of entropy, i.e., the Clausius-Duhem inequality, states that the time rate of change of the total entropy $Q$ is never less than the sum of the entropy influx
$S$ through surface $A$ of the body and the entropy $B$ supplied by the body sources. It is postulated that the above statement is valid for all parts of the body for all independent processes. Mathematically, it may be expressed as:

$$\gamma \equiv \frac{DQ}{Dt} - B - \oint_A S \cdot n dA \geq 0 , \quad (2.15)$$

where $\gamma$ is called *total entropy production*. For a continuous medium, it may be further written as:

$$\frac{D}{Dt} \int_V \rho s dV - \int_V \rho bdV - \oint_V S_k n_k dA \geq 0 , \quad (2.16)$$

where $b$ is the local entropy source per unit mass corresponding to $B$,

$$B \equiv \int_V \rho bdV .$$

When Eq. (2.16) is valid for all parts of the body, it may be localized as:

$$\rho\gamma \equiv \rho \dot{s} - \rho b - div S \geq 0 , \quad (2.17)$$

where $\gamma$ is the local entropy production corresponding to $\gamma$.

Substituting Eq. (2.14) into (2.4)$_2$, one may obtain:

$$\rho T b = \rho T b_1 + \rho \dot{u} - \sigma_{ij} v_{j,i} - q_i ;$$

carrying the above equation and Eq. (2.4) into (2.17), the Clausius-Duhem inequality may then be written as:

$$\rho \gamma \equiv \rho (\dot{s} - \frac{\dot{u}}{T}) + \frac{1}{T} \sigma_{ij} v_{j,i} + \frac{q_i}{T^2} T_i + \rho b_1 - div S_1 \geq 0 . \quad (2.18)$$

For a simple thermomechanical process, i.e., $S_1 = 0$ and $b_1 = 0$, then Eq. (2.18) may further be reduced to:

$$\rho \gamma \equiv \rho (\dot{s} - \frac{\dot{u}}{T}) + \frac{1}{T} \sigma_{ij} v_{j,i} + \frac{q_i}{T^2} T_i \geq 0 . \quad (2.19)$$
If the Helmholtz free energy $\Phi$ is introduced as:

$$\Phi \equiv u - Ts ,$$  \hspace{1cm} (2.20)

by taking derivative of the above equation and substituting into Eq. (2.19), the Clausius-Duhem inequality may be written alternatively as:

$$\rho \gamma \equiv -\frac{\rho}{T}(\dot{\Phi} + s \dot{T}) + \frac{1}{T}\sigma_{ij} v_{i,j} + \frac{1}{T^2} q_i T_{,i} \geq 0.$$  \hspace{1cm} (2.21)

This is a very useful form of the Clausius-Duhem inequality which will be used in formulating the material constitutive relation. The above form may also be written in terms of the 2nd Piola-Kirchhoff stress and Green strain.

### 2.3 Axioms for Formulating Constitutive Equations

The basic laws of continuum mechanics, namely, conservation of mass, balance of momenta, conservation of energy, and principle of entropy, as shown in Eqs. (2.9) - (2.11), and (2.14), lead to totally eight independent equations. The number of unknowns, $\rho$, $v_i$, $\sigma_{ij}$, $q_k$, $u$, $s$ and $T$ adds up to 19. In order to make the problem determinate, eleven additional equations must be constructed. The material characteristics are normally brought into the formulation by such ‘additional’ constitutive equations in terms of the constitutive variables. For a constitutive theory to represent a material adequately, certain physical and mathematical requirements must be satisfied. In this section, eight axioms for such requirements will be introduced to provide general guidance for formulating constitutive equations (Eringen, 1980). Correspondingly, the material constitutive relation will be derived and adapted to these axioms.
Axiom of Causality

The motion of the material points of a body and their temperatures are considered as self-evident observable effects in every thermomechanical behavior of the body; the remaining quantities (other than those derivable from the temperature and motion) that enter the expression of entropy of production are the 'causes' or the dependent constitutive variables.

This axiom provides a rule for selecting the constitutive variables of a material. In the treatment of thermomechanical behavior of a material, once the independent constitutive variables are selected, the remaining quantities in Eq. (2.9) - (2.11), (2.14) and (2.21), which are not derivable from the independent variables, consist of the dependent constitutive variables. In this case, the independent constitutive variables are:

\[ \mathbf{x} = \mathbf{x}(\mathbf{X}, t) \quad T = T(\mathbf{X}, t) \]  

(2.22)

The velocity \( \mathbf{v} \equiv \dot{x} \) is derivable from the motion and \( v_{i,j} \) from \( v \). The density of the body is related to the motion through Eq. (2.9). Accordingly, from Eq. (2.21), the dependent constitutive variables are the stress tensor \( \mathbf{\sigma} \), the heat vector \( \mathbf{q} \), the internal energy density \( u \), the entropy density \( s \), which will be expressed in terms of \( \mathbf{x}(\mathbf{X}, t) \) and \( T(\mathbf{X}, t) \).

Axiom of Determinism

The value of the thermomechanical constitutive functions \( (\mathbf{\sigma}, \mathbf{q}, u, s) \) at a material point \( \mathbf{X} \) of the body \( \mathcal{B} \) at time \( t \) is determined by the history of the motion and the temperature of all material points of \( \mathcal{B} \).
The axiom of determinism excludes the dependence of the material behavior at $X$ on any material points outside the body $B$ and the future events. Namely, the material behavior depends fully upon the past history of the motion and temperature within the body up to the present time. According to the above two axioms, the thermomechanical constitutive equations may be expressed as:

\[
\begin{align*}
\sigma(X, t) &= \hat{\sigma}[x(X', t'), T(X', t'), X, t] \\
q(X, t) &= \hat{q}[x(X', t'), T(X', t'), X, t] \\
u(X, t) &= \hat{u}[x(X', t'), T(X', t'), X, t] \\
s(X, t) &= \hat{s}[x(X', t'), T(X', t'), X, t]
\end{align*}
\]

where $\hat{\sigma}$ and $\hat{q}$ are tensor-valued and vector-valued functionals, respectively. $\hat{u}$ and $\hat{s}$ are scalar-valued functionals which depends on the real functions $x(X', t')$ and $T(X', t')$ with

$X' \in B \quad \text{and} \quad t' \leq t$.

**Axiom of Equipresence**

*At the outset all the constitutive functionals should be expressed in terms of the same list of independent constitutive variables until the contrary is deduced.*

The axiom of equipresence emphasizes the equal importance of all constitutive functionals and tells us not to be prejudiced against a certain class of variables and favor others in formulating the constitutive equations. At the outset, it is assumed to employ the same list of constitutive variables in all of the constitutive functions until the dependence on some of these variables can be eliminated by other axioms or approximations. Accordingly, one can see the same list of variables $x(X', t')$, $T(X', t')$, $X$, and $t$ in all the constitutive equations of Eq. (2.23).
Axiom of Objectivity

The constitutive response functionals must be form-invariant under arbitrary rigid motions of the spatial frame of reference and a constant shift of the origin of time.

This axiom is also called the material frame indifference. Physically it is evident that the material properties cannot depend on the motion of the observer. If two frames \( \bar{x} \) and \( x \) can be made to coincide with a rigid motion and a time shift, then the material frame indifference may be expressed as:

\[
\bar{x}(X, \bar{t}) = R(t)x(X, t) + b(t),
\]

(2.24)

where \( R(t) \) represents a proper orthogonal transformation, and \( \bar{t} \) denotes a translation, namely,

\[
RR^T = R^T R = I, \quad \det R = 1 \quad \bar{t} = t - a,
\]

where \( R^T \) is the transpose of the matrix \( R \), \( I \) is the unit matrix, and \( a \) is a constant. According to this axiom, the constitutive functionals must be form-invariant when \( x \) is replaced by \( \bar{x} \) for all the proper orthogonal transformations \( \{ R(t) \} \), translations \( \{ b(t) \} \) of the spatial frame of reference and all shifts \( \{ a \} \) of the origin of time.

The axiom of objectivity has impact on the constitutive functionals of Eq. (2.23). It should be noted that the following analysis will be restricted to the stress functional \( \sigma \) for the sake of brevity; the analysis for \( q \), \( u \) and \( s \) have similar procedures. According to Eq. (2.24), the stress functional \( \sigma \) may be written as:

\[
\hat{\sigma}[x(X', t'), T(X', t'), X, \bar{t}] = \hat{\sigma}[\bar{x}(X', \bar{t}'), T(X', \bar{t}'), X, \bar{t}],
\]

(2.25)

for all \( R(t') \), \( b(t') \) and \( a \) such that

\[
\bar{x}(X', \bar{t}') = R(t')x(X', t') + b(t') \quad \bar{t}' = t' - a.
\]

(2.26)
In the following, three special transformations of the frame \( \mathbf{x} \) will be considered: rigid translation, shift of time, and rigid rotations. For a rigid translation of spatial frame, taking \( \mathbf{R}(t') = \mathbf{I} \), \( a = 0 \), \( b(t') = -x(X, t') \), according to Eq. (2.26), one may obtain:

\[
\ddot{x}(X', \ddot{t}) = x(X', t') - x(X, t'), \quad \ddot{t} = t'.
\]

Substituting the above equation into Eq. (2.25), one can get

\[
\sigma = \dot{\sigma}[x(X', t') - x(X, t'), T(X', t'), X, t].
\] (2.27)

For a shift of time, selecting \( \mathbf{R}(t') = \mathbf{I} \), \( b(t') = 0 \) and \( a = t' \), from Eq. (2.26) one may obtain \( \ddot{t} = 0 \). Accordingly, \( \dot{\sigma} \) is independent of time \( t \), that is:

\[
\sigma = \dot{\sigma}[x(X', t'), T(X', t'), X].
\]

Introducing a new time variable,

\[
\tau' \equiv t - t' \geq 0, \quad 0 \leq \tau' \leq \infty,
\]

then Eq. (2.27) may be rearranged as:

\[
\sigma = \dot{\sigma}[x(X', t - \tau') - x(X, t - \tau'), T(X', t - \tau'), X].
\] (2.28)

In a rigid rotation of the spatial frame, taking \( b = 0 \), \( a = 0 \) and \( \mathbf{R}(t') \) arbitrary, then we have:

\[
\mathbf{R}(t)\dot{\sigma}[x(X', t - \tau') - x(X, t - \tau'), T(X', t - \tau'), X]\mathbf{R}^T(t) = \\
\dot{\sigma}[x(X', t - \tau') - x(X, t - \tau'), T(X', t - \tau'), X].
\] (2.29)

This is a restriction to make \( \dot{\sigma} \) objective. Accordingly, Eq. (2.28) subjected to the restriction of Eq. (2.29) is the most general constitutive equation for the stress in thermomechanical theories of continuous media.
Axiom of Material Invariance

The constitutive functionals may not change their forms when there exist certain crystallographic symmetries in the material. For example, if there is a reflection of the material frame of reference with respect to the plane $X_3 = 0$, the constitutive functionals will remain form-invariant when the material coordinates $(X_1, X_2, X_3)$ are changed to $(X_1, X_2, -X_3)$. Let $\{G\}$ represent a subgroup of the full group of orthogonal transformations for the material axes and $\{B\}$ denote the translations of these axes. Then the axiom of material invariance may be stated as:

Constitutive equations must be form-invariant with respect to a group of orthogonal transformations $\{G\}$ and translations $\{B\}$ of the material coordinates. These restrictions are the result of symmetry conditions implied by $\{G\}$ and $\{B\}$ in the material frame of reference $X$.

According to this axiom, the response functions will be form-invariant under the linear transformations of the form:

$$ \bar{X} = GX + B, \quad (2.30) $$

where

$$ GG^T = G^TG = I, \quad \det G = 1 $$

for all the members of the group $\{G\}$ and translations $\{B\}$. At $X$, $\{G\}$ represents the geometrical symmetries whereas $\{B\}$ denotes the inhomogeneities of the physical properties of the body. If the symmetry group $\{G\}$ is the full orthogonal group, then the material is called isotropic; if $\{G\}$ is the proper subgroup of the full orthogonal group, the material may be called transversely isotropic; in other cases, the material is called anisotropic. When the constitutive functionals do not depend on translations
\{B\} of the origin of material coordinates, then the material is called \textit{homogeneous}; otherwise, the material is called \textit{inhomogeneous}.

The restriction of the material invariance on the constitutive functionals uses the same techniques as in the axiom of objectivity except that in material invariance, the transformations are employed to material coordinates \(X_K\). Accordingly, the stress functional \(\dot{\sigma}\) must be subjected to the following restriction:

\[
\dot{\sigma}[x(X', t - \tau') - x(X, t - \tau'), T(X', t - \tau'), X] = \dot{\sigma}[x(GX' + B, t - \tau') - x(GX + B, t - \tau'), T(GX' + B, t - \tau'), GX + B].
\]  

(2.31)

**Axiom of Neighborhood**

\textit{The values of the independent constitutive variables at distant material points from \(X\) do not affect appreciably the values of the constitutive dependent variables at \(X\).}

This axiom states that the material response functionals at \(X\) are relatively insensitive to the motion and temperature histories of the distant points from \(X\). Namely in Eq. (2.23), the motions \(x(X', t')\) and temperature \(T(X', t')\) of material points \(X'\) that are far away from the point \(X\) will not contribute appreciably to the functionals \(\sigma, q, u\) and \(s\) at \((X, t)\). There exist two general formulations for this principle. One is the so-called \textit{smooth neighborhood}, the other is the \textit{attenuating neighborhood}.

In the smooth neighborhood case, the functions \(x(X', t')\) and \(T(X', t')\) are supposed to admit Taylor expansions about \(X = X'\) for all \(t' \leq t\), terminating with gradients of order \(M\) and \(N\) respectively, namely

\[
x(X', t') = x(X, t') + (X'_{K_1} - X_{K_1})x_{,K_1}(X, t')
\]  

(2.32)
\[
\begin{align*}
&+ \frac{1}{2!}(X'_{K_1} - X_{K_1})_1(X'_{K_2} - X_{K_2})_2\mathbf{x}_{,K_1K_2}(\mathbf{X},t') + \cdots \\
&+ \frac{1}{M!}(X'_{K_1} - X_{K_1})_1\cdots(X'_{K_M} - X_{K_M})_M\mathbf{x}_{,K_1\cdots K_M}(\mathbf{X},t')
\end{align*}
\]

\[
T(\mathbf{X}',t') = T(\mathbf{X},t') + (X'_{K_1} - X_{K_1})_1T_{,K_1}(\mathbf{X},t')
+ \frac{1}{2!}(X'_{K_1} - X_{K_1})_1(X'_{K_2} - X_{K_2})_2T_{,K_1K_2}(\mathbf{X},t') + \cdots
+ \frac{1}{M!}(X'_{K_1} - X_{K_1})_1\cdots(X'_{K_M} - X_{K_M})_MT_{,K_1\cdots K_M}(\mathbf{X},t')
\]

If the constitutive functionals are sufficiently smooth to approximate \( \mathbf{x}(\mathbf{X}',t') \) and \( T(\mathbf{X}',t') \) by the above expressions, then it is said that the material at \( \mathbf{X} \) satisfies the smooth neighborhood hypothesis. It is obvious that the material behavior of point \( \mathbf{X} \) is not affected by the histories of distant points from \( \mathbf{X} \). If substituting the above Taylor expansions into Eq. (2.28), the stress functional may further be arranged as:

\[
\sigma = \hat{\sigma}[\mathbf{x}_{,K_1}(t - \tau'), \mathbf{x}_{,K_1K_2}(t - \tau'), \cdots, \mathbf{x}_{,K_1\cdots K_M}(t - \tau'), T(t - \tau'), T_{,K_1}(t - \tau'), \cdots, T_{,K_1\cdots K_M}, \mathbf{D}_K, \mathbf{X}],
\]

(2.33)

where \( (K_1, K_2, \cdots, K_i, \cdots) = 1, 2, 3 \). \( \mathbf{D}_K \) is called material descriptors which express the directional dependence of the material properties at a material point \( \mathbf{X} \). Also the following notation is employed in the above equation:

\[
\mathbf{x}_{,K}(t - \tau') \equiv \mathbf{x}_{,K}(\mathbf{X}, t - \tau'), \quad T(t - \tau') \equiv T(\mathbf{X}, t - \tau').
\]

Here it is necessary to introduce a few definitions of material type. A material is said to be of mechanical grade \( M \) and thermal grade \( N \) or \( G(M, N) \), if and only if the constitutive functionals depend on the gradients of \( \mathbf{x}(\mathbf{X}, t - \tau') \) up to order \( M \) and gradients of \( T(\mathbf{X}, t - \tau') \) up to order \( N \). In a special case, thermomechanical materials of grade \( G(1, 1) \) are called simple materials in which the material constitutive behavior only involves the first-order mechanical and thermal gradients. For a simple material, the constitutive equations may have the form:

\[
\sigma = \hat{\sigma}[\mathbf{x}_{,K}(t - \tau'), T(t - \tau'), T_{,K}(t - \tau'), \mathbf{D}_K, \mathbf{X}].
\]

(2.34)
It should be noted that the simple materials contain many subclasses. For example, in the above equation if the dependence on \( \mathbf{x}, \mathbf{K} \) is eliminated, one may get a \textit{rigid material} \( G(0,1) \); when the dependence on \( T, \mathbf{K} \) is dropped one may obtain a \textit{non-heat-conducting material} \( G(1,0) \). It is noteworthy that the materials described by the constitutive equations of Eq. (2.33) are memory-dependent, namely, the constitutive equations at \((\mathbf{X}, t)\) depends on the history prior to time \( t \) of the deformation gradients, the temperature, and the temperature gradients.

If the dependence on the motion and temperatures of neighboring points is not sufficiently smooth to be represented by a few gradients or such gradients do not exist, the attenuating neighborhood hypothesis may be employed. In this case, an \textit{influence function} is introduced to favor the points near \( \mathbf{X} \). It is a decreasing function of \(| \mathbf{X}' - \mathbf{X} |\), and is defined as \( H = \hat{H}(| \mathbf{X}' - \mathbf{X} |) \) with a initial value of \( \hat{H}(0) = 1 \). The influenced distance between \( \mathbf{x}(\mathbf{X}', t') \) and \( \mathbf{x}(\mathbf{X}, t) \) may be defined as:

\[
\| \mathbf{x}(\mathbf{X}', t') - \mathbf{x}(\mathbf{X}, t') \|_H = \left\{ \int_{\mathcal{V}} | \mathbf{x}(\mathbf{X}', t') - \mathbf{x}(\mathbf{X}, t') |^2 \hat{H}(| \mathbf{X}' - \mathbf{X} |) d\mathbf{v}(\mathbf{X}') \right\}^{1/2} .
\]

(2.35)

The axiom of attenuating neighborhood now states (Eringen, 1966) that the response functionals are continuous in a neighborhood of \( \mathbf{X} \) in the function space normed according to the above equations for all \( t' \leq t \). Mathematically, for every \( \varepsilon > 0 \), there exist an \( a > 0 \) and \( b > 0 \) such that:

\[
\| \dot{\mathbf{\sigma}}[\mathbf{x}(\mathbf{X}', t'), T(\mathbf{X}', t'), \mathbf{X}, t] - \dot{\mathbf{\sigma}}[\mathbf{x}(\mathbf{X}, t'), T(\mathbf{X}, t'), \mathbf{X}, t] \| < \varepsilon
\]

(2.36)

whenever

\[
\| \mathbf{x}(\mathbf{X}', t') - \mathbf{x}(\mathbf{X}, t') \|_H < a , \quad \| T(\mathbf{X}', t') - T(\mathbf{X}, t') \|_H < b ,
\]

(2.37)

for all \( t' \leq t \) and \( \mathbf{X}' \) in \( \mathcal{B} \), where \( \| \dot{\mathbf{\sigma}} \| \) denotes \( \sqrt{tr \dot{\mathbf{\sigma}}^2} \). In other words, if the thermo-mechanical histories \( \mathbf{x}(\mathbf{X}', t'), T(\mathbf{X}', t') \) differ very little from their respective values
\( \mathbf{x}(\mathbf{X}, t') \) and \( T(\mathbf{X}, t') \) then the corresponding constitutive functionals (as shown in Eq. (2.23)) such as \( \dot{\mathbf{r}} \) are close together. The existence of the influence function \( H \) is to make the influence of points close to \( \mathbf{X} \) be stronger than that of the far points.

**Axiom of Memory**

*The values of the constitutive variables in the distant past do not affect appreciably the values of the constitutive functionals.*

This axiom is the counterpart of the principle of neighborhood in the time domain. According to this hypothesis, the memory of past motions and temperatures at any material point decays rapidly. There are so-called *smooth memory* and *fading memory* corresponding to the axioms of smooth neighborhood and attenuating neighborhood, respectively.

In the smooth memory case, the thermomechanical histories \( \mathbf{x}(\mathbf{X}', t') \) and \( T(\mathbf{X}', t') \) are expanded in Taylor series at \( t' = t \) for all \( \mathbf{X}' \in \mathcal{B} \) terminating with order \( m \) and \( n \) for a time domain \( \tau_0 < t' \leq t \), respectively, namely:

\[
\mathbf{x}(\mathbf{X}', t') = \mathbf{x}(\mathbf{X}', t) + (t' - t)\dot{\mathbf{x}}(\mathbf{X}', t) + \cdots + \frac{(t' - t)^m}{m!}\mathbf{x}^{(m)}(\mathbf{X}', t)
\]

\[
T(\mathbf{X}', t') = T(\mathbf{X}', t) + (t' - t)\dot{T}(\mathbf{X}', t) + \cdots + \frac{(t' - t)^n}{n!}T^{(n)}(\mathbf{X}', t), \tag{2.38}
\]

where

\[
\mathbf{x}^{(m)} = \frac{\partial^m \mathbf{x}}{\partial t^m}, \quad T^{(n)} = \frac{\partial^n T}{\partial t^n}.
\]

If the dependence of constitutive equations on \( \mathbf{x}(\mathbf{X}', t') \) and \( T(\mathbf{X}', t') \) can be approximated by the above expressions, then it is said that the constitutive functionals satisfy the axiom of smooth memory. Accordingly, the distant histories of motion and
temperature from time \( t \) of any material point \( \mathbf{X}' \) do not affect the material behavior at \( \mathbf{X} \) at time \( t \).

For a material with smooth memory, if the argument functions in Eq. (2.33) possess continuous partial derivatives with respect to \( \tau' \) at \( \tau' = t \), then the constitutive functionals in the range of \( \tau_0 < \tau' < t, -\infty < \tau_0 \), may be represented in terms of various time rates of the gradients of the argument functions:

\[
\sigma = \hat{\sigma}[\mathbf{x}, K_1, \dot{x}, K_1, \ddot{x}, K_1, \ldots; \mathbf{x}, K_1 K_2, \dot{x}, K_1 K_2, \ldots; \ldots; T, \dot{T}, \ddot{T}, \ldots;
T, K_1, \dot{T}, K_1, \ddot{T}, K_1, \ldots; T, K_1 K_2, \dot{T}, K_1 K_2, \ddot{T}, K_1 K_2, \ldots; \ldots; D_K, \mathbf{X}] .
\] (2.39)

It should be noted that \( \hat{\sigma} \) is no longer a functional but an explicit tensor-valued function of the arguments listed. These materials are called rate-type. A simple material of rate-type involving time rates of the deformation gradients up to order \( p \) and temperature and its gradients up to order \( m \) and \( n \) can be written as:

\[
\sigma = \hat{\sigma}[\mathbf{x}, K_1, \dot{x}, K_1, \ddot{x}, K_1, \ldots, \mathbf{x}^{(p)} K_1; T, \dot{T}, \ddot{T}, \ldots, T^{(m)};
T, K_1, \dot{T}, K_1, \ddot{T}, K_1, \ldots, T^{(n)} K_1; D_K, \mathbf{X}] .
\] (2.40)

If the the functions \( \mathbf{x}(\mathbf{X}', \tau') \) and \( T(\mathbf{X}', \tau') \) are not smooth enough to admit Taylor expansions of Eq. (2.38), the fading memory hypothesis (Coleman and Noll, 1961) may be used to smooth out past discontinuities in these functions. Similar to the axiom of attenuating neighborhood, an influence function is introduced to favor the time near the present. Let \( h(t - \tau') \) be such a time influence function in the time domain \(-\infty < \tau' \leq t \), then:

\[
h(0) = 1 \quad \lim_{\tau' \to -\infty} h(t - \tau') = 0 .
\] (2.41)

An example of the time-influence function is:

\[
h(t - \tau') = \exp[-\alpha(t - \tau')] \quad \alpha > 0 .
\]
The normed distance between $\mathbf{z}(\mathbf{X}', t')$ and $\mathbf{z}(\mathbf{X}', t)$ may be defined as:

$$
\|\mathbf{z}(\mathbf{X}', t') - \mathbf{z}(\mathbf{X}', t)\|_h = \left\{ \int_{-\infty}^t \left| \mathbf{z}(\mathbf{X}', t') - \mathbf{z}(\mathbf{X}', t) \right|^2 h(t - t') dt' \right\}^{\frac{1}{2}}.
$$

(2.42)

Then the axiom of fading memory states that the response functionals are continuous from below in a neighborhood of the thermomechanical history at time $t$ in the function space normed according to Eq. (2.42) for all $\mathbf{X}' \in \{B\}$.

**Axiom of Admissibility**

*All constitutive equations must be consistent with the basic principles of continuum mechanics, that is, they are subject to the principles of conservation of mass, balance of momenta, conservation of energy, and the Clausius-Duhem inequality.*

This axiom states that the formulations of the constitutive equations must adhere to the basic laws of motion. The dependence on certain constitutive variables may be eliminated whenever they can be obtained from the equation of motion. For example, the dependence on the history of mass density may be excluded since it can be obtained through the equation of continuity:

$$
\frac{\rho_0}{\rho(\mathbf{X}', t')} = \text{det} \frac{\partial \mathbf{x}_i(\mathbf{X}', t')}{\partial X'_{ik}}.
$$

(2.43)

The Clausius-Duhem inequality can be used to eliminate the dependence on certain constitutive variables. This will be addressed in the following section.

### 2.4 Thermodynamic Restrictions

By using the local Clausius-Duhem inequality, the constitutive functionals studied in the last section can be further restricted. In this section, the thermodynamic restric-
tions on two special classes of thermomechanical materials will be studied to provide general explicit constitutive equations for these materials.

**Thermomechanical Materials Without Memory**

Constitutive functionals of materials subjected to the smooth neighborhood hypothesis may be obtained from Eq. (2.39) by eliminating the time rates. Accordingly, the constitutive functionals of this type will be of the form:

$$\sigma = \dot{\sigma}[x_{K_1}, x_{K_1K_2}, \ldots, x_{K_1K_2\ldots K_M}, T, T_{K_1}, T_{K_1K_2}, \ldots, T_{K_1K_2\ldots K_N}, X]$$ \hspace{1cm} (2.44)

and similar expressions containing the same independent variables exist for $q$, $u$ and $s$. Because of the relation $\Phi = u - Ts$ (Eq. (2.20)), the Helmholtz free energy $\Phi$ is a function of the same arguments of the above equation. The restrictions of the local Clausius-Duhem inequality may be shown by substituting the above arguments into Eq. (2.21):

$$-\frac{\rho}{T}\left(\frac{\partial \Phi}{\partial x_{K_1}} \dot{x}_{K_1} + \frac{\partial \Phi}{\partial x_{K_1K_2}} \dot{x}_{K_1K_2} + \cdots + \frac{\partial \Phi}{\partial x_{K_1K_2\ldots K_M}} \dot{x}_{K_1K_2\ldots K_M} + \frac{\partial \Phi}{\partial T} \dot{T} + \frac{\partial \Phi}{\partial T_{K_1}} \dot{T}_{K_1} + \frac{\partial \Phi}{\partial T_{K_1K_2}} \dot{T}_{K_1K_2} + \cdots + \frac{\partial \Phi}{\partial T_{K_1K_2\ldots K_N}} \dot{T}_{K_1K_2\ldots K_N} + \dot{T}s\right)$$

$$+ \frac{1}{T} \sigma_{kj}X_{K,j} \dot{x}_{K,K} + \frac{1}{T^2} q_{k}X_{K,k}T_{K,k} \geq 0$$ \hspace{1cm} (2.45)

where

$$v_{k,j} = \dot{x}_{k,K}X_{K,j} \quad T_{,k} = T_{,K}X_{K,k}$$

This inequality must be satisfied for all independent thermomechanical processes. Since $\Phi$, $s$, $\sigma$ and $q$ are independent of rates of deformation gradients, temperatures and temperature gradients which are linear with their coefficients, the inequality cannot be valid for arbitrary independent variations of these rates unless all the
coefficients of these rates vanish separately, namely

\[-\rho \frac{\partial \Phi}{\partial x_{k,K}} + \frac{1}{T} \sigma_{kj} X_{K,j} \dot{x}_{k,K} = 0\]

\[\frac{\partial \Phi}{\partial T} + s = 0\]

\[q_k T_{k,k} \geq 0\]

\[\frac{\partial \Phi}{\partial x_{(K_1, K_2 \cdots K_R)}} = 0, \quad (R = 2, 3, \cdots, M)\]

\[\frac{\partial \Phi}{\partial T_{(K_1)}} = \frac{\partial \Phi}{\partial T_{(K_1 K_2 \cdots K_R)}} = 0, \quad (R = 2, 3, \cdots, N)\]

(2.46)

where the parenthesis enclosing the indices in Eq. (2.46)\textsubscript{4} and (2.46)\textsubscript{5} denotes the symmetric part, for example:

\[A_{(K_1 K_2)} \equiv \frac{1}{2!}(A_{K_1 K_2} + A_{K_2 K_1})\]

\[A_{(K_1 K_2 K_3)} \equiv \frac{1}{3!}(A_{K_1 K_2 K_3} + A_{K_3 K_1 K_2} + A_{K_2 K_3 K_1} + A_{K_3 K_2 K_1} + A_{K_1 K_3 K_2} + A_{K_2 K_1 K_3})\]

From the above equation, it is clear from Eq. (2.46)\textsubscript{2} and Eq. (2.46)\textsubscript{4} that the free energy \(\Phi\) is only a function of \(x_{,k}\) and \(T\). The general stress expression can be derived in terms of free energy \(\Phi\):

\[\sigma_{kj} = \rho \frac{\partial \Phi}{\partial x_{(k,K)}} x_{j,K}, \quad (2.47)\]

with the restriction that the antisymmetric part vanishes:

\[\rho \frac{\partial \Phi}{\partial x_{[k,K]}} x_{j,K} = 0. \quad (2.48)\]

where the indices within the square bracket \([\cdot]\) denote the antisymmetric part of the tensor. Accordingly, it is evident that the materials with constitutive equations of the form of Eq. (2.44) must be simple.

**Deformation-Rate-Dependent Thermomechanical Materials**

For deformation-rate-dependent thermomechanical materials subjected to the smooth neighborhood hypothesis, the constitutive equations will be of the form:

\[\sigma = \dot{\sigma} \langle x, K_1, x, K_1 K_2, \cdots, x, K_1 K_2 \cdots K_M; \dot{x}, K_1, \dot{x}, K_1 K_2, \cdots, \rangle\]

78
\[ \dot{x}_{,K_1,K_2,\cdots,K_M}, T, T_{,K_1}, T_{,K_1,K_2}, \cdots, T_{,K_1,K_2,\cdots,K_N}; X \] , 

(2.49)

and similar expressions are available for \( q, u \) (or \( \Phi \)) and \( s \). By substituting the above arguments into Eq. (2.21), one may obtain:

\[
-\frac{\rho}{T} \left( \frac{\partial \Phi}{\partial \dot{x}_{,K_1}} \dot{x}_{,K_1} + \frac{\partial \Phi}{\partial \dot{x}_{,K_1,K_2}} \dot{x}_{,K_1,K_2} + \cdots + \frac{\partial \Phi}{\partial \dot{x}_{,K_1,K_2,\cdots,K_M}} \dot{x}_{,K_1,K_2,\cdots,K_M} + \frac{\partial \Phi}{\partial \dot{x}_{,K_1}} \dot{x}_{,K_1} \right) \\
+ \cdots + \frac{\partial \Phi}{\partial \dot{T}_{,K_1}} \dot{T}_{,K_1} \\
+ \cdots + \frac{\partial \Phi}{\partial T_{,K_1,K_2,\cdots,K_N}} T_{,K_1,K_2,\cdots,K_N} + T_s \right) + \frac{1}{T} \sigma_{k,j} X_{K,j} \dot{x}_{,K} + \frac{1}{T^2} q_k X_{K,k} T, k \geq 0 .
\]

(2.50)

Since the above inequality is linear in the rates of the deformation gradients, temperatures and temperature gradients, by using the same arguments for Eq. (2.45), one may obtain:

\[
\frac{\partial \Phi}{\partial \dot{x}_{,K_1}} = \frac{\partial \Phi}{\partial \dot{x}_{,K_1,K_2}} = \cdots = \frac{\partial \Phi}{\partial \dot{x}_{,K_1,K_2,\cdots,K_M}} = 0 \\
\frac{\partial \Phi}{\partial \dot{T}_{,K_1}} = \frac{\partial \Phi}{\partial \dot{T}_{,K_1,K_2}} = \cdots = \frac{\partial \Phi}{\partial \dot{T}_{,K_1,K_2,\cdots,K_N}} = 0 \\
\frac{\partial \Phi}{\partial T} + s = 0 .
\]

(2.51)

From the above equation, it is clear that

\[ \Phi = \Phi(\dot{x}_{,K_1}, x_{,K_1,K_2}, \cdots, x_{,K_1,K_2,\cdots,K_M}, T) . \]

(2.52)

In general, the stress tensor may be decomposed into an elastic and a dissipative part as:

\[
\sigma = E \sigma(\dot{x}_{,K_1}, x_{,K_1,K_2}, \cdots, x_{,K_1,K_2,\cdots,K_M}, T) + D \sigma(\dot{x}_{,K_1}, x_{,K_1,K_2}, \cdots, \dot{x}_{,K_1,K_2,\cdots,K_M}, T, T_{,K}) ,
\]

(2.53)

where \( E \sigma \) and \( D \sigma \) denote the elastic and dissipative parts of the stress, respectively. Substituting the above equation into Eq. (2.50), one may obtain the following con-
constitutive equations:

\[ E \sigma_{kj} = \rho \frac{\partial \phi}{\partial x_{(k,K)}} T_j, K \]

\[ s = \frac{\partial \Phi}{\partial T} \]

\[ \Phi = \Phi(x_{,K}, T) \]  

\[ \frac{1}{T} D \sigma_{kj} v_{j,k} + \frac{1}{T^2} q_k T_{,k} \geq 0 . \]

(2.54)

It should be noted that \( D \sigma \) and \( q \) can be functions of deformation gradients and their rates although the free energy \( \Phi \) and the elastic part of stress is independent of the rates of the deformation. Eq. (2.54) imposes further restrictions on the forms of \( D \sigma \) and \( q \). It is clear that the rate-dependent thermomechanical materials can be non-simple.

2.5 Generalized Thermodynamic Forces and Fluxes

In this section, the Clausius-Duhem inequality will be further studied using a different approach. It can be considered as the inner product of the generalized forces and the generalized fluxes. To begin with, the Onsager theory of linear nonequilibrium thermodynamics will be introduced, then the Edelen's nonlinear theory will be briefly studied. It is inferred that the Clausius-Duhem inequality can be expressed in terms of a flux potential.

When the thermodynamic state variables or the independent variables in the constitutive equations are selected, then the Clausius-Duhem inequality will be of a form:

\[ \psi(X) = \sum_i J_i(F_j)(F_i - F_i^0) , \]

(2.55)

where the sets \( \{F_i\} \) and \( \{J_i\} \) are called thermodynamic forces and fluxes, respectively.
The set \( \{ F_i^0 \} \) refers to the equilibrium values of \( F_i \) at which the inequality vanishes, namely:

\[
\psi(F_i^0) = 0. \tag{2.56}
\]

An example of the thermodynamic forces and fluxes can be listed from Eq. (2.45), as shown in Table 2.1. It should be noted that the selection of \( \{ J_i \} \) and \( \{ X_i \} \) is arbitrary, but the entropy production must be nonnegative and always be given in the form of Eq. (2.55). In some cases, a few of the thermodynamic fluxes may vanish; then the entropy inequality may not contain corresponding terms.

In Onsager’s theory, the thermodynamic fluxes are assumed to be linear in the thermodynamic forces, namely:

\[
J_i = \sum_j L_{ij}(F_j - F_j^0), \tag{2.57}
\]

where \( L_{ij} \) are the constant phenomenological coefficients which satisfy the symmetric relations or the so-called Onsager reciprocity relations \( L_{ij} = L_{ji} \). Substituting Eq. (2.57) into Eq. (2.55), one may obtain:

\[
\psi = \sum_{i,j} L_{ij}(F_i - F_i^0)(F_j - F_j^0). \tag{2.58}
\]

Accordingly, the second law of thermodynamics is satisfied by the additional requirement that \( \{ L_{ij} \} \) be the coefficient matrix of a nonnegative quadratic form.

The Onsager theory has found many applications in chemical physics (Eringen, 1976). But this theory is only restricted to the linear case and the selection of the forces and fluxes is not unique. Edelen (1972) extended this linear theory to the nonlinear processes. To begin with, a collection of \( M \) functions

\[
J_\alpha = J_\alpha(W_1, \cdots, W_p; F_1, \cdots, F_M), \quad \alpha = 1, 2, \cdots, M \tag{2.59}
\]
Table 2.1: Thermodynamic Force and Corresponding Thermodynamic Flux

<table>
<thead>
<tr>
<th>Thermodynamic Force</th>
<th>Corresponding Thermodynamic Flux $J_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\dot{x}_{k,K}$</td>
<td>$-\frac{\rho}{T} \frac{\partial \Phi}{\partial x_{k,K}} + \frac{1}{T} \sigma_{kj} X_{K,j}$</td>
</tr>
<tr>
<td>$\dot{x}_{K_1K_2}$</td>
<td>$-\frac{\rho}{T} \frac{\partial \Phi}{\partial x_{K_1K_2}}$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$\dot{x}_{K_1K_2\ldots K_M}$</td>
<td>$-\frac{\rho}{T} \frac{\partial \Phi}{\partial x_{K_1K_2\ldots K_M}}$</td>
</tr>
<tr>
<td>$T_{K}$</td>
<td>$\frac{1}{\rho} q_k X_{K,k}$</td>
</tr>
<tr>
<td>$\dot{T}$</td>
<td>$-\frac{\rho}{T} \left( \frac{\partial \Phi}{\partial T} + s \right)$</td>
</tr>
<tr>
<td>$\dot{T}_{K_1}$</td>
<td>$-\frac{\rho}{T} \frac{\partial \Phi}{\partial T_{K_1}}$</td>
</tr>
<tr>
<td>$\dot{T}_{K_1K_2}$</td>
<td>$-\frac{\rho}{T} \frac{\partial \Phi}{\partial T_{K_1K_2}}$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$\dot{T}_{K_1K_2\ldots K_N}$</td>
<td>$-\frac{\rho}{T} \frac{\partial \Phi}{\partial T_{K_1K_2\ldots K_N}}$</td>
</tr>
</tbody>
</table>
of the $M$ thermodynamic forces $F_1, \ldots, F_M$ and $p$ state parameters $W_1, \ldots, W_p$ are defined to constitute an Onsager collection of thermodynamic fluxes if and only if

$$\frac{\partial J_i}{\partial F_j} = \frac{\partial J_j}{\partial F_i}, \quad i, j = 1, 2, \ldots, M$$  \hspace{1cm} (2.60)

holds for all $(W, F)$; $J_i(W_1, \ldots, W_p; 0, \ldots, 0) = 0$ holds for all $\{W\}$; $\psi = \sum_{i=1}^{M} J_i F_i$ is form-invariant under certain transformations; and $\psi = \sum_{i=1}^{M} J_i F_i \geq 0$ holds for all $(W, F)$. According to the above assumptions, it is inferred that there exists a flux potential:

$$\varphi(W, F) = \varphi_0(W) + \int_0^1 g(W, \tau F) \frac{d\tau}{\tau}$$  \hspace{1cm} (2.61)

which makes a collection of $M$ functions $J_i = J_i(W, F)$ constitute an Onsager collection if and only if

$$J_i = \frac{\partial \varphi(W, F)}{\partial F_i},$$  \hspace{1cm} (2.62)

where $g(W, F)$ have the properties that

$$g(W, F) \geq 0, \quad g(W, 0) = 0.$$  \hspace{1cm} (2.63)

Then one may obtain

$$\psi = \sum_{i=1}^{M} J_i F_i = g(W, F) \geq 0.$$  \hspace{1cm} (2.64)

Accordingly, for any function $g(W, F)$ satisfying Eq. (2.63), $\varphi$ and $J_i$ can be determined and the Clausius-Duhem inequality can then be satisfied.

### 2.6 Thermomechanical Approach with Internal Variables

From the above discussion, the conventional thermomechanical approach can be summarized as follows:
• A thermomechanical process must obey all the basic field equations: conservation of mass, balance of momentum, balance of moment of momentum, conservation of energy, and the Clausius-Duhem inequality. In other words, the thermomechanical process must be thermodynamically and mechanically admissible.

• By using the axiom of causality, the assumed independent constitutive variables are the motion and the temperature of the body, as shown in Eq. (2.22).

• From the axiom of determinism, it is postulated that the material behavior depends fully upon the past history of the motion and the temperature within the body up to the present time. The generalized constitutive equations for an admissible process are shown in Eq. (2.23).

• According to the axiom of objectivity, it is shown that the constitutive equations are independent of time. Equation (2.28) subjected to the restriction of Eq. (2.29) is the most general form of the constitutive equation.

• According to the axiom of smooth neighborhood, the generalized form of the constitutive equation can be further arranged into the form of Eq. (2.33) (similar expressions exist for $q$, $u$ and $s$).

• According to the axiom of smooth memory, Eq. (2.33) can be extended into a more specific form of Eq. (2.39).

• From the restriction of the Clausius-Duhem inequality, the explicit constitutive equations can be obtained. For example, the explicit constitutive equations for thermomechanical materials without memory are shown in Eq. (2.46) and (2.47), those for deformation-rate-dependent thermomechanical materials are shown in Eq. (2.54).
The above thermomechanical approach is called conventional because it strictly follows the basic axioms of continuum mechanics, the axioms for formulating constitutive equations, and the restriction of the Clausius-Duhem inequality.

The thermomechanical approach with internal variables was first proposed by Coleman and Gurtin (1967). In this theory, the constitutive functionals are assumed to be a function of the deformation gradient $\mathbf{x},T$, the temperature $T$, the temperature gradient $T_K$, and a set of internal variables $\xi$ or $(\xi_1, \xi_2, \ldots, \xi_N)$:

\[
\begin{align*}
\sigma(X, t) &= \dot{\sigma}[\mathbf{x}, T, T_K, \xi] \\
q(X, t) &= \dot{q}[\mathbf{x}, T, T_K, \xi] \\
\Phi(X, t) &= \dot{\Phi}[\mathbf{x}, T, T_K, \xi] \\
s(X, t) &= \dot{s}[\mathbf{x}, T, T_K, \xi],
\end{align*}
\]  

with an assumption:

\[
\dot{\xi}_i = g_i[\mathbf{x}, T, T_K, \xi].
\]

Substituting the above equation into the Clausius-Duhem inequality Eq. (2.21), one may obtain:

\[
\begin{align*}
-\frac{\rho}{T} \left( \frac{\partial \Phi}{\partial \mathbf{x}_K} \dot{\mathbf{x}}_K + \frac{\partial \Phi}{\partial T} \dot{T}_K + \frac{\partial \Phi}{\partial \xi_1} \dot{\xi}_1 + \cdots \right. \\
+ \left. \frac{\partial \Phi}{\partial \xi_N} \dot{\xi}_N + \dot{\mathbf{s}} \right) + \frac{1}{T} \sigma_{kj} X_{K,j} \dot{x}_{k,K} + \frac{1}{T^2} q_k X_{K,k} T_K \geq 0.
\end{align*}
\]  

Instead of using conventional arguments similar to Eq. (2.45) and (2.46), the term $\frac{\partial \Phi}{\partial \xi_i} \dot{x}_i$ is assumed as the dissipative part of the inequality, namely:

\[
\begin{align*}
-\frac{\rho}{T} \left( \frac{\partial \Phi}{\partial \mathbf{x}_K} + \frac{1}{T} \sigma_{kj} X_{K,j} \right) \dot{x}_{k,K} &= 0 \\
\frac{\partial \Phi}{\partial T} + s &= 0 \\
-\frac{\rho}{T} \left( \frac{\partial \Phi}{\partial \xi_1} \dot{\xi}_1 + \cdots + \frac{\partial \Phi}{\partial \xi_N} \dot{\xi}_N \right) + \frac{1}{T^2} q_k T_k \geq 0 \\
\frac{\partial \Phi}{\partial T_{(K)}} &= 0.
\end{align*}
\]  

Then the general stress expression has the same form as Eq. (2.47). But the Helmholtz free energy is now a function of $\mathbf{x},T$, and $(\xi_1, \xi_2, \ldots, \xi_N)$. The above
arguments are still controversial because according to the conventional thermomechanical approach $\Phi$ is only a function of $x_\kappa$ and $T$, as shown in Eq. (2.46) and (2.54), for the thermomechanical materials.

Tanaka (1982, 1985, 1986) extended the thermomechanical approach with internal variables to the constitutive modeling of shape memory alloys and the transformation superplasticity. The most important modeling step is to select the suitable internal state variables for different constitutive behavior. From a continuum-mechanics viewpoint, a suitable internal state variable should specify the position of a system in the range from the initial state to the final state. For example, a long-range order parameter which represents the extent of phase transition plays a very important role in the process of the second-order phase transition; the extent of reaction in the theory of reacting media and the volume fraction in the theory of mixtures also dominate in their corresponding processes.

By studying the phase transition characteristics of shape memory alloys, Tanaka (1982, 1985) proposed that the martensitic transformation fraction can be an internal variable to describe the thermomechanical behavior of shape memory alloys. In this theory, the thermomechanical behavior of the materials can be described by a set of state variables:

$$\Lambda = (\bar{\epsilon}, T, \xi),$$  \hspace{1cm} (2.68)

where $\bar{\epsilon}$ denotes the Green strain which is only a function of the deformation gradient $x_\kappa$, $T$ represents the temperature, and $\xi$ refers to the volume fraction of martensite. By using the same arguments as those of Coleman and Gurtin, the constitutive
equations for shape memory alloys may be obtained in the form:

$$\tilde{\sigma} = \rho_0 \frac{\partial \Phi}{\partial \tilde{\varepsilon}},$$  \quad (2.69)

where $\tilde{\sigma}$ denotes the 2nd Piola-Kirchhoff stress tensor corresponding to the Green strain $\tilde{\varepsilon}$, $\rho_0$ represents the mass density at the initial configuration. The above equation is only a different form of Eq. (2.47). The difference between Eq. (2.69) and Eq. (2.47) is that in Eq. (2.69) $\Phi$ is assumed to be a function of a set of state variables, as shown in Eq. (2.68). In Eq. (2.47) $\Phi$ is only a function of the deformation gradients $\bm{x}_{,K}$ and the temperature. The time rate of the internal variable $\xi$ is assumed to be governed by the set of state variables chosen, and may be expressed as:

$$\dot{\xi} = g(\tilde{\sigma}, T, \xi),$$  \quad (2.70)

which has the same form as Eq. (1.1). Further, assuming that the transformation phenomenology is fully determined by the set of state variables chosen, then the above equation must have the following form of solution:

$$\xi = g(\tilde{\sigma}, T).$$  \quad (2.71)

Tanaka then proposed a phenomenological formulation for $\xi$ which expresses $\xi$ as an exponential function of $\tilde{\sigma}$ and $T$. Later, Liang and Rogers (1990) proposed a refined exponential model and a cosine model, respectively, to improve the accuracy of Tanaka’s formulation.

In the following chapters of this thesis, the above thermomechanical approach with internal variables, especially Tanaka’s work, will be further extended to the modeling of the two-way shape memory effect and to the fully-coupled constitutive modeling of electromechanical effects in piezoelectricity.
Chapter 3

Modeling of the Two-Way Shape Memory Effect

In this chapter, a further refinement of the transformation kinetics expressions for shape memory alloys are derived after detailed analyses of the microstructural mechanism and training methods of the two-way shape memory effect. A one-dimensional thermomechanical constitutive model for two-way effect which undergoes thermoelastic martensitic transformation and its reverse transformation is presented based on the one-way effect model of Liang and Rogers (1990). This model reflects the essence of SMA in a very simple form, and therefore is easy to use in engineering design. It is revealed that the material parameters required by this model can be calculated or measured using standard material testing apparatus. The computer simulation results are in agreement with the experimental data that has been published in the open literature. It is demonstrated that developed theory can quantitatively predict and describe the behavior of shape memory alloys and be imposed on the two-way shape memory effect.
3.1 Introduction

The shape memory effect (SME) and two-way shape memory effect of Nitinol and other shape memory alloys have been intensively investigated in microstructural and thermomechanical aspects. An increasing number of practical applications have been proposed and demonstrated in recent years (Wayman, 1980; Duerig et al., 1982; Robinson, 1987) for some shape memory alloys (SMAs) such as the Ti-Ni type and copper-based alloys. In order to meet the requirements of engineering design, it is necessary to develop mathematical models to describe the thermomechanical behavior of shape memory alloy. Falk (1980, 1983) proposed a model which explained the thermoelastic transformation in terms of the Helmholtz free energy by applying Landau’s (1969) theory from statistical physics, and developed the stress-strain-temperature curves which are qualitatively in agreement with the experimental results. Müller (1979, 1980) developed a theory based on statistical mechanics, which could be used to describe the pseudoelastic behavior of SMA. Tanaka et al. (1982, 1986, 1988) developed a model based on the concept of the free energy driving force and the first and second laws of thermodynamics. Tanaka’s model is concerned with a one-dimensional metallic material which is undergoing either thermoelastic martensitic transformation or its reverse transformation. The internal state variables which are used to describe the processes of thermoelastic martensitic transformation are strain, temperature and martensitic transformation fraction. This model reflects the thermomechanics of SMA in a simple form. It has been used qualitatively to predict and describe the stress-strain relations and pseudoelastic effects. Based on Tanaka’s work, Liang and Rogers (1990) modified and extended this model to describe quantitatively the behavior of shape memory alloys and proposed a multi-dimensional constitutive model. Although a lot of models have been presented to predict and describe thermoelastic martensitic
transformation and related phenomena, no model has been suggested for the two-way shape memory effect.

The shape memory effect, i.e., one-way effect, refers to the mechanism by which a material which is deformed at a 'low' temperature ($< M_f$) reverts to its original shape by subsequent heating to a relatively 'high' temperature ($> A_f$). The shape recovery occurs during the reverse transformation of the deformed martensite to its parent phase. If the martensite strain limit (typically about 8% in NiTi alloys) is not exceeded, a specimen will completely regain its shape during heating. The one-way effect is irreversible, that is, the shape memory alloy can only remember its original high-temperature shape when the deformed specimen is heated, the specimen cannot recover to its deformed low temperature shape upon subsequent cooling. Contrariwise, the two-way shape memory effect, i.e., the reversible shape memory effect, refers to the mechanism by which the sample can reversibly and repeatedly recall both deformed and undeformed shapes when cooled and heated, respectively.

This chapter will describe a two-way effect constitutive model based on Liang and Roger's (1990) early work with one-way effect materials. Also to be presented is a further refinement of the important thermoelastic transformation kinetics expressions which ultimately determine the quantitative accuracy of the model. The model is utilized to predict the behavior of two-way effect materials and compare it to experimental data that has been published in the open literature.
3.2 The Mechanism and Training of the Two-Way Shape Memory Effect

3.2.1 Mechanism of the Two-Way SME

It is well known that the SME is based on thermoelastic martensitic transformation. In the thermoelastic transformation, the martensite plates form and grow continuously and the martensitic growth rate is proportional to the cooling rate. When cooling ceases, the growth of the martensite plates cease, but they continue to grow on subsequent cooling. When heated, the martensitic plates simply shrink and revert completely to their initial orientation. It seems that the thermoelastic martensitic plates grow and shrink with temperature as deformation of an elastic body under load. In contrast, in nonthermoelastic transformation, such as ferrous alloys or steels, martensitic plates nucleate and grow at a very high speed and do not continue to grow once they are formed. The transformation proceeds through nucleating new martensitic crystals in the remaining parent phase.

Intensive metallurgical studies have revealed the microstructural aspects of shape memory behaviors (Saburi and Wayman, 1979, 1980; Adachi and Perkins, 1984). Studies have shown that in the thermoelastic martensite, the martensitic plate may be internally twin-related or stacking fault-related. As we know, one habit plane and one deformation orientation constitute a variant. There are 24 martensitic variants, i.e., 24 crystallographically equivalent habit planes within one parent phase lattice. In the process of a thermoelastic transformation, an initial parent phase single crystal will transform into martensitic plate groups which are composed of four self-accommodating variants of martensitic plates. In other words, these 24 variants are arranged in 6 plate groups with 4 plate variants per plate group. The deforma-
tion of the four self-accommodating variants within one plate group are effectively canceled and therefore, the overall shape deformation for the group as a whole is approximately zero. When a specimen is deformed below the $M_f$ temperature, it has been demonstrated (Schroeder and Wayman, 1977) that with increasing stress only one of the four self-accommodating variants in a given plate group will begin to grow and dominate the whole specimen while the other three variants are converted into the orientation of the remaining one. The surviving variant is the one of the four variants that has the largest partial shear stress. The shape strain of the particular variant permits maximum deformation of the specimen in the direction of the applied stress. When the deformed specimen is heated, in the region of the reverse transformation between $A_s$ and $A_f$, a process of a single crystal of marten-
site reverting to a single crystal of the parent phase is initiated. Because the crystal structure of the martensite is less symmetric than the parent phase, e.g., the Cu-Zn martensite has an unsymmetric (9R) crystal structure, the reverse transformation is highly restricted in a unique crystallographic manner in which usually only one variant of the parent phase is created. This means that although the parent phase could produce 24 kinds of variants, only a single variant is created when the single crystal region of unsymmetric martensite transforms to parent phase. It has been shown that during the reverse transformation, all of the same variant (orientation) in parallel plates of the parent phase nucleate, grow and coalesce, and revert to the original single crystal and orientation of the parent phase. Perkins (1981) points out that the low elastic strain associated with the crystal structure change is one of the main reasons for the reversibility of the thermoelastic martensite. Another reason for this unique behavior is that groups of mutually accommodating variants are formed which cancel out the volume change due to martensite growth. In addition, twinning or stacking faults within individual plates also accommodate the transformation
strains. Based on the fact that shape memory alloys have anomalously low elastic shear modulus $C' = \frac{1}{2}(C_{11} - C_{12})$ near the transformation temperature, it is conjectured that the characteristics of the elastic shear modulus near the transformation temperature plays a very important role in determining the properties of transformation and that the thermoelastic transformation is driven by a local elastic instability (Castán and Planes, 1987). The above crystallographic discussion is the essence of the shape memory effect.

The two-way shape memory effect has been studied since the late 1960's. The two-way SME was attributed to the stress-induced martensitic transformation due to a special residual stress field (Perkins, 1974; Nagasawa et al., 1974; Wasilewski, 1975; Nishida, 1984). If a specimen is deformed plastically at a temperature below $M_f$, a characteristic internal stress field will be built up. When the deformed specimen is heated above $A_f$ and again cooled to a temperature below $M_s$, the martensitic transformation will initiate as a result of the stored strain energy associated with the residual stress. It is this residual stress field that assists and controls the martensitic transformation to the low-temperature deformed state. The martensitic transformation is controlled through the growth and deformation mode of the intermediate phase and the martensitic variants by the stress field. More specifically, the effect of the residual stress on thermoelastic martensitic transformation is by changing the martensitic transformation to the extent that only a single variant of martensite forms upon cooling. The dislocation induced by the two-way shape memory training inside martensite is the evidence of the residual strain field (Zhu and Yang, 1988). It is investigated, in the CuZnAl alloys, that a great number of dislocations in the parent phase could favor the growth of only one of the 24 possible martensite variants. All of these dislocations in the parent phase have the same Bergers vector, and a single
dislocation in the martensite could dissociate into two dislocations (dislocation pair) during reverse transformation (Ríos-Jara and Gueñín, 1987). It is noted that these dislocations have the second lowest elastic energy in the parent phase of CuZnAl alloys. Therefore once they are formed, these dislocations will exist as stable defects in the parent phase (Ríos-Jara and Gueñín, 1987).

3.2.2 Training of the Two-Way SME

As stated above, the residual strain field within the SMA is the key to the two-way SME. The process of setting up such a characteristic strain field is called the training of the two-way SME. Through training, the sites of the internal stress field which control the martensitic transformation upon cooling are created inside the high-temperature parent phase. These sites are usually a certain pattern of irreversible defects such as dislocations caused by deformation.

The conventional training approach is termed SME cycling (Schroeder and Wayman, 1977; Perkins, 1981). The basic training process is by deforming a specimen below its $M_f$ temperature to produce the preferred variant, then heating it to above its $A_f$ temperature and cooling the specimen to the temperature below $M_f$ again. A schematic illustration of a typical training process and the components of training strain are shown in Figure 3.1. The first step in training the material is to deform the specimen to a certain strain at a temperature below $M_f$ in order to create dislocations for the characteristic stress field. The training strain, $\varepsilon_T$, must be greater than the inelastic strain limit of the material, $\varepsilon_L$, which is a critical value for inducing true plastic deformation. The specimen is unloaded by releasing the stress and then heating the specimen to a temperature above $A_f$ before cooling the specimen to the deformation
Figure 3.1: Schematic Illustration of the Training and the Components of Strain. OA: Loading; AB: Unloading; BC: Heating; CD: Cooling.
temperature again. This completes the first training cycle. The value of $\varepsilon_T$ is usually fixed and used for numerous training cycles, as described below. The training strain, $\varepsilon_T$, is a very important parameter in determining the optimum two-way shape memory effect. If $\varepsilon_T$ is too low, the residual stress fields created by dislocations are too small. Conversely, if the $\varepsilon_T$ is too high, unnecessary internal defects will be induced and will inhibit a high value of two-way SME. For polycrystalline NiTi, the maximum two-way SME can be obtained for training strains between 10% and 20% and training cycles of about 20 (Escher et al., 1990). The stress-strain relation changes with increasing the number of training cycles. The stress-strain-temperature diagrams of the training process are shown in Figure 3.2. The first cycle of the training is a typical deformation curve for SMA deformed below $M_f$. With an increasing number of training cycles, the pseudoelastic deformation disappears and the reorientation of the martensitic variants becomes easier. In addition, transformation strengthening of the parent phase due to the dislocation deposition during the training cycling can improve the reproducibility of the normal shape memory effect (Perkins, 1981). The results of optical microscopy of martensite revealed that non-trained specimens exhibit a microstructure of randomly distributed martensitic variants inside any grain, but the martensitic variants in a trained specimen are oriented parallel in specific directions (Escher et al., 1990).

Another method for training the two-way SME is the so-called stress-induced martensite (SIM) training (Perkins, 1981). A specimen is repeatedly strained at a temperature above $A_f$ (but less than $M_d$) to produce stress-induced martensite and then reverse transformation by release of the applied load via the pseudoelastic effect. After the SIM training, the stress-induced martensite are created in the parent phase and the martensitic variants are reoriented (Zhu and Yang, 1988). The two-way SME
Figure 3.2: Changes in Stress-Strain-Temperature Behavior During the Training Process (after Escher et al.).
may be exhibited on cooling below $M_f$.

From the above discussion, it is obvious that both SME and SIM training must induce a severe plastic deformation which is essential to the two-way effect. Therefore, the shape recovery is incomplete in both the heating and cooling processes for the trained SMA. From this viewpoint, it seems that the two-way effect is obtained at the expense of the SMA's capacity of shape recovery.

Unlike other materials, the micromechanism of plastic deformation of SMA is very complex because of the peculiarity of the thermomechanical behaviors such as pseudoplasticity (one-way effect), two-way effect and pseudoelasticity. Figure 3.3 is schematic illustration of stress-strain curves for an SMA at different deformation temperatures, $T_d$ ($T_1$, $T_2$, $T_3$ temperature levels), relative to martensite transformation range (Perkins, 1981; Sato and Tanaka, 1986). At $A_f < T_1 < M_d$, the strain induced in the loading processes recovers immediately in the unloading process. The behavior by which shape recovery occurs fully on unloading is termed pseudoelastic effect. At $T_3 < M_f$, the strain recovery can only occur upon heating the material to a temperature above $A_f$. This is the so-called shape memory effect. At $A_r < T_2 < A_f$, only part of the reverse transformation occurs during the unloading processes; the other part of the strain recovers in the heating process. This thermomechanical behavior is a transition state between the pseudoelasticity and the shape memory effect, which is called partial pseudoelasticity (Delaey et al., 1974). It is well known that there are components of both pseudoelasticity and SME in shape memory behavior. At lower deformation temperatures, the shape memory effect dominates over pseudoelasticity, i.e., most of the initial strain is recovered by heating instead of unloading. At higher temperatures, the pseudoelastic effect is more pronounced than the SME (Perkins,
Figure 3.3: Schematic Stress-Strain Curves Showing the Pseudoelastic, Partial Pseudoelastic Effect and Pseudoplastic Effect Behaviors at Different Deformation Temperatures $T_T$ Relative to the Martensite Transformation Range:
(a) Pseudoelastic Effect ($A_f < T_T = T_1 < M_d$)
(b) Partial Pseudoelastic Effect ($A_s < T_T = T_2 < A_f$)
(c) Shape Memory Effect ($T_T = T_3 < M_f$)
($\alpha$, Martensite; $\alpha'$, Martensite formed by or altered by stress; $\beta$, parent phase)
1981). In general, the strain, $\varepsilon$, obtained under an external load can be expressed as (Zhang and Hornbogen, 1988):

$$\varepsilon = \varepsilon_e + \varepsilon_{pe} + \varepsilon_{pl} + \varepsilon_{1w} + \varepsilon_{2w}$$  \hspace{1cm} (3.1)

where $\varepsilon_e, \varepsilon_{pe}, \varepsilon_{pl}, \varepsilon_{1w}, \varepsilon_{2w}$ represent the elastic, pseudoelastic, plastic, one-way effect and two-way effect components of the strain, respectively. $\varepsilon_e$ and $\varepsilon_{pe}$ can be removed by unloading. $\varepsilon_{1w}$ and $\varepsilon_{2w}$ can be released by complete transformation into the parent phase. In the SME training process, because the deformation temperature is below $M_f$, there is no pseudoelastic component of the strain, i.e., $\varepsilon_{pe} = 0$. The two-way effect component, $\varepsilon_{2w}$, is included in $\varepsilon_{1w}$, as shown in Figure 3.1. In this way, the SME training strain, $\varepsilon_T$, is given by:

$$\varepsilon_T = \varepsilon_e + \varepsilon_{pl} + \varepsilon_{1w}.$$  \hspace{1cm} (3.2)

### 3.3 Transformation Kinetics of Thermoelastic Martensitic Transformation

Transformation kinetics deal mainly with the relationship between the martensitic fraction and other state variables such as temperature and strain. A typical schematic curve of the martensitic fraction and temperature is shown in Figure 3.4. Intensive studies (Magee, 1968; Koistinen et al., 1959) on transformation kinetics have resulted in an understanding of the relations between the martensitic fraction, $\xi$, and the standard quenching temperature, $T_q$. The calculation of transformation kinetics is based on the concept of the phase equilibrium and the free energy driving force, $\Delta G^{A\rightarrow M}$. The change of martensitic fraction, $d\xi$, can be expressed as (Magee, 1968):

$$d\xi = -\bar{V}(-1 - \xi)K \frac{d\Delta G^{A\rightarrow M}}{dT}$$ \hspace{1cm} (3.3)
Figure 3.4: Typical Thermocycling Curves of Martensitic Fraction vs. Temperature.
where \( \bar{V} \) is the average volume of the pre-formed martensitic plates and \( K \) is a constant. Integrating from \( M_s \), where \( \xi = 0 \), to \( T_q \), assuming \( \bar{V} \) and \( \frac{d\Delta G^{A-M}}{dT} \) are constants yields:

\[
1 - \xi = \exp[\alpha(M_s - T_q)]
\]

(3.4)

where \( \alpha = \bar{V}K\frac{d\Delta G^{A-M}}{dT} \). Equation (3.4) is in agreement with experimental results for higher quenching temperatures which are in the upper part of the temperature interval between \( M_s \) and \( M_f \) for ferrous materials. \( \alpha \) is found to be \(-0.011\) (Koistinen et al., 1959). Equation (3.3) can also be expressed in another form as:

\[
\frac{d\xi}{\xi} = V'K'\frac{d\Delta G^{A-M}}{dT}dT
\]

(3.5)

Integrating from \( M_f \), where \( \xi = 1 \), to any quenching temperature, \( T_q \), yields:

\[
\xi = \exp[\beta(T_q - M_f)]
\]

(3.6)

where \( \beta \) is a constant similar to \( \alpha \). As we know, Equation (3.4) provides a good approximation around \( M_s \) but a poor description around \( M_f \) (Magee, 1968). Contrarily, Equation (3.6) gives a good representation around \( M_f \) but a poor one near \( M_s \). A combination of both equations may be used to describe the relationship of the martensitic fraction and temperature in a piecewise manner. The basic transformation kinetics equations (3.4) and (3.6) have been extended to predict the transformation and reverse transformation of thermoeelastic martensite, respectively, and are in fairly good agreement with the experimental results (Sato and Tanaka, 1988). If \( M_0 = \frac{(M_s + M_f)}{2} \) and \( A_0 = \frac{(A_s + A_f)}{2} \), assuming that both equations yield a \( \xi \) of 0.5 at \( M_0 \) and \( A_0 \), and result in \( \alpha = \beta \), then relations between \( \xi \) and \( T \) for the martensitic transformation may be expressed as:

\[
\begin{cases}
\xi = 1 - \frac{1}{2} \exp[a_M(M_0 - T)^n] & \text{if } M_0 \geq T \geq M_f \\
\xi = \frac{1}{2} \exp[a_M(T - M_0)^n] & \text{if } M_s \geq T \geq M_0
\end{cases}
\]

(3.7)

102
where \( n \) is an introduced parameter which accounts for the term \( \frac{d\Delta G}{dT} \) not being a constant in most cases. For common NiTi SMAs, \( n \) is found to be 0.75-1 (Liang, 1996). For reverse martensitic transformation, the relations of \( \xi \) vs. \( T \) are assumed to be:

\[
\begin{align*}
\xi &= 1 - \frac{1}{2} \exp[a_A(A_0 - T)^n] & A_0 \geq T \geq A_s \\
\xi &= \frac{1}{2} \exp[a_A(T - A_0)^n] & A_f \geq T \geq A_0
\end{align*}
\]  

(3.8)

If \( M_f \) and \( M_s \) are defined to correspond to, respectively, the transformation of 99% and 1% of austenite (a necessary assumption when using experimental parameters), \( a_M \) and \( a_A \) can be derived as

\[
a_M = -3.912 \times 2^n/(M_s - M_f)^n \quad \text{and} \quad a_A = -3.912 \times 2^n/(A_f - A_s)^n.
\]

If the martensitic transformation begins from a state which has mixed austenite and martensite phases denoted by \((\xi_A, T_A)\), or the reverse transformation commences from a mixed-phase state denoted by \((\xi_M, T_M)\), as shown in Figure 3.5, Equations (3.7) and (3.8) can be written as:

\[
\begin{align*}
\xi &= 1 - \frac{1-\xi_A}{2} \exp[a_M(M_0 - T)^n] & M_0 \geq T \geq M_f \\
\xi &= \frac{1-\xi_A}{2} \exp[a_M(T - M_0)^n] & M_s \geq T \geq M_0
\end{align*}
\]  

(3.9)

and

\[
\begin{align*}
\xi &= \frac{\xi_M - \xi_m}{2} \exp[a_A(A_0 - T)^n] & A_0 \geq T \geq A_s \\
\xi &= \frac{\xi_M + \xi_m}{2} \exp[a_A(T - A_0)^n] & A_f \geq T \geq A_0
\end{align*}
\]  

(3.10)

The influence of stress on the phase transformation is approximately linear for many shape memory alloys, as verified by the experimental evidence (Pops, 1970; Goldstein and Kabacoff, 1987) and theoretically indicated by the Clausius-Clapeyron equation under uniaxial stress. A schematic diagram of transition temperatures vs. applied stress is shown in Figure 3.5. It is assumed that the \( M_f \) and \( A_f \) lines are straight and parallel to the \( M_s \) and \( A_s \) lines, respectively. Two more material constants indicating the influence of stress on the transition temperature are assumed as well:

\[
\begin{align*}
C_M &= \tan(\alpha) \\
C_A &= \tan(\beta)
\end{align*}
\]  

(3.11)
Figure 3.5: Influence of Applied Stress on Martensitic Transition Temperature.
From Figure 3.5, the stress range in which stress-induced martensite can be obtained may be expressed as:

\[ C_M(T - M_f) \geq \sigma \geq C_M(T - M_s) \]  \hspace{1cm} (3.12)

whereas the stress range for the reverse transformation may be given as:

\[ C_A(T - A_s) \geq \sigma \geq C_A(T - A_f) \]  \hspace{1cm} (3.13)

Now, the transition temperatures under stress can be expressed in terms of the initial transition temperatures, \( C_M, C_A \) and \( \sigma \).

\[
\begin{align*}
M_s^\sigma &= M_s + \frac{\sigma}{C_M} \\
M_f^\sigma &= M_f + \frac{\sigma}{C_M} \\
M_0^\sigma &= M_0 + \frac{\sigma}{C_M}
\end{align*}
\]  \hspace{1cm} (3.14)

and

\[
\begin{align*}
A_s^\sigma &= A_s + \frac{\sigma}{C_A} \\
A_f^\sigma &= A_f + \frac{\sigma}{C_A} \\
A_0^\sigma &= A_0 + \frac{\sigma}{C_A}
\end{align*}
\]  \hspace{1cm} (3.15)

where \( \sigma \) represents the magnitude of stress which affects the transition temperatures. For the case of uniaxial loading, when the strain is below the inelastic strain limit of the material, \( \epsilon_L \), and no true plastic deformation is induced within the material, \( \sigma \) is simply equal to the external applied stress, i.e., \( \sigma = \sigma_{ap} \). When true plastic deformation is involved or the material is under plastic cyclic deformation, \( \sigma \) denotes the difference between the applied stress, \( \sigma_{ap} \), and the residual stress associated with plastic deformation, \( \sigma_{re} \):

\[ \sigma = \sigma_{ap} - \sigma_{re} \]  \hspace{1cm} (3.16)

Equation (3.16) means that the applied external load and the residual internal stress have different effects on the martensitic transition temperatures. It is well known that applied external stress enhances the martensite transition temperatures (Pops, 1970) because the external stress increases the transformation free energy. It has been
found only in recent years that the martensitic transition temperatures are lowered
and the complete transformation cycle is shifted to lower temperatures with increasing
residual stress (Zhang and Hornbogen, 1987; Goldstein, 1987). The reason that the
residual stress decreases the transition temperatures is attributed to the fact that
the residual stress aids and initiates R structure formation at a lower temperature
(Goldstein, 1987) and impedes the transformation front due to dislocations induced
during plastic deformation (Zhang and Hornbogen, 1988). Figure 3.6 shows the effects
of the plastic deformation on the martensitic transition temperatures. Based on
Equations (3.9) - (3.16), the kinetic expressions for the thermoelastic martensitic
transformation are given by:

\[
\begin{align*}
\xi &= 1 - \frac{1 - \xi_m}{2} \exp[a_M(M_0 + \frac{\sigma}{C_M} - T)^n] \quad M_0^\sigma \geq T \geq M_f^\sigma \\
\xi &= \frac{1 - \xi_m}{2} \exp[a_M(T - \frac{\sigma}{C_M} - M_0)^n] \quad M_f^\sigma \geq T \geq M_0^\sigma,
\end{align*}
\] (3.17)

by:

\[
\begin{align*}
\xi &= \xi_M - \frac{\xi_m}{2} \exp[a_A(A_0 + \frac{\sigma}{C_A} - T)^n] \quad A_0^\sigma \geq T \geq A_f^\sigma \\
\xi &= \frac{\xi_m}{2} \exp[a_A(T - A_0 - \frac{\sigma}{C_A})^n] \quad A_f^\sigma \geq T \geq A_0^\sigma
\end{align*}
\] (3.18)

for the reverse transformation.

Equation (3.17) and (3.18) are the improved kinetic martensitic expressions which
will be used to model the shape memory effect.

### 3.4 Constitutive Model of Two-Way SME Alloys

#### 3.4.1 Constitutive Equations

Tanaka et al. (1982, 1986, 1988) showed that the thermomechanical behavior of SMA
which undergoes stress-induced martensitic transformation or reverse transformation
can be qualitatively described by the following constitutive equation:

\[\dot{\sigma} = D\dot{\varepsilon} + \Theta\dot{T} + \Omega\dot{\xi}\] (3.19)
Figure 3.6: Typical Electrical Resistivity vs. Temperature Curves For a Deformed SMA
where $\tilde{\sigma}$, $\tilde{\varepsilon}$ and $T$ denote the second Piola-Kirchhoff stress tensor, the Green strain tensor and the temperature, respectively, while the superposed dot represents the material time derivative. The coefficients $D$ and $\Theta$ are the Young's modulus and the thermoelastic tensor, respectively. The transformation tensor, $\Omega$, is associated with the volume change due to phase transformation. Equation (3.19) is suitable for a material under one-dimensional tension or compression stress, which is smaller than the yield stress.

Based on Tanaka's early work, Liang and Rogers (1990) extended and modified Equation (3.19) in order to predict and describe quantitatively the behavior of SMA such as free recovery, fully-restrained recovery and controlled recovery. To simplify the expression of the constitutive relations, assuming that $D$, $\Omega$, $\Theta$, $\tilde{\varepsilon}$ and $\xi$ are independent of time and integrating Equation (3.19) from the initial state $(\tilde{\sigma}_0, \tilde{\varepsilon}_0, \xi_0, T_0)$ to a state $(\tilde{\sigma}, \tilde{\varepsilon}, \xi, T)$ with respect to time results in the following constitutive relation:

$$\tilde{\sigma} - \tilde{\sigma}_0 = D(\tilde{\varepsilon} - \tilde{\varepsilon}_0) + \Omega(\xi - \xi_0) + \Theta(T - T_0)$$

(3.20)

where variables with subscript '0' represent the values of initial conditions. Combining Equation (3.29) with (3.17) or (3.18) yields a very good quantitative approximation to describe the one-way behavior of SMA in the course of the stress-induced martensitic transformation or the reverse transformation (Liang, 1990), and is in agreement with experimental data (Dye, 1990). The selection of the martensitic fraction expressions from Equation (3.17) or (3.18) is based on whether it is a martensitic transformation or a reverse transformation.

This model can be extended to accommodate two-way SME. It is implied, from Equation (3.20), that the martensitic transformation or the reverse transformation is
only determined by the final state variables, and is not related to the transformation process. As discussed above, training of the two-way SME introduces a characteristic residual stress field into the shape memory alloy. Although the training process involves plastic cycling of a relatively complex nature, the overall effect on the SMA is to simply introduce a stable and definite residual stress field. Let \( \sigma_{re} \) represent the magnitude of the characteristic residual stress field associated with the plastic deformation. The following equation may then be used to describe the two-way shape memory effect:

\[
\bar{\sigma} - \bar{\sigma}_{re} - \bar{\sigma}_0 = D(\ddot{\epsilon} - \dot{\epsilon}_0) + \Omega(\xi - \xi_0) + T(T - T_0)
\]

(3.21)

where \( \bar{\sigma}_0, \dot{\epsilon}_0 \) and \( \xi_0 \) represent the “initial” second Piola-Kirchhoff stress tensor, the Green strain tensor and martensitic fraction of the trained material, respectively. We should keep in mind that the initial point of the “initial” is the trained state.

In principle, the magnitude of the residual stress field, \( \sigma_{re} \), can be measured by non-destructive methods, such as x-ray diffraction (Noyan and Cohen, 1987), or can be calculated. As we know, the two-way effect and the associated residual stress, \( \sigma_{re} \), are induced by the SME or SIM training which is basically a cyclic plastic deformation process. Therefore, the residual stress, \( \sigma_{re} \), could be represented as (Meyers, 1985):

\[
\sigma_{re} = K'(\epsilon_{pl})^{n'}
\]

(3.22)

where \( K' \) denotes the cyclic strength coefficient and \( n' \) is the cyclic strain-hardening exponent. The value of \( n' \) varies between 0.10 and 0.20, with an average value close to 0.15. From Equation (3.2), \( \epsilon_{pl} \) for SME training can be expressed as:

\[
\epsilon_{pl} = \epsilon_T - \epsilon_{1w} - \epsilon_e
\]

(3.23)

Because the values of \( \epsilon_T, \epsilon_{1w} \) and \( \epsilon_e \) can be obtained from the training, \( \epsilon_{pl} \) can be
calculated. A direct measurement of plastic true strain can be approximated by:

\[ \varepsilon_{pl} = \ln \frac{L}{L_0} \]  

(3.24)

where \( L_0 \) is the length of the specimen before training, \( L \) represents the length of the specimen after training at a temperature above \( A_f \). Substituting Equation (3.23) into (3.22), we have:

\[ \sigma_{re} = K'(\varepsilon_T - \varepsilon_{1w} - \varepsilon_e)^{n'} \]  

(3.25)

Although a lot of data on \( K' \) and \( n' \) are available, it is still not easy to determine accurately the values of \( K' \) and \( n' \) for an SMA with a specific composition and heat treatment condition. It is, therefore, necessary to make some approximations to estimate the residual stress corresponding to the residual strain. It is shown that residual strain increases in the early cycles and has a constant value in the latter cycles (Tobushi et al., 1989), in addition, the training cycles are usually 20 or less (Escher et al., 1990). If we neglect the cyclic strain-hardening effect of the material, residual stress can be approximated by bilinear expressions between stress and strain for simple tension (Johnson and Mellor, 1983):

\[ \sigma_{re} = P\varepsilon_{pl} \]  

(3.26)

where \( P \) represents the plastic modulus. \( P \) can be measured easily by tension testing. Substituting Equation (3.26) into Equation (3.21), one can obtain:

\[ \tilde{\sigma} - \sigma_0 = D(\tilde{\varepsilon} + \frac{P}{D}\varepsilon_{pl} - \varepsilon_0) + \Omega(\xi - \xi_0) + \Theta(T - T_0) \]  

(3.27)

Combination of Equations (3.21) and (3.22) can be done to describe and predict the two-way effect. Equation (3.27) can be used approximately to analyze the behavior of the SME-trained materials.
3.4.2 Determination of the Material Parameters

From the above discussion, it is obvious that the material parameters required to predict the two-way SME are the four transition temperatures ($M_s$, $M_f$, $A_s$ and $A_f$), two stress influence coefficients ($C_M$ and $C_A$), the Young's modulus ($D$), the transformation tensor ($\Omega$), and the thermoelastic tensor ($\Theta$). The four transition temperatures can be accurately measured by using a differential scanning calorimeter (DSC). The phase transformation start temperatures, $M_s$ and $A_s$, are defined as the temperature at which one percent of the parent phase has been transformed. $M_f$ and $A_f$ are defined to correspond to 99% transformed phase. $C_M$ and $C_A$ are measured by examining the mechanical behavior as described by Liang (1990) or by the definition from Equation (3.11) and Figure 3.5. The thermoelastic tensor, $\Theta$, can be measured directly from experiments as described by Dye (1990). The value of $\Theta$ is different for martensite and austenite, but is much smaller than the other material constants. It affects the pre-transformation and post-transformation stage and is negligible during the phase transformation (Liang, 1990). Accordingly, a constant thermoelastic tensor is assumed.

The Young's modulus of an SMA can be determined directly from mechanical testing. The Young's modulus varies significantly with temperature within the phase transformation range. Figure 3.7 shows the relationship between Young's modulus of a Nitinol alloy and temperature during the heating process (Cross et al., 1970). Accordingly, the following equation may be used to calculate the Young's modulus (Liang, 1990):

$$D = D_A + \xi(D_M - D_A)$$  \hspace{1cm} (3.28)

where $D_A$ and $D_M$ denote the Young's modulus of austenite and martensite, re-
spectively. The martensitic fraction $\xi$ is a function of stress and temperature, as is governed by Equation (3.17) or (3.18) for martensitic or reverse transformation, respectively.

The phase transformation tensor, $\Omega$, can be determined by analysis of the thermo-mechanical behavior of a piece of SMA wire undergoing an isothermal mechanical loading and unloading process. In an isothermal process, there is no temperature change; therefore, the $T$ term in Equation (3.19) is zero ($\dot{T} = 0$). Considering a case which has loading and unloading temperatures ranging from $M_s$ to $A_s$ and 100% austenite before loading for an untrained SMA, there is no stress-induced martensite involved ($\xi = 0$) if the stress is less than the elastic stress limit, $\bar{\sigma}_e = C_M(T - M_s)$, as is given in Equation (3.12) (the related elastic strain limit is $\bar{\varepsilon}_e = \bar{\sigma}_e/D$). When the applied stress is greater than the elastic stress limit, $\bar{\sigma}_e$, the martensite will be stress-induced, and Equation (3.20) will become:

$$\bar{\sigma} - \bar{\sigma}_0 = D(\bar{\varepsilon} - \bar{\varepsilon}_0) + \Omega(\xi - \xi_0)$$  \hspace{1cm} (3.29)

with initial conditions, $\bar{\sigma}_0 = \bar{\sigma}_e$, $\bar{\varepsilon}_0 = \bar{\varepsilon}_e$, $\bar{\sigma}_e = D\bar{\varepsilon}_e$ and $\xi_0 = 0$. Substituting the initial conditions into Equation (3.29), one can obtain:

$$\bar{\sigma} = D\bar{\varepsilon} + \Omega\xi$$  \hspace{1cm} (3.30)

From consideration of the unloading process initiating at $\xi = 1$ and $\bar{\sigma} = 0$, the following relation can be determined from Equation (3.30):

$$\Omega = -D\bar{\varepsilon}_L$$  \hspace{1cm} (3.31)

where $\bar{\varepsilon}_L$ represents the recovery strain limit which is considered a temperature-independent material constant, as shown in Figure 3.3(c). The above equation provides an indirect approach to calculated $\Omega$. 

112
Figure 3.7: Young’s Modulus of Nitinol vs. Temperature (after Cross et al., 1970).
3.5 Numerical Illustration

When heating a deformed SMA to a temperature above $A_s$ without applied external load and restrained conditions, the material will begin to recover liberally to its original shape. This process is called free recovery. The thermomechanical behavior for free recovery can be described by means of Equations (3.18) and (3.21). If we neglect the influence of the residual stress on the martensite transition temperatures, Equations (3.18) and (3.21) can be rewritten for $\sigma = 0$ as:

\[
\begin{align*}
\bar{\epsilon} &= \bar{\epsilon}_0 - \left[ \Omega(\xi - \xi_0) + \Theta(T - T_0) + \sigma_{re} \right] / D \\
\xi &= \xi_M - \xi_M^* \exp[a_A(A_0 - T)^n] \quad A_0 \leq T \leq A_s \\
\xi &= \xi_M^* \exp[a_A(T - A_0)^n] \quad A_s \leq T \leq A_0
\end{align*}
\]  

with the initial conditions:

\[
\begin{align*}
\bar{\epsilon}_0 &= \epsilon_{re} \\
T_0 &= A_s \\
\xi_0 &= \xi_M^* = \frac{\epsilon_{re}}{\epsilon_L}
\end{align*}
\]  

where $\epsilon_{re}$ is the residual strain after unloading, as is shown in Figure 3.3(b). Since free recovery only occurs above $A_s$, the austenite start temperature can be assumed as initial condition for the temperature in Equation (3.32).

Table 3.1 lists the material properties used for computer simulation. The comparison between free strain recovery from the recovery strain limit $\epsilon_L$ and a typical two-way strain of about 3.0% is shown in Figure 3.8. The computer simulation curves are in very good agreement with the two-way experimental results (Kovneristyy et al., 1986; Liu and McCormick, 1988). The influence of the residual stress on the free recovery process at a residual strain of 6.0% is demonstrated in Figure 3.9. It is obvious that the direct influence of the residual stress on the free recovery is not significant. This coincides with the conclusion (Liu and McCormick, 1990) that there is no unique relationship between the two-way memory effect and the plastic strain. The training of the two-way effect is through influencing the microstructural aspects (which is
Figure 3.8: Free Recovery of a One-Way Strain and a Two-Way Strain.
Table 3.1: Mechanical Properties for Computer Simulation

<table>
<thead>
<tr>
<th>°C</th>
<th>GPa</th>
<th>MPa/°C</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>( M_f )</td>
<td>( M_s )</td>
<td>( A_s )</td>
<td>( A_f )</td>
</tr>
<tr>
<td>9.0</td>
<td>18.4</td>
<td>34.5</td>
<td>49.0</td>
</tr>
</tbody>
</table>
related to $\Omega$, $\Theta$) of the SMA. It should be noted that the material parameters such as $\Omega$ and $\Theta$, change with temperature (Tanaka, 1988) instead of constants. Accordingly, the above calculation is just an approximation.

### 3.6 Concluding Remarks

A thermomechanical constitutive model for materials which are subjected to thermoelastic martensitic transformation or its reverse transformation is proposed. This model is based upon the concept of free energy and the basic laws of thermodynamics. It reflects the essence of SMA in a very simple form, and therefore is easy to use in engineering design. The relations for the internal variable, $\xi$, is proposed based on transformation kinetics of SMA. It is revealed that the material parameters required by this model can be calculated or measured using standard material testing apparatus. It is demonstrated that this model can be used quantitatively to describe the SME and can be imposed on the two-way SME. It is noted that multidimensional models for the two-way SME still need to be developed and more experimental verifications need to be performed.
Figure 3.9: Influence of the Residual Stress on the Free Recovery.
Chapter 4

A Macroscopic Phenomenological Formulation for Coupled Electromechanical Effects in Piezoelectricity

In this chapter, a phenomenological formulation of polarization reversal of piezoelectric materials is proposed based on the dynamics of domain switching. This formulation provides a method to describe the hysteresis in piezoelectricity as well as in electromagnetics. It is shown that a good approach to describe the nonlinear induced strain-field behavior and electromechanical hysteresis in piezoelectricity is by combining the macroscopic phenomenological aspects with the microscopic material properties. A one-dimensional thermo-electro-mechanical constitutive model for piezoceramics which undergo polarization reversal is presented using a continuum mechanics approach. This model is based on thermodynamic principles and reflects the essence of the electromechanical behavior of piezoceramics in a simple form. It is illustrated that this theory can describe the electromechanical behavior of piezoceramics simply and reasonably well.
4.1 Introduction

The microstructural and electromechanical aspects of piezoelectric ceramics are currently being studied extensively. An increasing number of practical applications have been demonstrated and proposed in recent years for piezoelectric ceramics. With the technological trend for less driving power and miniaturization of devices, piezoelectric materials are playing a more and more important role in intelligent material systems and structures.

Ferroelectric ceramics are composed of many crystallites with randomly-distributed polar axis orientations. When a piece of ferroelectric ceramics is poled by applying a high electric field during processing, the ceramics become piezoelectric and are called piezoceramics. Lead zirconate titanate (PZT) type ceramics, which have numerous applications, are the most researched piezoceramics.

Piezoceramics exhibit a linear relationship between the components of the electric-field vector and the induced strain tensor components. But the linear relationship is only valid for a low field (usually less than 100 V/mm for commercial PZTs). For large fields, the nonlinearity and hysteresis appear during the increase and decrease of the field and influence the accuracy of large induced-strain actuation. The electric field-induced strain hysteresis may result in a multi-valued output problem, namely, for a given input value, the output strain can be one of the many values (Anderson and Crawley, 1989).

In piezoceramics, the relationship between electric displacement $D$ and applied field $E$ forms a hysteresis loop. A typical hysteresis loop for electric displacement $D$ and
Figure 4.1: Schematic Diagram of a Typical Hysteresis Loop for Piezoelectric Ceramics.
applied field $E$ is shown in Fig. 4.1. The important parameters which determine the shape of the hysteresis loop are spontaneous polarization $P_s$, remanent polarization $P_r$, and coercive field $E$. There is general agreement that the hysteresis loop is a consequence of the delayed responses of polarization reversal and domain switching. The hysteresis arises from the energy needed to switch the domain and polarization during each cycle of the field. Although the hysteresis loop appears to be independent of time, it is known that each point on the loop is a function of time because the polarization reversal and domain switching depend not only on material properties and temperatures but also on time rates and the magnitude of the applied field. More specifically, the polarization $P$ is a function of applied field $E$ and time of domain switching $t$: $P = f(E, t)$ (Pulvari and Kuebler, 1958).

The hysteretic behavior of the piezoceramics is a complex issue, sensitive to applied stresses and temperature. Figure 4.2 exhibits the influence of an applied stress field on the hysteresis loop (Arndt, Schmidt and Vogel, 1984). It is shown that a parallel compressive stress decreases the spontaneous and remanent polarization as well as the coercive field, but a compressive stress perpendicular to the field only slightly increases the coercive field. The temperature is also an important factor which influences the hysteresis loop. Figure 4.3 shows that the remanent polarization $P_r$ and coercive field $E_c$ are functions of temperature (Gerthsen and Krüger, 1976). It is shown that the whole hysteresis loop shrinks with increasing temperature. It is thus revealed that the hysteresis loop is a manifestation of the coupled electromechanical behavior of piezoceramics.

In order to exploit piezoceramics for engineering applications and describe the non-linear behavior of the material, it is necessary to formulate a theoretical model which
Figure 4.2: Dielectric Hysteresis Loop for Different Biasing Compressive Stresses (after Arndt, Schmidt and Vogel, 1984): (a) Parallel Compressive Stresses; (b) Compressive Stresses Perpendicular to the Field.
can predict the performance of piezoceramics under various conditions. In this paper, a general phenomenological formulation for polarization reversal dynamics will be proposed. The factors which affect the electromechanical hysteresis will be analyzed. A procedure for formulating a mathematical model for piezoelectricity will be presented based on the detailed analyses of domain switching dynamics and on the material properties. A one-dimensional thermo-electro-mechanical constitutive model will be formulated using a continuum mechanics approach.

4.2 Preliminary Considerations

Much work has been done on single crystals from a microscopic viewpoint. Among these investigations, the thermodynamic theory of single crystals has been intensively researched. Devonshire (1949) put forward a phenomenological theory of ferroelectricity which proposes an expression of the free energy in terms of polarization P, temperature T and stress σ, namely, \( G = f(P, \sigma, T) \). The relationships between the parameters of electric field, polarization, and temperature can be inferred from this theory. Devonshire's work has been extended by numerous studies and is qualitatively in agreement with experimental results. Even in recent years, this theory has been used in thermodynamic studies of \( PbTiO_3 \) single crystals (Haun et al., 1987; Rossetti et al., 1990). This theory has also been extended to model the dielectric properties of \( BaTiO_3 \) polycrystals (Shaikh and Vest, 1989).

Another aspect of the microscopic studies of ferroelectricity is the domain switching process and polarization reversal (Pulvari and Kuebler, 1958). Through either the direct observation of domain wall motion during polarization reversal or measurement of the switching transient under pulsing conditions, the domain switching processes
Figure 4.3: Remanent Polarization $P_r$, Coercive Field $E_c$, and Ratio $P_r^3/E_c$ Shown as Functions of Temperature for PLZT Ceramics (after Gerthsen and Krüger, 1976).
have been intensively studied in ferroelectric materials. It has been shown that nucleation of domains and domain wall motion are the responsible mechanisms for the domain switching process. The domain switching time can be expressed as a function of applied field and the associated material properties. But, the mechanism of the domain configuration and the switching process are still not well understood. On the other hand, most of the practical ferroelectric materials are in polycrystalline ceramic form and more parameters, such as grain boundary and intergranular stresses will complicate the microscopic formulation. In addition, the single crystal properties must be averaged over the crystalline orientations. Therefore, it is difficult at the present time to propose a satisfactory theory of ferroelectrics from a microscopic viewpoint.

It is significant that a macroscopic theory of ferroelectric materials can be formulated in terms of external parameters such as applied electric field, stress state, temperature, dielectric properties and geometric configurations. If such a macroscopic theory can be developed, it will be very easy to use in the prediction of piezoelectric behavior. It is noted that formulation of such a theory with acceptable precision would still be a tremendous task. From a physical viewpoint, Chen and co-workers (Chen and Peercy, 1979; Chen and Montgomery, 1980) proposed a semi-macroscopic phenomenological theory for the hysteresis and butterfly loops in ferroelectric materials based on the relationship between the domain switching under an external electric field and the number of dipoles aligned in the direction of the field. This theory suggested that the mechanical stress $\sigma$ and electric displacement $D$ are functions of the mechanical strain $\epsilon$, the absolute temperature $T$, the external electric field $E$, and the number of aligned dipoles $N$:

$$\begin{cases} 
\sigma = \sigma(\epsilon, T, E, N) \\
D = D(\epsilon, T, E, N).
\end{cases} \quad (4.1)$$
It is noted that the $N$ is the effective number of aligned dipoles and is determined by the projection of the dipole moments to $E$. The number of aligned dipoles $N$ obeys the rate law:

$$\dot{N} = h(\epsilon, T, E, N).$$  \hspace{1cm} (4.2)

This theory represents a new method to characterize the behavior of the ferroelectric materials. It is mentioned that the parameter $N$ has only a pure physical sense and is microscopic. The parameter $N$ cannot match the real polycrystal material where dipoles are arranged in randomly-distributed domains.

Based on the first and second law of thermodynamics, Bassiouny and co-workers (Bassiouny, Ghaleb and Maugin, 1988, 1989) developed a complete phenomenological theory for coupled electromechanical hysteresis effects. Two sets of state variables are used to describe the coupled electromechanical process. One set is the normal state variable such as temperature, the total strain $\epsilon$ and the total polarization. The second type of state variable refers to internal variables such as plastic strain and residual polarization by which actual state depends also on the past history. In Bassiouny’s theory, an introduced electric polarization $P^{\text{int}}$ is chosen as an internal variable. The free energy per unit volume can be expressed as a function of total strain $\epsilon$, reversible polarization $P^r$, internal variable $P^{\text{int}}$ and temperature $T$:

$$\Phi = \Phi(\epsilon, P^r, P^{\text{int}}, T),$$  \hspace{1cm} (4.3)

where $P^r$ is the difference between total electric polarization and residual polarization. $T$ denotes the absolute temperature measured from a reference temperature, $T_0$. After using the Clausius-Duhem inequality and decomposition of the dissipation inequality, a phenomenological theory for coupled electromechanical hysteresis is proposed by analogy of Drucker’s inequality in plasticity and the assumption of loading
functions. This theory provides a very good method of treating electromechanical hysteresis in a continuous media. It seems that the theory is purely a mathematical issue because too many assumptions and approximations are used. For example, the coercive field is assumed to be a constant in the loading functions and the material coefficients are introduced by expanding the free energy function \( \Phi \) in terms of the independent variables.

As discussed before, a combination of the macroscopic phenomenological method and the microscopic physical method will perhaps give a better prediction of the material behavior. In modeling the coupled electromechanical behavior in piezoelectricity, a macroscopic phenomenological approach based on thermodynamics will yield a reasonable macroscopic theory which can describe the material behavior both phenomenologically and physically. This approach will consider the characteristics of the material properties and the domain dynamics.

### 4.3 Domain Dynamics

#### 4.3.1 Concepts of Domain Dynamics

The detailed dynamics of changes in domain configuration under an applied field or stress are complex. When a single crystal or a crystallite is under an applied electric field, the new domains are first nucleated and formed mainly at the surface. They grow forward through the thickness of the crystal, and then expand sideways and coalesce until all the region is occupied by the newly-formed domains. It has been shown that the sideways growth of domains plays an important role in the domain switching process. The sideways wall velocity at low electric field may be expressed
as:

$$v = v_\infty \exp \left( -\frac{\alpha}{E} \right),$$  \hspace{1cm} (4.4)

where $\alpha$ is the activation field for sideways wall motion and $v_\infty$ is a constant. $\alpha$ is a function of specimen thickness $h$ and temperature, the value of $\alpha$ decreases linearly with increasing temperature. The relationship between activation field $\alpha$ and thickness $h$ can be obtained as follows:

$$\alpha = \alpha_0 (1 + \frac{h_0}{h}).$$  \hspace{1cm} (4.5)

The above relations are valid for $h \geq h_0$, where $h_0$ is the surface-layer thickness where the domain switching behavior is independent of the thickness. The coercive field can be explained physically as a field where the reversed domains cover exactly one-half of the volume of the sample. The coercive field can also be expressed as a function of specimen thickness $h$ in a similar way as activation field $\alpha$:

$$E_c = E_\infty (1 + \frac{h_0}{h}),$$  \hspace{1cm} (4.6)

where $E_\infty$ is a constant.

### 4.3.2 Basic Theories of Polarization Reversal Dynamics

Electrical techniques are widely used for studying the polarization reversal dynamics. The Miller-Savage technique of analyzing the electric behavior of ferroelectric materials under square pulses is the most reliable and popular method used to study the domain wall motion and domain switching time $t_s$ (Fatuzzo and Merz, 1967). The important parameters which can be measured experimentally are the switching time, the maximum switching current $i_{\text{max}}$, and the normalized shape of the pulse. The switching time $t_s$ is basically defined as the time necessary to reverse a certain
fraction (e.g., 95%) of the total polarization. In practice, $t_s$ is usually defined as the time necessary for the switching current $i$ to drop to a certain fraction of its maximum value $i_{\text{max}}$. The terms $t_s$ and $i_{\text{max}}$ are related to each other with spontaneous polarization, $P_s$, i.e., $2P_s = i_{\text{max}}t_s f$, where $f$ is a shape factor which depends on the shape of the switching curve (generally ranges from 0.43 to 1.0). Merz (1954) found experimentally that both $t_s$ and $i_{\text{max}}$ in $\text{BaTiO}_3$ change with the electric field $E$ following an exponential law at low electric field ($< 15KV/cm)$:

$$i_{\text{max}} = i_0 \exp \left( -\frac{\alpha}{E} \right)$$

$$t_s = t_0 \exp \left( \frac{\alpha}{E} \right),$$  (4.7)

where $i_0$ and $t_0$ are constants. Stadler (1958) extended Merz’s measurements to the high-field range (10 ~ 100KV/cm) and found that the switching time $t_s$ follows a power law of the type: $t_s = aE^{-b}$, where $a$ and $b$ are constants. This coincides with the fact that for high electric field, only the first power term of the expanded exponential law of Eq. (4.7) remains significant.

Numerous theories for polarization reversal dynamics have been proposed either empirically or theoretically. Based on Merz’s (1954) experimental results, Landauer, Young and Drougard (1956) proposed a polarization theory for $\text{BaTiO}_3$. The switching rate, $dP/dt$, varies with $\exp(-\frac{\alpha}{E(t)})$ as follows:

$$\frac{dP}{dt} = \nu(P_s - P) \exp \left( -\frac{\alpha}{E(t)} \right),$$  (4.8)

where $E(t)$ is the instantaneous field, $\nu$ is a rate constant which is independent of field and polarization, and, $(P_s - P)$ represents the fact that the probability of forming new domains is proportional to the residual volume of the original polarization. If the peak voltage value of the sine wave is twice the coercive voltage or larger, one can take $E(t)$ to be $\dot{E}t$ and integrate Eq. (4.8):

$$\frac{P}{P_s} = 1 - 2 \exp \left[ \frac{\nu\alpha}{E} \tilde{F}(\alpha/E) \right]$$  (4.9)
where \( F'u = \frac{e^{-u}}{u} - \int_{u}^{\infty} \frac{e^{-x}}{x} \, dx \), and can be evaluated by mathematical series. The above equation exhibits the time dependence of the instantaneous polarization during the switching process and defines the shape of the hysteresis loop.

Pulvari and Kuebler (1958) developed a phenomenological theory of polarization reversal in \( BaTiO_3 \) single crystals based on the electric transient response under pulsed voltage. From the experimental results, it has been shown that the current transient can be expressed as \( i = i_{\text{max}} f(\frac{t}{t_{\text{max}}}) \) and \( \int_{0}^{\infty} i \, dt = \text{constant} \), where \( t_{\text{max}} \) is the domain switching time corresponding to \( i_{\text{max}} \). It has also been demonstrated experimentally that \( t_{\text{max}} \) can be represented as a function of electric field: \( \frac{1}{t_{\text{max}}} = \nu E \exp(-\alpha/E) \). \( \alpha \) and \( \nu \) have the same physical meaning as introduced above. A approximation function, \( f \), which satisfies the physical principle of the transient response of the pulsed voltage is assumed as follows:

\[
 f\left(\frac{t}{t_{\text{max}}}\right) = e^{\frac{1}{2}} \left(\frac{t}{t_{\text{max}}}\right) \exp\left[-\frac{1}{2} \left(\frac{t}{t_{\text{max}}}\right)^2\right]. \tag{4.10}
\]

According to the stated assumption and experimental evidence, one can obtain the polarization switching dynamics as follows:

\[
 \frac{P}{P_s} = 1 - 2 \exp\left[-\frac{1}{2} \left(\nu t E e^{-\frac{\alpha}{E}}\right)^2\right]. \tag{4.11}
\]

It has been demonstrated that the dynamic expression of Eq. (4.11) is in agreement with some experimental results.

In addition to Landauer's theory and Pulvari's model, which are empirical and based on experimental results, theoretical models have also been proposed to account for the polarization reversal dynamics. Avrami (1939, 1940, 1941) has demonstrated that the extended area \( A_{ee}(t) \) of a thin ferroelectric material can be related to the actual
area $A_c(t)$ which is covered by the domains at time $t$:

$$A_0 - A_c(t) = A_0 \exp\left[-\left(\frac{A_{\infty}(t)}{A_0}\right)\right], \quad (4.12)$$

where $A_0$ represents the cross-sectional area of the thin ferroelectric surface. $A_{\infty} = \sum_{i=1}^{n} a_i(t)$, and $a_i(t)$ is the area of the $i$th domain at time $t$ while the influence of the other $(n - 1)$ domains to the $i$th domain is neglected. For the sake of simplicity, a fixed number of nucleation sites $N_1$ and the same extended area $a(t)$ for every domain are assumed for a unit surface area. In a thin layer of ferroelectric material, the total extended area per unit area of the electrode (in the same sense as $A_{\infty}(t)$ in Eq. (4.12) if $A_0 = 1$ is assumed) for the positive half-period of a sinusoidal electric field can be expressed as (Janta, 1971):

$$A_{\infty}(t) = N_1 a(t) = 2N_1 h_0 + \frac{v_{\infty}}{\omega} g(\omega t, \frac{E_0}{\alpha})], \quad (4.13)$$

where $g(\omega t, \frac{E_0}{\alpha}) = \int_0^{\infty} \exp\left(-\frac{\alpha}{E_{0 \text{nn}} x}\right) dx$. The velocity of the (apparent) sideways motion of the domain walls is assumed, as in Eq. (4.4). $\omega$ is the angular frequency of the electric field, $v_{\infty}$ represents a constant which has the same meaning as in Eq. (4.4), $h_0$ denotes the initial domain half-width, and $E_0$ is the amplitude of the sinusoidal field. Setting $A_0 = 1$ in Eq. (4.4), one can obtain the expression for mutual overlap of growing domains per unit cross-sectional area as $A_c(t) = 1 - \exp[-A_{\infty}(t)]$. The relative polarization can be expressed in terms of $A_{\infty}(t)$ as:

$$\frac{P(t)}{P_s} = 2A_c(t) - 1 = 1 - 2 \exp[-A_{\infty}(t)]. \quad (4.14)$$

For steady-state hysteresis loops, one may use the symmetry condition $P(\pi) = -P(0)$ for a sinusoidal field to obtain the initial value $h_0$. Combination of Eq. (4.13) and Eq. (4.14) gives:

$$\frac{P(t)}{P_s} = 1 + [1 + \tanh\left(N_1 \frac{v_{\infty}}{\omega} g(\pi, \frac{E_0}{\alpha})\right)] \exp[-2N_1 \frac{v_{\infty}}{\omega} g(\omega t, \frac{E_0}{\alpha})]. \quad (4.15)$$
It should be noted that the above simplified theory is based on the sideways expansion of randomly-distributed 180° domains which are assumed to have the same size and form. Nucleation and forward wall motion are neglected. It has been shown that in the case of cylindrical domains, an expression similar to Eq. (4.15) may be obtained (Jnata, 1971). It should be noted that the above polarization reversal theories are only valid for single crystals. For polycrystals or ceramics, the situation may be quite different.

Polycrystalline ceramics are more important than single crystals because it is easier to maintain quality control with them and they are simple and inexpensive to produce. However, their domain dynamics and domain configuration are more complex than those for single crystals. The grain size and the grain boundary properties are very important factors which influence the domain dynamics. Figure 4.4 represents spontaneous polarization of $BaTiO_3$ as a function of grain size (Shaikh and Vest, 1989). The spontaneous polarization increases with the increase in grain size because as grain size decreases, the grain boundary area (an amorphous state) increases where ferroelectricity does not exist. The spontaneous polarization of large-grain polycrystals of $BaTiO_3$ in an unpolarized state can be calculated from the single-crystal values as follows (Arlt and Sasko, 1980):

$$P_{cer} = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{P_\alpha \cos \theta \cos \phi}{\pi^2} d\theta d\phi,$$

(4.16)

where $P_\alpha = \sqrt{2}P_{sc}$, and $P_{sc}$ is the spontaneous polarization of the single-crystal $BaTiO_3$ along the $<001>$ direction. Integrating the above equation yields $P_{cer} = 0.19P_{sc}$. For poled piezoceramics, the spontaneous polarization is larger than $0.19P_{sc}$.
Figure 4.4: Variation of Spontaneous Polarization of $BaTiO_3$ Ceramics at 70° C with Grain Size (after Shaikh and Vest, 1989).
4.3.3 Phenomenological Formulation of the Polarization Reversal Dynamics for Ceramics

As previously discussed, the mechanism and process of polarization reversal are still not well understood. The basic theories of polarization reversal introduced here are only approximate descriptions of the hysteresis loop. Therefore, it is very difficult to formulate a dynamic theory of polarization reversal by its mechanism or process. Phenomenological formulation may provide an alternative method to describe and predict the hysteresis loop in a simple form. As we know, the instantaneous polarization \( P \) or relative polarization \( P/P_s \) is an exponential function of the applied field. It is postulated that an exponential function of electric field or a combination of exponential functions may provide a better formulation. If a typical hysteresis loop is represented as \( \frac{P}{P_s} = f(E) \), where \( E \) may be a function of time, the basic characteristics of the hysteresis loop may be expressed as at \( E = \pm E_c, \frac{\partial f}{\partial E} = 0 \), because two points \( (\pm E_c, 0) \) are the points of inflection of the hysteresis loop, at \( P > 0, \frac{\partial^2 f}{\partial E^2} < 0 \) and at \( P < 0, \frac{\partial^2 f}{\partial E^2} > 0 \). It is fortunate that the hyperbolic tangent function \( \frac{P}{P_s} = \tanh k(E \pm E_c) \) can match all the characteristics of the hysteresis loop. It is therefore logical that a phenomenological formulation for a steady-state loop may be proposed as follows:

\[
\begin{align*}
\frac{P}{P_s} &= \tanh k(E - E_c) & \frac{dE}{dt} > 0 \\
\frac{P}{P_s} &= \tanh k'(E + E_c) & \frac{dE}{dt} < 0,
\end{align*}
\]

(4.17)

where \( E_c \) represents the amplitude of the coercive field. \( k \) and \( k' \) are newly introduced material parameters. For the case of general ferroelectric ceramics in which the domains are randomly distributed, \( k \) is equal to \( k' \). For piezoelectric ceramics, there is a difference between the polarization along the poling direction and the polarization antiparallel to the poling direction; different \( k \) and \( k' \) values are suitable for this case. It should be noted that the remanent polarization \( P_r \) can be obtained easily from Eq.
(4.17) as \( P_r = P_s \tanh k'E_c \). The virgin curve may be expressed as \( P = P_s \tanh kE \).

The shape of the steady-state hysteresis loop can be totally predicted and described by the two parameters \( k \) and \( E_c \) through Eq. (4.17). Figure 4.5 is an illustration of typical hysteresis loops calculated from Eq. (4.17) for different \( k \) and \( E_c \) values. It is demonstrated that different \( k \) values give different curvatures and the amplitude of the coercive field, \( E_c \), determines the horizontal size of the loop. It is necessary that the material parameter \( k \) and the coercive field \( E_c \) could be determined either theoretically or experimentally. As discussed above, the coercive field \( E_c \) is a function of stress state, temperature, frequency of the electric field, geometrical shape, and material properties such as grain size. Therefore the coercive field \( E_c \) can be represented mathematically in terms of external variables and material properties:

\[
E_c = \hat{E}_c(\sigma, T, \omega, h, d),
\]

where \( h \) represents the thickness of the specimen as in Eq. (4.6) and \( d' \) denotes the diameter of the grain size. It seems very difficult at the current time to evaluate \( E_c \) theoretically. Gerthsen and Krüger (1976) calculated the coercive field assuming that polarization reversal is basically determined by \( 90^\circ \) reorientation. This requires \( P_s^3/E_c \) to be constant (as is shown in Fig. 4.3 for \( P_r^3/E_c \)):

\[
E_c = \beta YQ^2P_s^3,
\]

where \( \beta \) is a correcting coefficient considering the different domain alignments in adjacent grains, \( Y \) is the elastic modulus, \( Q \) represents the electrostriction constant and \( P_s \) is the spontaneous polarization. Because \( P_s \) is also a complicated function of external variables and material properties, \( E_c \) still cannot be determined except when \( P_S \) is known. If it is assumed that the effect of external variable and material
Figure 4.5: Typical Hysteresis Loops Calculated from the Proposed Phenomenological Polarization Reversal Model.
properties to the coercive field $E_c$ are independent of each other, $E_c$ may be obtained analogous to the determination of fatigue strength (Shigley and Mitchell, 1983):

$$E_c = C_\sigma C_T C_\omega C_h C_d E_{co},$$  (4.20)

where $C_\sigma$, $C_T$, $C_\omega$, $C_h$ and $C_d$ represent stress-state factor, temperature factor, frequency factor, geometric factor and grain size factor for coercive field, respectively. $E_{co}$ denotes a coercive field at a reference state. By experimental parametric study, all the coefficients may be determined. The material parameter $k$ may also be expressed as a function of external variables and material properties:

$$k = \dot{k}(\sigma, T, \omega, h, d),$$  (4.21)

and analogous to the arguments for the coercive field, the material parameter $k$ may be obtained as:

$$k = k_\sigma k_T k_\omega k_h k_d k_0,$$  (4.22)

where $k_\sigma$, $k_T$, $k_\omega$, $k_h$ and $k_d$ have similar meanings as the same coefficients of coercive field. $k_0$ denotes a specific $k$ value at a reference state. All the coefficients of $k$ may also be determined by experimental parametric study as in the determination of fatigue strength.

From an engineering viewpoint, the above polarization reversal phenomenology is significant due to its simplicity. This formulation does provide an approach to describe the hysteresis loop in a sense similar to Paris’s law predicting the crack-propagation behavior of fatigue. On one hand, there is no polarization reversal dynamic theory in existence which can describe the hysteresis loop in ceramics because of the complicated essence of the problem and because too many factors influence the hysteresis of ceramics. Although Chen’s semi-microscopic theory and Bassiouny’s thermodynamic theory may reveal some features of domain switching in ceramics, they are
still in the mathematical stage and are too complicated to apply to engineering. On the other hand, the above phenomenological dynamic theory may be used directly in engineering with acceptable precision. With the development of ferroelectric memory appliances, \( k \) and \( E_c \) may be used as the criteria to evaluate whether or not the memory property of ferroelectric ceramics or thin films is sufficient. In these cases, the stress components are zero, and the frequency dependence of \( k \) and \( E_c \) at low-range frequencies is negligible (as shown in Fig. 4.6 for \( E_c \)). Therefore, \( k \) and \( E_c \) are material constants which depend only on material properties. In addition, the above polarization reversal phenomenology may be used to describe the electromagnetic hysteresis loops.

### 4.4 Thermodynamic Formulation for Coupled Electromechanical Effects in Piezoelectricity

As discussed above, formulation of a microscopic theory for piezoceramics is difficult due to the complexity of the problem. From a material viewpoint, the factors which influence the nonlinearity of piezoceramics are composition, ceramic structure, grain size, internal defects and internal stress. From a mechanical aspect, applied stress, temperature and applied electric field may affect the nonlinear behavior of piezoceramics. In Chen’s semi-microscopic theory, the number of aligned dipoles \( N \) cannot describe the polycrystal material in an engineering sense where dipoles are arranged in randomly-distributed domains. Bassiouny’s theory does provide a reasonable approach to predict the electromechanical hysteresis in a continuous media. But this theory is still in a purely mathematical stage and is not easy to use in engineering applications. In this section, a one-dimensional macroscopic phenomenological formulation will be proposed based on the first and second laws of thermodynamics using
Figure 4.6: The Frequency Dependence of Coercive Field $E_c$ for PLZT Ceramics (after Gerthsen and Krüger, 1976).
a continuum mechanics approach. In this study, the material under consideration is polycrystal, and the polarization reversal and electromechanical processes may be considered as processes where both polarization inertia and gradients are neglected.

Analogous to Tanaka's model of SMA materials, which are subjected to thermoelastic martensitic transformation or its reverse transformation (Tanaka, 1986), a one-dimensional ceramic material is considered which is undergoing either polarization or reverse polarization. The energy balance equation and Clausius-Duhem inequality in the current x-coordinate may be expressed as:

$$
\rho \ddot{U} - \sigma L - \rho E \cdot \dot{D} + \frac{\partial q_{\text{sur}}}{\partial x} - \rho q = 0
$$

$$
\rho \dot{S} - \rho \dot{q} + \frac{\partial}{\partial x} \left( q_{\text{sur}} \frac{\partial T}{\partial x} \right) \geq 0, \quad (4.23)
$$

where $\rho$ is the density in the current deformed configuration, $L$ is the deformation velocity, and $\sigma$, $U$ and $q_{\text{sur}}$ denote the Cauchy stress, the internal energy density and the heat flux from the surroundings, respectively. $E$ refers to the local electric field at a given point within the material and $D$ is the electric displacement. The superposed dot denotes a material derivative. $T$, $q$ and $S$ represent the temperature, internal heat source, and entropy density, respectively. According to Tanaka's definition, the above energy balance equation and Clausius-Duhem inequality may be written in the original configuration state, $X$, as:

$$
\rho_0 \ddot{U} - \tilde{\sigma} \tilde{\varepsilon} - \rho_0 E \cdot \dot{D} + f^{-1} \frac{\rho_0}{\rho} \frac{\partial q_{\text{sur}}}{\partial X} - \rho_0 \xi = 0
$$

$$
\rho_0 \dot{S} - \rho_0 \dot{q} + f^{-1} \frac{\rho_0}{\rho} \frac{\partial q_{\text{sur}}}{\partial X} - f^{-1} \frac{\rho_0}{\rho} \frac{\partial T}{\partial X} \frac{\partial T}{\partial X} \geq 0, \quad (4.24)
$$

where $\rho_0$ is the mass density with reference to the original configuration, $f$ is the deformation gradient, $\tilde{\sigma} = \frac{\rho_0}{\rho} \frac{\sigma}{f^2}$ is defined as the second Piola-Kirchhoff stress, $\tilde{\varepsilon} = \frac{f^2 - 1}{2}$ is the Green strain.
The thermomechanical state of a ceramic material at a given point at time \( t \) is completely determined by a set of state variables. It is considered (Bassiouny, Ghaleb and Maugin, 1988) that the temperature \( T \), the total strain \( \varepsilon \) and the instantaneous polarization \( P \) define the coupled electromechanical behavior for a reversible process. In order to use the proposed phenomenological polarization reversal formulation of Eq. (4.17), we assume that the set of state variables which describe the coupled electromechanical process of piezoceramics are the total strain \( \varepsilon \), the temperature \( T \), the relative polarization \( \frac{P}{P_s} \) (which will be denoted by \( \xi \)), and the electric field \( E \). It should be noted that except at the external surface of the specimen, the external electric field does not act on a given point within the material. It may be considered as a pseudo-state variable whose effect on a given point within the material is through a so-called local electric field \( \mathbf{E}_L \). The relative polarization \( \xi \) determines the polarization state of a given point. Such a state variable is assumed to reveal the implied relationship between the stress state and an applied electric field \( \mathbf{E} \) in an explicit form. The assumed general state variable may be expressed as:

\[
\Lambda \equiv (\varepsilon, T, \xi, E).
\] (4.25)

The energy function corresponding to this case, which is the driving force of the electromechanical process, is a function of the state variable \( \Lambda \) and is given by:

\[
\Phi(\Lambda) = U - TS - \mathbf{D} \cdot \mathbf{E}
\] (4.26)

Considering Eq. (4.25) and taking derivatives of Eq. (4.26) yields:

\[
\dot{\Phi} = \frac{\partial \Phi}{\partial \varepsilon} \dot{\varepsilon} + \frac{\partial \Phi}{\partial T} \dot{T} + \frac{\partial \Phi}{\partial \xi} \dot{\xi} + \frac{\partial \Phi}{\partial E} \dot{E} = \mathbf{U} - \mathbf{T} \mathbf{S} - \mathbf{D} \cdot \mathbf{E} - \dot{\mathbf{D}} \cdot \mathbf{E}.
\] (4.27)

Substituting Eq. (4.27) and the energy balance equation (Eq. (4.24)\(_1\)) into the Clausius-Duhem inequality (Eq.(4.24)\(_2\)) and expressing every term in the original
configuration $X$ gives:

$$
\left( \frac{\sigma}{\rho_0} - \frac{\partial \Phi}{\partial \varepsilon} \right) \dot{\varepsilon} - (S + \frac{\partial \Phi}{\partial T}) \dot{T} - (D + \frac{\partial \Phi}{\partial E}) \dot{E} - \frac{\partial \Phi}{\partial \xi} \dot{\xi} - \frac{1}{\rho T} f^{-1} q_{\text{sur}} \frac{\partial T}{\partial X} \geq 0.
$$

(4.28)

From the thermodynamics of continuous media, the coefficients of $\dot{\varepsilon}$ and $\dot{T}$ should vanish. Therefore the inequality of Eq. (4.28) yields the electromechanical constitutive equation:

$$
\sigma = \rho_0 \frac{\partial \Phi}{\partial \varepsilon} = \sigma(\varepsilon, T, \xi, E) \\
D = -\frac{\partial \Phi}{\partial E} = \dot{D}(\varepsilon, T, \xi, E).
$$

(4.29)

Taking the derivative of the above equation, the rate form of the electromechanical constitutive equation may be given as:

$$
\dot{\sigma} = Y \dot{\varepsilon} + \Theta \dot{T} + \Omega \dot{\xi} + \Gamma \dot{E} \\
\dot{D} = \Delta \dot{\varepsilon} + \Pi \dot{T} + \Xi \dot{\xi} + \Upsilon \dot{E},
$$

(4.30)

where $Y$, $\Theta$, $\Omega$, $\Gamma$, $\Pi$, $\Xi$ and $\Upsilon$ represent Young's modulus, thermoelastic tensor, polarization reversal tensor, electromechanical tensor, thermoelectrical tensor, electrical-polarization tensor and electrical tensor, respectively. These material properties so derived may be expressed in terms of the corresponding energy function:

$$
Y = \rho_0 \frac{\partial^2 \Phi}{\partial \varepsilon^2} \\
\Theta = \rho_0 \frac{\partial^2 \Phi}{\partial \varepsilon \partial T} \\
\Omega = \rho_0 \frac{\partial^2 \Phi}{\partial \varepsilon \partial \xi} \\
\Gamma = \rho_0 \frac{\partial^2 \Phi}{\partial \varepsilon \partial E} \\
\Delta = -\Pi / \rho_0 \\
\Xi = -\frac{\partial^2 \Phi}{\partial E^2}.
$$

(4.31)

Eq. (4.30) and (4.31) provide a simple coupled electromechanical model which describe the stress-strain relation in an explicit form. However, a detailed experimental procedure is needed to determine the material constants which are defined in Eq. (4.31). It should be noted that the `one-dimensional' is only required for the induced strain and the external applied stress but not for the applied field.
4.5 A Special Case

In an isothermal electromechanical process, where there is no temperature change, the
$T$ term in Eq. (4.30)$_1$ will vanish ($\dot{\bar{T}} = 0$). Therefore, Eq. (4.30)$_1$ may be rewritten
as:

$$\dot{\bar{\sigma}} = Y \dot{\bar{\varepsilon}} + \Omega \dot{\bar{\xi}} + \Gamma \dot{\bar{E}}.$$  \hfill (4.32)

For a low-frequency case where the stress-strain process may be considered as quasi-
static, the stress $\bar{\sigma}$, strain $\bar{\varepsilon}$ and relative polarization $\bar{\xi}$ can be envisioned as functionals
of the electric field $\bar{E}$ where the independent variable is time $t$. Integrating every term
of the above equation over $\bar{E}$ gives:

$$\bar{\sigma} - \bar{\sigma}_0 = Y(\bar{\varepsilon} - \bar{\varepsilon}_0) + \Omega(\bar{\xi} - \bar{\xi}_0) + \Gamma(\bar{E} - \bar{E}_0)$$  \hfill (4.33)

where $\bar{\sigma}_0$, $\bar{\varepsilon}_0$, $\bar{\xi}_0$ and $\bar{E}_0$ represent initial stress, strain, relative polarization and initial
electric field, respectively. The physical meaning of the above equation is very clear.
The first two terms represent a basic mechanical stress-strain relation. The third term
denotes the contribution of the polarization reversal to the stress-strain relation of the
material. The fourth term gives the linear reverse piezoelectric effect. For a stress-
free case with zero initial conditions, if the proposed phenomenological formulation
of polarization reversal for a virgin curve ($\xi = \tanh kE$) is used, the above equation
may be simplified as:

$$\bar{\varepsilon} = -\frac{\Omega}{Y} \tanh kE - \frac{\Gamma}{Y} E.$$  \hfill (4.34)

The above equation may provide a nonlinear relationship between the induced strain
and the applied field. Due to lack of experimental data, the material constants re-
quired by the model are still not available and we cannot verify the above model
at the current time. However, the validation of our model may be illustrated indi-
rectly. As shown in Figure 4.7, the nonlinear field-strain experimental data, $E_3$ vs. $\varepsilon_1$
(Anderson and Crawley, 1989), can be matched 'perfectly' with the above equation through curve-fit techniques. The expression for the theoretical curve in Fig. 8 is 

\[ \varepsilon = 0.665E - 206 \tanh 0.002E \]  

(E is in V/mm and \( \varepsilon \) is in microstrain). In principle, any experimental curve can be matched by regression analysis but may not match the physical principle. In our case, the theoretical expression matches both the experimental data and the physical principle of the nonlinearity.

### 4.6 Concluding Remarks

A general phenomenological formulation of polarization reversal is proposed in a simple form. This formulation provides a very good method to describe the hysteresis in piezoelectricity as well as in electromagnetics. With the development of ferroelectric memory devices, the material parameters \( k \) and \( E_c \) may be used as criteria to evaluate the memory quality of the materials. A good approach to describe the nonlinear induced strain-field behavior and electromechanical hysteresis in piezoelectricity is by combining a macroscopic phenomenological method with microscopic material properties. A one-dimensional thermo-electro-mechanical constitutive model for piezoceramics which undergo polarization reversal is presented using a continuum mechanics approach. This model is based on the basic laws of thermodynamics and reflects the essence of the electromechanical behavior of piezoceramics in a simple form. It is significant that the proposed model provides an explicit form of the stress-strain relation which may facilitate engineering design of piezoelectric actuators. It is illustrated that this theory can describe the electromechanical behavior of piezoceramics simply and reasonably well. It should be noted that material constants for the proposed model are not yet determined due to lack of experimental data at the present time.
Figure 4.7: Comparison of Experimental Data (after Anderson and Crawley, 1989) and Theoretical Prediction.
Chapter 5

Conclusions and Recommendations

5.1 Conclusions

The thermomechanical approach with internal variables, which is a combination of thermodynamic principles and continuum mechanics has been thoroughly analyzed. A one-dimensional constitutive model for two-way shape memory effect and a one-dimensional constitutive model for piezoceramics have been developed, respectively, based on the approach introduced in this study.

- The thermomechanical approach with internal variables is analyzed in detail to provide a solid background theory for modeling the two-way shape memory effects and piezoelectricity. This approach is based on the field equations of motion, the axioms for formulating constitutive equations, and the restrictions of the Clausius-Duhem inequality. Therefore, it reflects the physical essence of constitutive behavior of materials.

- A one-dimensional constitutive theory for two-way shape memory effect is developed based on Tanaka’s and Liang’s early work.
1. The mechanism of the two-way shape memory effect is analyzed and a residual stress $\sigma_{rr}$ is introduced as a controlling parameter for the two-way shape memory effect.

2. A further refinement of the transformation kinetics expressions for two-way shape memory alloy which undergoes thermoelastic martensitic transformation and its reverse transformation is derived.

3. It is demonstrated that the material parameters required by this model can be calculated or measured using a standard material testing apparatus.

4. A numerical study is conducted and the effectiveness of this model is verified.

- A fully-coupled one-dimensional constitutive model for piezoceramics is developed by using the thermomechanical approach with internal variables.

1. It is significant that a new internal variable, the polarization fraction, is introduced to relate the macroscopic thermo-electro-mechanical behavior of a piezoceramic with its micro-properties.

2. A phenomenological formulation of polarization reversal for piezoelectric material is proposed based on the dynamics of domain switching. This formulation provides a good method to describe the hysteresis in piezoelectricity as well as in electromagnetics.

3. A fully-coupled thermo-electro-mechanical model for piezoceramics is proposed by using the restriction of the Clausius-Duhem inequality.

4. A numerical example shows that the developed theory can describe and predict the electromechanical behavior of piezoceramics simply and reasonably well.
5.2 Recommendations

In this thesis, based on the study of the thermomechanical approach, a one-dimensional constitutive model for two-way shape memory effects and a fully-coupled one-dimensional thermo-electro-mechanical constitutive model for piezoceramics have been proposed. It has been shown that both models can predict the constitutive behavior of the materials very well. However, much work is needed to explore the characteristics and limitations of the developed models prior to their uses in engineering design. The following are a few suggestions that are proposed for further study of the two constitutive models.

For the constitutive model of the two-way shape memory effect, the following issues need to be researched further:

- The second-rank tensors $D$, $\Omega$ and $\Theta$ are assumed to be constant in the linear model of the two-way shape memory effect, but further research has shown that these tensors are functions of the martensitic fraction (Brinson, 1992). Accordingly, much more work is needed to determine the amplitudes of these tensors and how to incorporate these tensors into the constitutive model.

- Further experimentation of the parameter $\sigma_{re}$ is necessary. The effect of this parameter on the kinetics of the thermoelastic martensitic transformation also needs to be studied further.

For the fully-coupled thermo-electro-mechanical model for piezoceramics, the following topics require further consideration:

- The approach and experimentation on how to determine the second-rank ten-
sors: $Y, \Theta, \Omega, \Gamma, \Pi, \Xi$ and $\Upsilon$.

- The different effects of the residual polarization and reversible polarization on the constitutive behavior of piezoceramics need to be clarified.

- In the developed model, all the second-rank tensors $Y, \Theta, \Omega, \Gamma, \Pi, \Xi$ and $\Upsilon$ are assumed to be constant. This assumption needs to be experimentally verified. If they are functions of the set of state variables chosen, then these functions need to be determined.

- The feasibility of modeling three-dimensional piezoelectricity should be conducted.
References


Delaey, L., Krishnan, R.V., Tas, H. and Warlimont, H., 1974, "Thermoelasticity, Pseudoelasticity and Memory Effects Associated With the Martensitic Transforma-


Dye, T., 1990, *An Experimental Verification of a One-Dimensional Constitutive Model of Shape Memory Alloys*, Master’s Thesis, Department of Mechanical Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, August.


Eringen, A.C., 1966, “A Unified Theory of Thermomechanical Materials”, *Internat-


Liu, Y. and McCormick, P.G., 1990, “Factors Influencing the Development of Two-

157


Vita

Xiaodong Zhang was born in Zhouzi County, Shaanxi Province, China, on June 10, 1963. He grew up in his hometown where he graduated from the Zhouzi High School in 1980. After high school he attended Xi’an Jiaotong University, and majored in Casting. In Xi’an Jiaotong University, he received a Bachelor’s degree from the Mechanical Engineering Department in 1984 and a Master’s degree from the Materials Science and Engineering Department in 1987. Then he went to Tianjin University and taught physical metallurgy and heat treatment principles for two and one-half years. He came to Virginia Polytechnic Institute and State University in January 1990. In December 1992 he graduated with a Master’s degree in Mechanical Engineering. His thesis study involved a thermomechanical approach to the modeling of materials with internal variables.

Xiaodong Zhang